MOLECULAR ANALYSIS OF GENETIC DIVERSITY AND PHYLOGENETIC RELATIONSHIPS IN COFFEA

J. CROS, Ph. LASHERMES, Ph. MARMEY, F. ANTHONY, S. HAMON, A. CHARRIER

ORSTOM, Laboratoire de Ressources Génétiques et Amélioration des Plantes Tropicales, 911 Av. Agropolis, BP 5045, 34032 Montpellier, France

Introduction

*Coffea* subgenus *Coffea* consists of approximately 100 taxa so far identified. Commercial coffee production relies on only two species: *Coffea arabica* and *Coffea canephora*, but many species are a valuable gene reservoir for different breeding purposes (Berthaud & Charrier, 1988).

Coffee genetic resources have been analyzed using geographical distribution, cytological observations and taxonomic data including agro-morphological and biochemical characteristics (Charrier, 1977; Berthou & al., 1980; Louarn, 1982; Anthony & al., 1989; Clifford & al., 1989; Rakotomalala & al., 1993). However, structure of the genetic diversity and phylogenetic relationships between species remain imprecise.

Molecular marker techniques are particularly suitable for genetic diversity analysis. DNA-based marker techniques such as Random Amplified Polymorphic DNA (RAPD) and Restriction Fragment Length Polymorphism (RFLP) have been recently adapted to coffee germplasm at ORSTOM. In this communication, we report two preliminary studies.

Use of RAPD for assessing variation in coffee

The RAPD technique (Williams et al., 1990) based on the polymerase chain reaction (PCR) offers a new class of DNA markers which present particular interests. This approach provides many advantages over the RFLP/southern blotting approach to revealing polymorphisms. It is faster, does not use cloned probes and is independent of prior DNA sequence information. The amplification protocol differs from the standard PCR conditions in that only a short single random oligonucleotide is employed as a primer. Number and size of fragments generated by the RAPD system strictly depend on the nucleotide sequence of the primer used and on the source of the template DNA, resulting in a genome-specific "fingerprint" of random DNA fragments. Use of such fragments as genetic markers in *Coffea* was investigated. Nineteen coffee samples representing major *Coffea* species (*arabica*, *canephora*, *congensis*, *eugenioides*, *liberca*, *restinosa*, *stenophylla*, *pseudozanguebariae*) were studied. *C. arabica*, *C. canephora* and *C. liberica* were represented by different plants which were chosen to display a wide genetic variability. In addition, we analysed Hibrido de Timor, a tetraploid genotype that presents a phenotype like *C. arabica* and combines important resistance to coffee berry disease (CBD) and to most rust races (Moreno, 1989).
Twenty-three arbitrary oligonucleotides were used singly as primers for the amplification of random DNA sequences from genomic DNA. DNA extraction and RAPD experiments were carried out as described by Lashermes & al. (1993). RAPD were scored as dominant markers (presence versus absence), and a similarity index (D) expressing the probability that a RAPD in one sample is also found in another was calculated according to Wetton & al. (1987) for all possible pairwise comparisons between accessions.

All samples generated comparable numbers of amplified products (mean of 4.3 amplified fragments per primer) with the exception of C. pseudozanguebariae and C. sp. A801 which produced consistently less fragments. This result can be related to the relatively low genome size determined for C. pseudozanguebariae by laser flow cytometry (Cros & al. 1993).

Intraspecific variation was easily detected in C. canephora and C. liberica. Comparisons between canephora yielded D-values of 0.62 to 0.73. Liberica variety liberica yielded D=0.66 when compared with the liberica variety dewevrei. On the other hand, the primers assayed failed to reveal polymorphism between C. arabica. Hibrido de Timor was found to be slightly different from arabica samples. In particular, results showed that Hibrido de Timor shares a marker (common amplified product) with one accession of C. canephora, confirming that Hibrido de Timor most likely originated from a spontaneous interspecific cross between C. arabica and C. canephora.

Considerable interspecific genetic variation was evidenced within the range of Coffea species analysed. More than 50% of the amplified DNA fragments differed between all pairwise species. A hierarchical clustering analysis (Benzecri 1973) was performed to generate a dendrogram showing genetic relationships between accessions (Figure 1). The coffee species from West Africa (C. canephora, C. liberica, C. congensis and C. stenophylla), and from the highland forest of Kenya and Ethiopia (C. arabica and C. eugenioides) present a high similarity and constitute a first group. The different accessions of C. canephora as well as the two accessions of C. liberica cluster before they join the clusters of other species. C. resinosa originated in Madagascar and appears distantly related to all species surveyed and forms a second group. C. pseudozanguebariae indigenous to the costal region of East Africa also presents a low similarity with all species analysed. Only C. sp. A801, accession unidentified, appears closely related to C. pseudozanguebariae. These results are consistent with the classification in three sections proposed by Berthaud (1986) based on morphological and cytological studies: Erythrocoffea, Mozambicoffea and Mascarocoffea including respectively Coffea species from West and Central Africa, the Indian Ocean coast of East Africa and Madagascar.

![Figure 1: Dendrogram showing genetic relationships between Coffea accessions constructed by hierarchical clustering analysis using RAPD markers.](image-url)
RFLP analysis of chloroplastic DNA

Analysis of the chloroplast genome (cpDNA) has proved to be a very powerful approach for determining phylogenetic relationships among populations and species (Hooglander et al., 1993; Dally & Second, 1990; White, 1990; Sytsma & Gottlieb, 1986; Wilson et al., 1992). The cpDNA is circular and usually ranges in size from 120 to 150 kilobase pairs (Pillay, 1993). Its gene organization and nucleotide order are extremely conserved through evolutionary time and make it an ideal target for plant phylogenetic study. Estimates of cpDNA variability can be obtained by comparison of cpDNA restriction fragment length polymorphism (RFLP) (Jansen et al., 1990; Bremer, 1991).

The purpose of the current study is 1) to determine levels of cpDNA variation within and between Coffea species and 2) to determine relationships among Coffea species by chloroplast DNA phylogeny.

<table>
<thead>
<tr>
<th>Species</th>
<th>Accession (Code/population)</th>
<th>Origin</th>
</tr>
</thead>
<tbody>
<tr>
<td>C. arabica</td>
<td>ET 12-5</td>
<td>Ethiopia</td>
</tr>
<tr>
<td>C. arabica</td>
<td>Hibrido de Timor</td>
<td>Timor island</td>
</tr>
<tr>
<td>C. bertrandi</td>
<td>HASK</td>
<td>Madagascar</td>
</tr>
<tr>
<td>C. brevipes</td>
<td>Mt Cameroun</td>
<td>Cameroon</td>
</tr>
<tr>
<td>C. brevipes</td>
<td>Kumba Loum</td>
<td>Cameroon</td>
</tr>
<tr>
<td>C. brevipes</td>
<td>var. heterocalyx</td>
<td></td>
</tr>
<tr>
<td>C. canephora</td>
<td>IF A25</td>
<td>Congo</td>
</tr>
<tr>
<td>C. canephora</td>
<td>BC8 (de la Nana)</td>
<td>Cameroon</td>
</tr>
<tr>
<td>C. congensis</td>
<td>CA9</td>
<td>Cameroon</td>
</tr>
<tr>
<td>C. congensis</td>
<td>CB9</td>
<td>Cameroon</td>
</tr>
<tr>
<td>C. eugenioides</td>
<td>A16</td>
<td>Kenya</td>
</tr>
<tr>
<td>C. farafanganesis</td>
<td>A 208</td>
<td>Madagascar</td>
</tr>
<tr>
<td>C. humilis</td>
<td>G2</td>
<td>Ivory Coast</td>
</tr>
<tr>
<td>C. liberica</td>
<td>EB25 (var. dewevrei)</td>
<td>Madagascar</td>
</tr>
<tr>
<td>C. millotii</td>
<td>CM1</td>
<td>Madagascar</td>
</tr>
<tr>
<td>C. perieri</td>
<td>H35</td>
<td>Kenya</td>
</tr>
<tr>
<td>C. pseudozanguebariae</td>
<td>IB9</td>
<td>Tanzania</td>
</tr>
<tr>
<td>C. racemosa</td>
<td>IA10</td>
<td>Mozambique</td>
</tr>
<tr>
<td>C. racemosa</td>
<td>PA4 (Shimba)</td>
<td>Kenya</td>
</tr>
<tr>
<td>C. sessiliflora</td>
<td>FB1 (Assabli)</td>
<td>Ivory Coast</td>
</tr>
<tr>
<td>C. stenophylla</td>
<td>FA21 (Ira)</td>
<td>Ivory Coast</td>
</tr>
<tr>
<td>C. stenophylla</td>
<td>FA41 (Shimba)</td>
<td></td>
</tr>
</tbody>
</table>

Table 1: Accessions used for RFLP analysis of chloroplastic DNA.

Twenty-five accessions representing 15 Coffea species were analysed. Code and origin of samples are listed in table 1. All material came from the ORSTOM genetic resources collection except IF A25, a cultivated clone of C. canephora. C. brevipes var. heterocalyx is an accession the origin of which is unknown (Anthony, 1991). Total DNA samples were extracted from leaves. Freeze-dried tissue (1 g dry weight) was mechanically ground to fine powder, dispersed in MATAB extraction buffer (Tris-HCl 0.1M, NaCl 1.25M, EDTA 0.02M, MixedAlkyTriMethylAmmoniumBromide 2%, βmercaptoethanol) and incubated at 60°C for 30 min. with slow rocking. After two consecutive chloroform:isoamyl alcohol (24:1) extractions the aqueous phase was transferred. DNA was precipitated with isopropanol and resuspended in TRIS-EDTA buffer. Following a second precipitation with ethanol, DNA was washed in 76% ethanol, dried onto a Kimwipe and resuspended in TRIS-EDTA buffer. Following a second precipitation with ethanol, DNA was washed in 76% ethanol, dried onto a Kimwipe and resuspended in TRIS-EDTA buffer. Total DNA from each sample was digested singly for completion with Eco RI and with Eco RV restriction enzymes. After electrophoresis, restriction fragments were transferred to a nylon membrane (Hybond N+) under alkaline condition.

Fifteen clones of Lactuca sativa cpDNA SacI fragments served as heterologous chloroplast DNA probes for southern analysis (Jansen and Palmer, 1987). These fragments represent approximately 95% of the
Lactuca chloroplastic genome (150Kb). CpDNA inserts were radiolabelled by random priming. Prehybridization, hybridization and wash were performed in accordance with manufacturer's recommendations. Autoradiography was carried out using Amersham Hyperfilm-MP film at -80°C.

Hybridization of the 15 cpDNA probes to Southern blots obtained for both enzymes revealed a total of 101 restriction fragments but only 33 fragments were phylogenetically informative. Differences between theendonuclease/probe combinations were observed. RFLPs were detected by only 6 of the 15 cpDNA probes, and digestion with EcoRV revealed significantly more polymorphism than the EcoRI digests. The chloroplast genome of Coffea species showed a high degree of colinearity with the Lactuca genome, as has been observed in angiosperm and several other Rubiaceae (Bremer and Jansen, 1991).

Interpretation of banding patterns was undertaken. Presence of either restriction-site mutation or length mutation was considered. It was assumed that variability was due to restriction-site polymorphism if the sum of the two restriction fragments equaled the size of a missing fragment. Six different restriction-site mutations were unambiguously identified.

Combinations of these six restriction-site mutations produced 8 different plastotypes. Pairwise distance (Jaccard index) between all individual genotypes were calculated and entered into a distance matrix program (PHILYP computer package; Felsenstein, 1987) to generate a UPGMA tree (unweighted pair group of matrix average). The number of site-mutations was considered too low to construct parsimony phylogenetic trees.

Figure 2: Dendrogram showing genetic relationships between Coffea species based on preliminary chloroplastic data.

The clustering analysis is presented in figure 2. Despite the low number of site mutations considered, we were able to distinguish most Coffea species.

The main separation was found between the group constituted by C.canephora / C.congensis / C.brevipes , and other species, but caution needs to be exercised in interpretation of this dendrogram. Identification of additional site mutation changes is required for further analysis.

Intraspecific variations were found in C. liberica and C. brevipes. The identification of several restriction-site changes between liberica variety liberica and liberica variety dewevrei was to be expected since the two varieties present large morphological differences and their hybrid shows a reduced fertility (Louam, 1993). Regarding C. brevipes, it has been already proposed to consider the taxa C. brevipes variety heterocalyx distinct from C. brevipes (Louam, 1992).
Conclusion and prospects

In recent years, molecular marker techniques have gained widespread applications in many fields of plant genetics and evolution. With regards to the analysis of the genetic diversity of coffee, both RAPD and RFLP marker techniques offer unique opportunity and will be of great interest.

RAPD assay provides a highly effective and convenient means to "fingerprint" coffee plants. This method should therefore be of high value for germplasm characterization and genetic resource maintenance in Coffea. Applications could include fingerprinting of genotype, identification of duplicate samples and analysis of genetic diversity in a collection. The usefulness of RAPD markers for genetic mapping has been largely reported (Carlson et al., 1991; Klein-Lankhorst et al., 1991; Martin et al., 1991). In connection with assisted backcross-breeding (Tanksley et al., 1989), RAPD technology is obviously a very powerful tool to increase the effectiveness of introgression to cultivated species of desirable traits (e.g. rust resistance) from wild coffee material or spontaneous hybrid such as Hibrido de Timor.

RFLP analysis of chloroplastic DNA appeared as a very attractive approach for determining phylogenetic relationships among Coffea species. Highly informative genetic variations were evidenced in this preliminary study despite the limited number of restriction enzymes assayed. This study on chloroplastic DNA will be extended to a higher number of species.

Summary

Molecular markers techniques provide suitable tools for analysis of genetic diversity and phylogenetic relationships in Coffea species. Two preliminary studies are reported:

- Arbitary oligonucleotides were used as primers to amplify genomic DNA of different coffee accessions (representing major Coffea species) by polymerase chain reaction. Extensive interspecific variation was observed, intraspecific variation was easily detected in C. canephora and C. liberica. Random amplified DNA markers appeared to be of high value for characterization, analysis and utilization of coffee genetic resources.

- A set of 15 chloroplastic DNA fragments from Lactuca sativa were used to probe southern blot of total DNA of 25 Coffea accessions digested by 2 restriction endonucleases. Nine of the 30 enzymes/probes combinations show polymorphism. Six polymorphic restriction-sites were evidenced. Height different plastotypes were identified. The low level of interspecific variation detected was enough to give phylogenetic information.

References


Louarn J., 1992. La fertilité des hybrides interspécifiques et les relations génomiques entre caféiers diploïdes d'origine africaine (Genre Coffea sous genre Coffea) Thèse, Université de Paris -Sud centre d'Orsay. 200p.


QUINZIÈME COLLOQUE SCIENTIFIQUE INTERNATIONAL SUR LE CAFÉ

Montpellier, 6-11 juin 1993

Volume I

20 JAN. 1994

ORSTOM
CENTRE DOCUMENTATION
Montpellier

Association Scientifique Internationale du Café (ASIC)
42, rue Scheffer, 75116 Paris