Theor Appl Genet (2002) 104:894–900
DOI 10.1007/s00122-001-0798-8

**F. Anthony · M. C. Combes · C. Astorga · B. Bertrand**
**G. Graziosi · P. Lashermes**

# The origin of cultivated *Coffea arabica* L. varieties revealed by AFLP and SSR markers

Received: 8 January 2001 / Accepted: 20 June 2001 / Published online: 8 February 2002


**Abstract** Molecular markers were used to assess polymorphism between and within the genetic bases of coffee (i.e. Typica and Bourbon) spread from Yemen since the early 18th century that have given rise to most arabica cultivars grown world-wide. Eleven *Coffea arabica* accessions derived from the disseminated bases were evaluated by amplified fragment length polymorphism (AFLP) using 37 primer combinations and simple-sequence repeats (SSRs) produced by six microsatellites. Four cultivars growing in Yemen and 11 subspontaneous accessions collected in the primary centre of diversity of the species were included in the study in order to define their relationship with the accessions derived from the genetic bases of cultivars. One hundred and seven AFLP markers were used to calculate genetic distances and construct a dendrogram. The accessions derived from the disseminated bases were grouped separately, according to their genetic origin, and were distinguished from the subspontaneous accessions. The Yemen cultivars were classified with the Typica-derived accessions. Except for one AFLP marker, all AFLP and SSR markers present in the cultivated accessions were also detected in the subspontaneous accessions. Polymorphism among the subspontaneous accessions was much higher than among the cultivated accessions. It was very low within the genetic bases, confirming the historical documentation on their dissemination. The results enabled a discussion of the genetic diversity reductions that successively occurred during the dissemination of *C. arabica* from its primary centre of diversity.

Communicated by H.C. Becker

F. Anthony (✉) · C. Astorga · B. Bertand
CATIE, Ap. 59, 7170 Turrialba, Costa Rica
e-mail: fanthony@catie.ac.cr
Fax: +506-556-1533

M.C. Combes · P. Lashermes
IRD, BP 5045, 34032 Montpellier Cedex 1, France

G. Graziosi
Università deli Studi di Trieste, I-34148 Trieste, Italy

**Keywords** *Coffea arabica* · Coffee · AFLP markers · SSR markers · Domestication

## Introduction

Commercial coffee production relies mainly on two species, *Coffea arabica* L. (70%) and *C. canephora* Pierre (30%). The better cup quality is associated with the allotetraploid (2n=4x=44) species *C. arabica*, which has its primary centre of diversity in the southwestern Ethiopia highlands (Sylvain 1955), the Boma Plateau of Sudan (Thomas 1942) and Mount Marsabit of Kenya (Anthony et al. 1987). The first written mention of coffee was from Razes, a 10th century Arabian physician (Smith 1985), but coffee exploitation may have begun at least 1,500 years ago in southwestern Ethiopia (Lejeune 1958). Arabica coffee plants could have been introduced to Yemen as early as 575 AD (Wellman 1961) or just three to four centuries ago (Eskes 1989). Two genetic bases spread from Yemen, and these have given rise to most of the commercial Arabica cultivars grown world-wide. They have been described as two distinct botanical varieties: *C. arabica* var. arabica (usually called *C. arabica* var. typica Cramer) and *C. arabica* var. *bourbon* (B. Rodr.) Choussy (Krug et al. 1939; Carvalho et al. 1969). Historical data indicate that the Typica genetic base originated from a single plant from Indonesia that was subsequently cultivated in the Amsterdam botanical gardens in the early 18th century (Chevalier and Dagron 1928; Carvalho 1946). The Bourbon genetic base originated from coffee trees introduced from Mocha (Yemen) to the Bourbon Island (now Réunion) in 1715 and 1718 (Haarer 1956).

Genetic diversity conserved in coffee field genebanks has been assessed using agro-morphological characters and, more recently, by DNA-based genetic markers. The results have enabled the characterisation of the two genetic bases of the cultivated material. Bourbon-derived cultivars are characterised by a more compact and upright growth habit, higher yield and better cup quality than the Typica-derived cultivars (Carvalho et al. 1969). Random

amplified polymorphic DNA (RAPD) markers have shown a distinction between cultivated and wild material (Orozco-Castillo et al. 1994; Lashermes et al. 1996; Anthony et al. 2001). Wild and cultivated coffee from Ethiopia was classified according to their geographic origin (i.e. southwestern and south/southeastern Ethiopia) using agro-morphological observations (Montagnon and Bouharmont 1996) and RAPD markers (Anthony et al. 2001). The recent development of the amplified fragment length polymorphism (AFLP) approach (Vos et al. 1995), which enables simultaneous analysis of a large number of marker loci throughout the genome, appears to be remarkably powerful (Powell et al. 1996). The application of AFLP markers on coffee has proven to be effective in detecting introgression fragments in lines and derivatives of a spontaneous interspecific hybrid (*C. arabica* × *C. canephora*) (Lashermes et al. 2000). The simple-sequence repeats (SSRs) or microsatellite-based markers are also very attractive for plant genetic studies as they can be efficiently analysed by rapid and simple polymerase chain reaction (PCR) assays. Eleven microsatellite loci were recently identified in coffee, of which six were polymorphic within *C. arabica* (Combes et al. 2000).

In the investigation reported here, AFLP markers were used to assess polymorphism between and within Typica- and Bourbon-derived accessions and Yemen cultivars. Their relationship with wild material was defined analysing accessions derived from subspontaneous trees collected in the centre of origin of the species *C. arabica*. Microsatellite loci were used to confirm the structure of the genetic diversity detected by the AFLP markers and to compare the polymorphism within the genetic groups. The results are discussed in relation to historical data on coffee cultivation for outlining the main steps of coffee domestication.

## Materials and methods

### Plant material

The 26 accessions used in this study are presented in Table 1. Typica and Bourbon cultivars were selected from the CATIE field genebank that conformed to the botanical varieties described by Krug et al. (1939). Five commercial cultivars were analysed: two Typica-derived cultivars (Blue Mountain and Pluma Hidalgo), two Bourbon-derived cultivars (Caturra and Villasarchi) and a cultivar derived from a (Typica × Bourbon) × Bourbon hybrid and cultivated worldwide as Catuai (Carvalho et al. 1991). A Typica-derived mutant (Cera), a Bourbon-derived mutant (Mokka) and a Bourbon-derived di-haploid plant (n=22) were also included in the study. Four cultivars collected in the Popular Democratic Republic of Yemen by IPGRI (Eskes 1989) completed the sampling of cultivated material. Eleven subspontaneous-derived accessions were used: four accessions from the FAO collection in Ethiopia (FAO 1968), five accessions from the ORSTOM (now IRD) collection in Ethiopia (Guillaumet and Hallé 1978), Rume Sudan originating from southern Sudan (Thomas 1942) and Geisha from southwestern Ethiopia (Jones 1956). The subspontaneous-derived accessions included in this study were selected to be representative of the genetic diversity detected by RAPD markers in wild coffee (Anthony et al. 2001). The term "subspontaneous" refers to material derived from wild coffee and cultivated locally in Ethiopia. One individual was analysed per accession except for cv. Catuai, which was represented by two plants.

### AFLP analysis

DNA was isolated from lyophilised leaves according to Agwanda et al. (1997). The AFLP procedure was performed essentially as described by Vos et al. (1995) with minor modifications for coffee DNA (Lashermes et al. 2000). An aliquot of 500 ng genomic

**Table 1** Studied accessions of coffee

| Accession | Description | Origin | Source[a] |
|---|---|---|---|
| Typica | Cultivar conform to *C. arabica* var. typica | Brazil | T996 |
| Bourbon | Cultivar conform to *C. arabica* var. bourbon | Brazil | T995 |
| Blue Mountain | Commercial cultivar (Typica) | Jamaica | T977 |
| Pluma Hidalgo | Commercial cultivar (Typica) | Mexico | T3629 |
| Caturra | Commercial cultivar (Bourbon) | Brazil | Ct7 |
| Villasarchi | Commercial cultivar (Bourbon) | Costa Rica | T17603 |
| Catuai 8 | Commercial cultivar (Typica × Bourbon) | Brazil | C8 |
| Catuai 10 | Commercial cultivar (Typica × Bourbon) | Brazil | C10 |
| Cera | Typica-derived mutant | Brazil | T2309 |
| Mokka | Bourbon-derived mutant | La Réunion | T2313 |
| HA | Bourbon-derived di-haploid (n=22) | Côte-d'Ivoire | 113/B |
| PDRY-7 | Cultivar (IPGRI collection) | Yemen | T21236 |
| PDRY-14 | Cultivar (IPGRI collection) | Yemen | T21239 |
| PDRY-15 | Cultivar (IPGRI collection) | Yemen | T21240 |
| PDRY-22 | Cultivar (IPGRI collection) | Yemen | T21242 |
| E-12 | Subspontaneous (FAO collection) | Ethiopia | T4950 |
| E-18 | Subspontaneous (FAO collection) | Ethiopia | T4474 |
| E-238 | Subspontaneous (FAO collection) | Ethiopia | T4759 |
| E-536 | Subspontaneous (FAO collection) | Ethiopia | T4905 |
| ET-5 | Subspontaneous (ORSTOM collection) | Ethiopia | T16693 |
| ET-6 | Subspontaneous (ORSTOM collection) | Ethiopia | T17177 |
| ET-32B | Subspontaneous (ORSTOM collection) | Ethiopia | T17205 |
| ET-52 | Subspontaneous (ORSTOM collection) | Ethiopia | T16733 |
| ET-59 | Subspontaneous (ORSTOM collection) | Ethiopia | T16739 |
| Geisha | Subspontaneous | Ethiopia | T2722 |
| Rume Sudan | Subspontaneous | Sudan | T2724 |

[a] T, C and I indicate CATIE (Costa Rica), CICAFE (Costa Rica) and IDEFOR (Ivory Coast) sources, respectively

**Table 2** The 37 primer combinations used in the study

| E+3[a] | M+3[a] | E+3 | M+3 | E+3 | M+3 |
|---|---|---|---|---|---|
| AAC | CAA | ACA | CAG | ACT | CAC |
| AAC | CAC | ACA | CAT | ACT | CAT |
| AAC | CAG | ACA | CTG | ACT | CTA |
| AAC | CTA | ACA | CTT | ACT | CTG |
| AAC | CTC | ACC | CAA | ACT | CTT |
| AAC | CTG | ACC | CAG | AGC | CAA |
| AAC | CTT | ACG | CAA | AGC | CAC |
| AAG | CAA | ACG | CAT | AGC | CAG |
| AAG | CTA | ACG | CTA | AGC | CTC |
| AAG | CTG | ACG | CTC | AGC | CTG |
| AAG | CTT | ACG | CTG | AGG | CTG |
| ACA | CAA | ACG | CTT | | |
| ACA | CAC | ACT | CAA | | |

[a] E+3 and M+3: 3′ end-selective nucleotides of the primers complementary to the *Eco*- and *Mse*-adapter, respectively

DNA was digested with restriction enzymes *Eco*RI and *Mse*I. Restriction fragments were ligated with double-strand *Eco*RI-and *Mse*I-adapters. A preamplification was done using the appropriate primer combinations with one added selective nucleotide: E+A/M+C. The code following E or M corresponds to the selective nucleotides at the 3′ -end of the *Eco*RI and *Mse*I primers, respectively. The reaction mix was diluted 1/30, and 10 µl was used for the final amplification with two primers, each having three selective nucleotides. Thirty-seven AFLP primer combinations generated clear patterns with coffee accessions and detected polymorphism among the accessions used in this study (Table 2).

### Microsatellite analysis

Six microsatellites that previously showed clear polymorphisms in *C. arabica* (Combes et al. 2000) were used in this study. The SSRs were derived from a partial genomic library (*C. arabica* cv. Caturra), enriched for $(TG)_{13}$ motifs (Vascotto et al. 1999). Reaction mixtures for the PCR amplification of microsatellite loci contained 25 ng of genomic DNA, 10 m*M* Tris-HCI, pH 9.0, 0.1% Triton X-100, 1.5 m*M* of $MgCl_2$, 0.2 pmol of each primer, 0.2 m*M* of dCTP, dGTP, dTTP, 0.01 m*M* of dATP, 0.8 mCi of [$^{33}$ P]-dATP (Amersham Pharmacia, Piscataway, N.J.) and 0.5 U of *Taq* DNA polymerase (Promega, Madison, W.s.) in a 25-µl final volume. Reactions were performed in a PTC-200 thermocycler (MJ Research). The amplification cycle consisted of an initial 2-min denaturation at 94 °C, followed by 5 cycles of denaturation at 94 °C for 45 s, a 1-min primer annealing at 60 °C, with a 1 °C decrease in temperature at each cycle, and 1 min 30 s elongation at 72 °C. Then, 30 cycles of 45 s at 90 °C, 1 min at 55 °C and 1 min 30 s at 72 °C were performed, and followed by a final 8-min elongation at 72 °C. Amplification products were separated on 6% denaturing polyacrylamide gel with 8 *M* urea and 1× TBE. A radioactively labelled 10-bp ladder DNA was used as a size standard.

### Data analysis

Only AFLP bands showing a clear polymorphism were scored as present (1) or absent (0). These were designated according to the restriction enzymes, the primer combinations used and fragment size, then coded from 1 to 107 in order to simplify their identification. The fragments amplified by the six primer pairs targeting polymorphic microsatellite loci were coded by allele and designated according to fragment size (bp) as estimated by Combes et al. (2000). The genetic distance between pairs of accessions was estimated on the basis of the Jaccard coefficient (Jaccard 1908). The matrix of genetic distances was analysed by the unweighted pair-group method using arithmetic averages (UPGMA) (Sneath and Sokal 1973) of the TREE procedure of TREECON software (Van de Peer and De Wachter 1994), version 1.3. The bootstrap method (Felsenstein 1985) was employed to evaluate the reliability of tree topologies. The genetic identity and the genetic distances (Nei 1972) between the groups of the dendrogram were estimated using POPGENE software, version 1.21 (University of Alberta and Center for International Forestry Research) for dominant markers.

## Results

### Diversity structure

A total of 107 AFLP polymorphic markers were detected among the studied accessions and scored for genetic analysis. These markers were used to construct a dendrogram by group average clustering (UPGMA) using genetic distances between accession pairs (Fig. 1). The

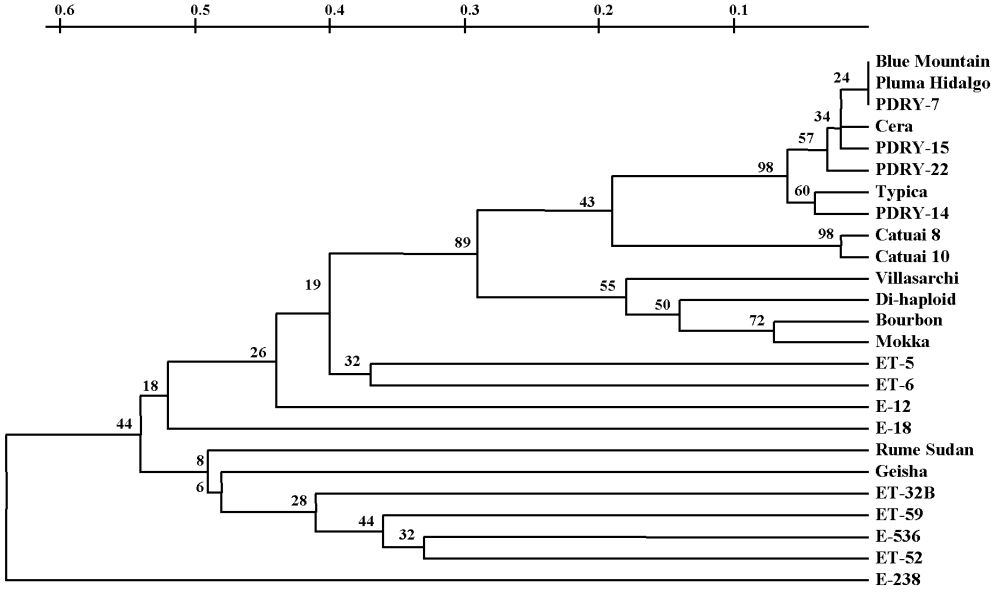


**Fig. 1** Dendrogram generated after UPGMA using AFLP-based genetic distance. *Numbers* on the branches are bootstrap values (%) obtained from 200 replicate analyses.

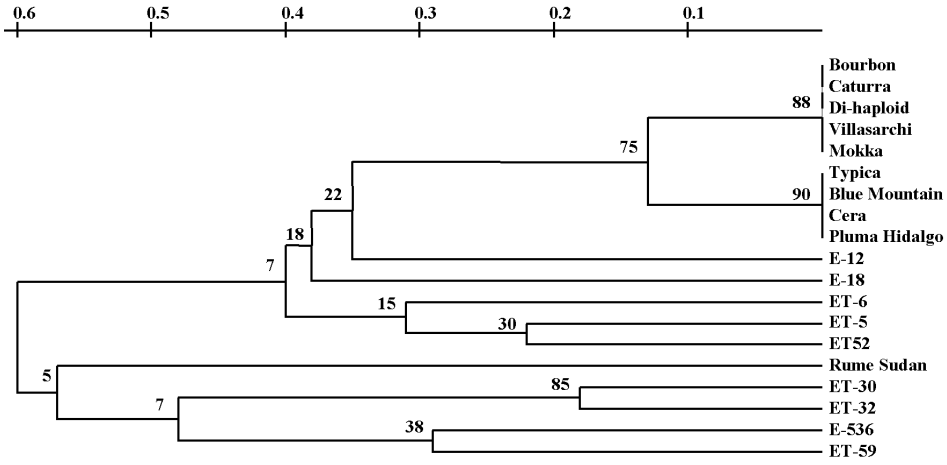


**Fig. 2** Dendrogram generated after UPGMA using SSR-based genetic distance. *Numbers* on the branches are bootstrap values (%) obtained from 200 replicate analyses.

Typica- and Bourbon-derived accessions were classified into two distinct groups according to their genetic origin, each group being supported by a high bootstrap value (89%). The four Yemen cultivars were clearly grouped with the Typica-derived accessions. The classification of the di-haploid accession with the Bourbon-derived accessions conformed to its genetic origin (i.e. Bourbon). The two plants of cv. Catuai were grouped with the Typica-derived accessions and Yemen cultivars. The subspontaneous-derived accessions were clearly separated from the cultivated accessions and did not constitute a group, but rather chains, with poor bootstrap values. The subspontaneous-derived accession from southern Ethiopia (E-238) was classified separately from the southwestern accessions.

The dendrogram generated after UPGMA using SSR-based genetic distances (Fig. 2) was similar to the dendrogram constructed with AFLP data. The Typica- and Bourbon-derived accessions were grouped according to their genetic origins, with high bootstrap value (75%), and classified separately from the subspontaneous-derived accessions.

## Between-group genetic distances

The genetic distances (Nei 1972) calculated from AFLP data were low between the Typica and Bourbon groups, 0.17 on average. Cv. Catuai was shown to be close to the Typica- (0.095) and Bourbon-derived (0.085) accessions. Higher values were estimated between cultivated and subspontaneous-derived accessions. As the subspontaneous-derived accessions studied did not constitute a homogeneous group, it is more appropriate to give the range of calculated values rather than the average. The genetic distances ranged from 0.40 to 0.71 between the accessions classified in the Typica group and the subspontaneous-derived accessions, and from 0.37 to 0.70 between the Bourbon- and the subspontaneous-derived accessions. The lowest values in the Typica (0.40) and Bourbon (0.37) groups were found between two accessions (Typica and PDRY-14) and the subspontaneous-derived accession ET-6, and between the Bourbon-derived accession Mokka and the subspontaneous-derived accession ET-5. The highest values were generated by the subspontaneous-derived accession E-238. This accession presented high genetic distances with all of the Typica- (0.68–0.71) and Bourbon-derived (0.67–0.70) accessions, and variable genetic distances with the other subspontaneous-derived accessions, ranging from 0.48 (Geisha) to 0.67 (E-12). Except for accession E-238, the maximums were significantly lower; 0.61 between the Yemen cultivar PDRY-14 and the subspontaneous-derived accession Geisha, and 0.58 between the Bourbon-derived accession Mokka and the subspontaneous-derived accession ET-59.

## Characterisation of the genetic groups

Of the 107 AFLP polymorphic markers 96 were present in the subspontaneous-derived accessions, while only 59, 55 and 54 were present in the Bourbon-derived accessions, Typica-derived accessions and Yemen cultivars, respectively (Table 3). Most of the markers detected in these accessions were also detected in the subspontaneous-derived accessions: 52 markers (88%) of Bourbon-derived accessions, 52 markers (95%) of Typica-derived accessions, and 50 markers (93%) of Yemen cultivars. Thirty-seven markers were present in all of the cultivated accessions, from which 36 were polymorphic in the subspontaneous-derived accessions. One marker was found to be specific to the cultivated material. Seven markers common to all Typica-derived accessions and Yemen cultivars were not detected in Bourbon-derived accessions. Seven markers present in all Bourbon-derived accessions were not detected in Typica-derived accessions and Yemen cultivars. The two plants of cv. Catuai presented 53 markers, of which 44 were detected in both Typica- and Bourbon-derived accessions and nine in either Typica- or Bourbon-derived accessions. One marker present in all Typica- and Bourbon-derived accessions was not found in the cv. Catuai. The subspontaneous-derived accession E-238 differed from the other

**Table 3** Distribution of 107 AFLP markers in three Typica-derived accessions, four Yemen cultivars, four Bourbon-derived accessions and 11 subspontaneous-derived accessions

| Number of markers | Typica | Yemen | Bourbon | Subspontaneous |
|---|---|---|---|---|
| 36 | 1[a] | 1 | 1 | + |
| 1 | 1 | 1 | 1 | 0 |
| 2 | 1 | 1 | + | 1 |
| 4 | 1 | 1 | + | + |
| 6 | 1 | 1 | 0 | + |
| 1 | 1 | 1 | 0 | 0 |
| 1 | + | + | 0 | + |
| 1 | + | 0 | + | + |
| 1 | + | + | 1 | + |
| 7 | 0 | 0 | 1 | + |
| 1 | 0 | 0 | + | + |
| 1 | + | 0 | 0 | + |
| 1 | + | 0 | 0 | 0 |
| 2 | 0 | + | 0 | 0 |
| 6 | 0 | 0 | + | 0 |
| 36 | 0 | 0 | 0 | + |

[a] 0, Absent; 1, present in all accessions; +, polymorphic

**Table 4** Number of alleles and products sizes for six microsatellite loci

| Locus | Product size (bp) | Allele number |
|---|---|---|
| M20 | 252–268 | 5 |
| M24 | 139–169 | 6 |
| M25 | 168–170 | 2 |
| M29 | 120–122 | 2 |
| M32 | 112–135 | 6 |
| M47 | 100–132 | 7 |

subspontaneous-derived accessions by the presence of three specific markers and the absence of five markers.

The six SSRs surveyed allowed for the identification of a total of 28 alleles among the four Typica-, five Bourbon- and ten subspontaneous-derived accessions studied (Table 4). The number of alleles per locus ranged from 2 to 7 and was lower within the Typica- and Bourbon-derived accessions (8 within each group) than within the subspontaneous-derived accessions, the latter, which presenting all of the alleles identified in the study (28). Seven alleles were present in all Typica- and Bourbon-derived accessions and Yemen cultivars. Two alleles (M24-147 and M24-149) discriminated between the Typica- and Bourbon-derived accessions.

### Within-group polymorphism

The number of polymorphic markers varied significantly within the three main groups of the dendrogram. Whereas 94 AFLP markers were polymorphic within the subspontaneous-derived accessions, only 7 and 14 were respectively polymorphic within the Typica-derived accessions and the Yemen cultivars, and within the Bourbon-derived accessions (Table 3). The Typica-derived accessions only presented five polymorphic markers, of which three were absent in the Bourbon-derived accessions. Seven polymorphic markers within the Bourbon-derived accessions were absent in the Typica-derived accessions. No difference was detected between the two plants of cv. Catuai.

The SSR markers confirmed the low polymorphism detected in cultivated accessions by AFLP markers and the high polymorphism within the subspontaneous-derived accessions, 0% and 100% of the identified alleles, respectively.

## Discussion

### Diversity of the genetic bases Typica and Bourbon

Genetic diversity, as expressed by the number of markers detected, and polymorphism appear to be much weaker in the cultivated accessions than in the subspontaneous-derived accessions. Only about half of the 107 AFLP markers in the study were observed in the Typica- and Bourbon-derived accessions, but 90% of the markers were found in the subspontaneous-derived accessions. Polymorphism was revealed to be much more important within the subspontaneous-derived accessions (94 AFLP markers) than within the Typica- (5) and Bourbon-derived (14) accessions. The diffusion of coffee and the selection that followed have therefore strongly reduced the genetic diversity present in the subspontaneous coffee of Ethiopia. Polymorphism was reduced during the cycles of selection, due to the homogenisation of genetic structures, favoured by the predominant autogamy of the species *C. arabica* (Carvalho et al. 1991).

With the exception of one AFLP marker, all of the markers of the cultivated accessions were also found in the subspontaneous-derived accessions. This result confirms the Ethiopian origin of the Typica and Bourbon genetic bases (Anthony et al. 2001). The cultivated accessions appeared to be more differentiated from the subspontaneous-derived accession E-238 collected in southern Ethiopia than the other accessions studied, which originated in southwestern Ethiopia. The classification based on the AFLP markers confirmed the partition of the subspontaneous coffee into two groups, separated by the Great Rift Valley (Anthony et al. 2001). The origin of the cultivar genetic bases is, in fact southwestern Ethiopia, where Lejeune (1958) situated the first cultivation of coffee.

Genetic differentiation appears to be weak between the Typica- and Bourbon-derived accessions, as shown by the estimation of their genetic distance (0.17). The differences between the two genetic bases only concerned 22% of the AFLP markers and 7% of the SSR markers. This weak differentiation does not permit an explanation of the vigour shown by the hybrids (Typica × Bourbon), selected in Brazil under the name cv. Mundo Novo (Carvalho et al. 1969). The heterosis of the hybrids

could be a result of complementary action of certain genes that condition the whole plant, rather than the genetic differentiation of their parents.

The polymorphism detected within the Typica and Bourbon genetic bases was weak as determined by the AFLP markers and nonexistent as determined by the SSR markers. Only 5 and 14 markers identified in the Typica- and Bourbon-derived accessions were revealed to be polymorphic. The higher polymorphism in the Bourbon genetic base indicates that it was constituted from the descendants of several individuals and not from one single individual as in the Typica genetic base. This result confirms the historical data given by Haarer (1956), in which several introductions took place from Yemen to Reunion Island. Given the importance of the commercial exchanges in the 18th and 19th centuries, it is also probable that re-introductions to Reunion Island could have taken place from other islands in the Indian Ocean such as Mauritius, where seeds from Yemen were introduced around 1715 (Wellman 1961), or from South East Asia.

## Diversity of Yemen cultivars

The Yemen cultivars presented a number of AFLP and SSR markers similar to the Typica- and Bourbon-derived accessions. These cultivars were classified unambiguously with the Typica-derived accessions by the AFLP markers. This result is in accordance with historical data on the diffusion of the Typica genetic base (Chevalier and Dagron 1928; Chevalier 1929) from Yemen to Amsterdam, via Java. The presence of a microsatellite allele common to the Bourbon-derived accessions can be interpreted as a result of a common origin of the Typica and Bourbon genetic bases – from the same pool of Ethiopian coffee introduced to Yemen. According to Haarer (1956), the coffee-trees growing in Reunion Island Introduced from Yemen in 1715 and 1718. Yemen appears to be the first centre of dispersion for coffee outside of Ethiopia, as suggested by Meyer (1965).

The four Yemen cultivars revealed little polymorphism between them, although they represent three of the six morphological types described by the prospectors of the IPGRI mission (Eskes 1989): (1) Essaii (PDRY-7), (2) Tessawi (PDRY-14) and Katii (PDRY-l5) and (3) a local type (PDRY-22). Only 4 of the 107 AFLP markers in the study revealed differences at the level of their DNA. Therefore, they have differentiated little since their introduction from Ethiopia.

## History of coffee dissemination

The history of coffee cultivation is charactersied by the successive reductions of diversity within the two subpopulations of wild coffee introduced from Ethiopia to Yemen (Fig. 3). According to different authors, the first reduction occurred when coffee was introduced to Yemen, between 1,500 and 300 years ago. It was, without a doubt, the most Draconian in terms of loss of diversity even if it is likely that several introductions followed from Ethiopia to Yemen. The classification by the AFLP markers and the estimation of genetic distances have shown that the cultivars are closer to the subspontaneous coffee of the west side of the Great Rift Valley than to the east side. If the coffee of Yemen came from seeds harvested in the Harar region, as affirmed by Wellman (1961), it was an intermediate step in the introduction to Yemen, since no references have mentioned the existence of wild coffee on the east side of the tectonic break.

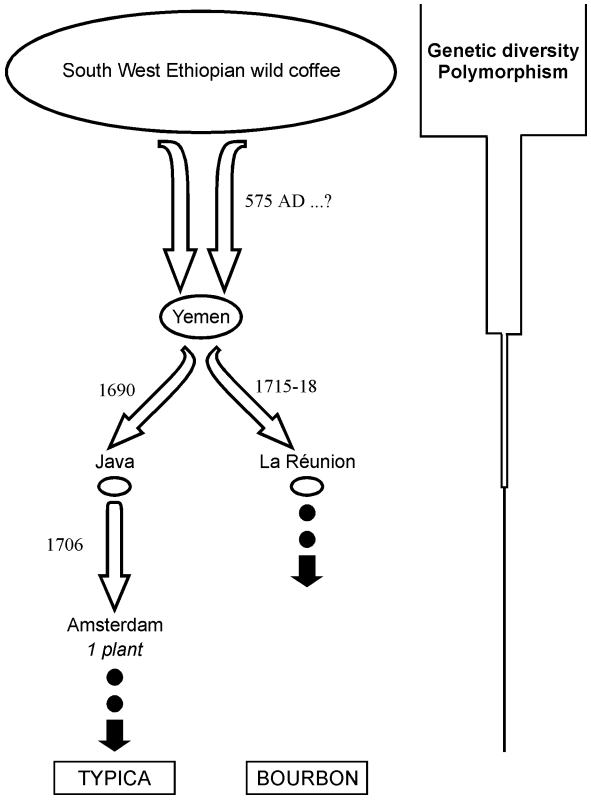


**Fig. 3** Schematic representation of the main steps in the history of coffee cultivation

Further reductions in diversity were produced during the introduction of coffee to Java, then Amsterdam and Reunion Island at the beginning of the 18th century. These introductions included a limited number of seeds and plants, a single plant to Amsterdam and several descendants to Reunion Island. For the origin of the cultivated coffee on Reunion Island, Haarer (1956) stated that the first introduction in 1708 was composed of seeds and 60 small plants, of which none survived. Only two plants from the second introduction in 1715 survived. This information testifies to the difficult conditions for cultivation that were faced by the first coffee plants introduced. This intense selection doubtlessly contributed to the reduction of introduced polymorphism. Coffee culture then spread rapidly to the American continent and Indonesia from seeds produced by the auto-fertilisation of some surviving coffee trees.

AFLP effectiveness for coffee characterisation

AFLP markers proved to be effective for characterising the two genetic bases which were diffused in the 18th century, identifying their hybrids and distinguishing the subspontaneous coffee plants. The 37 primer combinations used in this study enabled the identification of 107 AFLP polymorphic markers among the 25 accessions analysed, which represents an average of 2.9 polymorphic markers per combination. The AFLP markers permitted the classification of the Typica- and Bourbon-derived accessions by their genetic origin, confirming the existence of two genetic bases for the cultivated material. Fourteen AFLP markers could be used to characterise the cultivars issued from the Typica and Bourbon genetic bases. The cultivars were clearly separated from the Ethiopian subspontaneous-derived accessions, as shown in previous studies based on the RAPD markers (Orozco-Castillo et al. 1994; Lashermes et al. 1996; Anthony et al. 2001). The AFLP markers detected in the cv. Catuai were consistent with its hybrid origin.

**Acknowledgements** The research was funded by IRD (Institut de recherche pour le développement, ex ORSTOM, Paris, France) and the European Community through the International Scientific Cooperation Programme (INCO-DC contract ERB1C18CT970181). C. Astorga was funded by a MSc grant from CATIE (Centro Agronómico Tropical de Investigación y Enseñanza, Turrialba, Costa Rica).

## References

Agwanda CO, Lashermes P, Trouslot P, Combes MC, Charrier A (1997) Identification of RAPD markers for resistance to coffee berry disease, *Colletotrichum kuhawae*, in arabica coffee. Euphytica 97:241–248

Anthony F, Berthaud J, Guillaumet JL, Lourd M (1987) Collecting wild *Coffea* species in Kenya and Tanzania. Plant Genet Resour News1 69:23–29

Anthony F, Bertrand B, Quiros 0, Lashermes P, Berthaud J, Charrier A (2001) Genetic diversity of wild coffee (*Coffea arabica* L.) using molecular markers. Euphytica 118:53–65

Carvalho A (1946) Distribuição geografica e classificação botânica do gênero *Coffea* com referência especial à espécie Arabica. V. Origem e classificaçãbotânica do *C. arabica* L. Bol Super Serv Caf 21:174–180

Carvalho A (1952) Taxonomia de *Coffea arabica* L. Caracteres morfologicos dos haploides. Bragantia 12:201–212

Carvalho A, Ferwerda FP, Frahm-Leliveld JA, Medina PM, Mendes AJT, Monaco LC (1969) Coffee. In: Ferwerda FP, Wit F (eds) Outlines of perennial crop breeding in the tropics. Veenman & Zonen, Wageningen, pp 189–241

Carvalho A, Medina Filho HP, Fazuoli LC, Guerreiro Filho O, Lima MMA (1991) Aspectos genéticos do cafeeiro. Rev Bras Genet l4:135–183

Chevalier A (1929) Les caféiers du globe, fascicule I. Généralités sur les caféiers. P. Lechevalier, Paris

Chevalier A, Dagron M (1928) Recherches historiques sur les débuts de la culture du caféier en Amérique. Commun Actes Acad Sci Colonial Paris

Combes MC, Andrzejewski S, Anthony F, Bertrand B, Rovelli P, Graziosi G, Lashermes P (2000) Characterisation of microsatellite loci in *Coffea arabica* and related coffee species. Mol Ecol 9:1178–1180

Cros J, Combes MC, Trouslot P, Anthony F, Hamon S, Charrier A, Lashermes P (1998) Phylogenetic relationships of *Coffea* species: new evidence based on the chloroplast DNA variation analysis. Mol Phylogenet Evol 9:109–117

Eskes AB (1989) Identification, description and collection of coffee types in P.D.R. Yemen. IPGRI, Rome

FAO (1968) FAO coffee mission to Ethiopia 1964–1965. FAO, Rome

Felsenstein J (1985) Confidence limits on phylogenies: an approach using the bootstrap. Evolution 39:783–791

Guillaumet JL, Hallé F (1978) Echantillonnage du matériel récolté en Ethiopie. Bull IFCC 14:13–18

Haarer AE (1956) Modern coffee production. Leonard Hill, London

Jaccard P (1908) Nouvelles recherches sur la distribution florale. Bull SOC Vaudoise Sci Nat 44:223–270

Jones PA (1956) Notes on the varieties of *Coffea arabica* in Kenya. Mon Bull Coffee Board Kenya 21:305–309

Krug CA, Mendes JET, Carvalho A (1939) Taxonomia de *Coffea arabica* L. Bol Tec no 62. Instituto Agronomico de Estado, Campinas, Brazil

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

Lashermes P, Combes MC, Robert J, Trouslot P, D'Hont A, Anthony F, Charrier A (1999) Molecular characterisation and origin of the *Coffea arabica* L. genome. Mol Gen Genet 261:259–266

Lashermes P, Andrzejewski S, Bertrand B, Combes MC, Dussert S, Graziosi G, Trouslot P, Anthony F (2000) Molecular analysis of introgressive breeding in coffee *(Coffea arabica* L.). Theor Appl Genet 100:139–146

Lejeune JBH (1958) Rapport au Gouvernement Impérial d'Ethiopie sur la production caféière. FAO, Rome

Meyer FG (1965) Notes on wild *Coffea arabica* from southwestern Ethiopia, with some historical considerations. Econ Bot 19:136–151

Montagnon C, Bouharmont P (1996) Multivariate analysis of phenotypic diversity of *Coffea arabica*. Genet Res Crop Evol 43:221–227

Nei M (1972) Genetic distance between populations. Am Nat 106:283–292

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

Powell W, Morgante M, Andre C, Hanafey M, Vogel J, Tingey S, Rafalski A (1996) The comparison of RFLP, RAPD, AFLP and SSR (microsatellite) markers for germplasm analysis. Mol Breed 2:225–238

Smith RF (1985) A history of coffee. In: Clifford MN, Willson KC (eds) Coffee botany, biochemistry and production of beans and beverage. Croom Helm, London Sydney, pp 1–12

Sneath PHA, Sokal RR (1973) Numerical taxonomy. W.H. Freeman, San Francisco

Sylvain PG (1955) Some observations on *Coffea arabica* L. in Ethiopia. Turrialba 5:37–53

Thomas AS (1942) The wild arabica coffee on the Boma Plateau, Anglo-Egyptian Sudan. Emp Exp Agric 10:207–212

Van de Peer Y, De Wachter R (1994) TREECON for Windows: a software package for the construction and drawing of evolutionary trees for the Microsoft Windows environment. Compnt Applic Biosci 10:569–570

Vascotto F, Degli Ivanissevich S, Rovelli P, Anthony F, Anzueto F, Lashermes P, Graziosi G (1999) Microsatellites in *Coffea arabica*: construction and selection of two genomic libraries. In: Proc 3rd Int Sem Biotechnol Coffee Agroind. Londrina, Brazil, pp 125–128

Vos P, Hogers R, Bleeker M, Reijans M, Van der Lee T, Hornes M, Frijters A, Pot J, Peleman J, Kuiper M, Zabeau M(1995) AFLP: a new technique of DNA fingerprinting. Nucleic Acids Res 23:4407–4414

Wellman FL (1961) Coffee, botany, cultivation and utilization. Leonard Hill Books, London