POLYMORPHIC MICROSATELLITES IN *COFFEA ARABICA*

R. METTULIO*, P. ROVELLI*, F. ANTHONY**, F. ANZUETO***, P. LASHERMES****, G. GRAZIOSI*

*Dipartimento di Biologia, Università di Trieste, Trieste Italy.
**CATIE, Turrialba, Costa Rica.
***PROMECAFE Network, Guatemala City, Guatemala.
****IRD, Montpellier, France

INTRODUCTION

DNA polymorphisms are being applied on an increasingly wider scale in the field of biotechnology, as for example in variety characterisation, agronomic traits identification and marker assisted breeding programs. Some DNA polymorphisms have also been found in *Coffea arabica* through different technical approaches such as RFLP (Lashermes et al, 1996a), RAPD (Orozco-Castillo, 1994; Lashermes et al, 1996b) and AFLP (Lashermes et al, 1999). However, it has been reported that the degree of polymorphism of this species is relatively low (Paillard et al., 1993; Paillard et al., 1996). *C.a.* is an autogamous species and this reproductive strategy undoubtedly leads to a high degree of homozygosity. Furthermore, the genetic base of most coffee cultivars is rather narrow (Bertraud J.and Charrier A, 1988) thus reducing the degree of variability.

Microsatellites are highly polymorphic DNA repetitive sequences. They have been found in all animal and vegetal species so far analysed and are therefore expected to be present in *C. arabica* too. Moreover, they have a relatively high mutation rate which should increase the heterozygosity of coffee, thus compensating for the restricted genetic base. We carried out screening and selection of two genomic libraries enriched in Simple Sequence Repeats (SSR) to identify polymorphic microsatellites, in view of constructing a low density genetic map of *Coffea arabica*.

MATERIAL AND METHODS.

Samples.

The genomic libraries were constructed from DNA of a Caturra plant of the IRD collection. The polymorphism were assessed on a F2 population (Caturra x Ethipia ET30, IRD) and on a limited number of cultivars.

DNA extraction.

Particular care was taken in DNA extraction and purification when preparing the genomic libraries. The methods reported by Murray and Thompson (1980) and Orozco-Castillo et al (1994) were modified as reported by Vascotto et al, 1999).

Preparation of the genome libraries.

Two genomic libraries were prepared, one enriched in $(TG)_n$ and the other in $(ATC)_n$. The basic approach adopted for the enrichment in SSRs has already been reported by Rafalski et al. (1996), Morgante et al. (1998). The relevant selection steps were a first enrichment through magnetic beads conjugated with biotin-strepavidin oligonucleotides complementary to TG and ATC and a second degree selection performed on the clones by colony hybridisation with TG or ATC oligonucleotides. The DNA fragments obtained following the magnetic beads step were cloned into the *EcoRI* site of λ-ZAPII (Stratagene, La Jolla, Calif.).

Sequencing and primer design.

The positive clones were sequenced using a Thermo Sequenase dye terminator cycle pre-mix kit (Amersham Pharmacia Biotec) following the producer's instructions. The reaction products were analysed on an ABI 373A sequencer. The sequences were screened for presence of SSRs and primers were designed for the positive sequences using the on-line programs Primer3 (Whitehead Institute for Biomedical Research, Cambridge, Massachusetts, USA) or *Primers!* (Williamstone Enterprises). A constant tail KS (5'-TCGAGGTCGACGGTATC-3') was added to one of the primers for each primer pair.

Amplification and analysis of the microsatellites.

To avoid using radioactive labels, we developed a three primer system: two primers were locus specific primers while the third primer was complementary to the constant tail of one of the primers and was tagged by either 6-FAM (6-carbossifluoresceine) or JOE (2',7'-dimetossi-4',5'-dicloro-6-carbossifluoresceine). Amplification conditions are reported elsewhere (Vascotto et al, 1999). The amplified fragments were run on sequencing gels in an automatic sequencer ABI 373A and alleles were identified via GENESCAN 672 (Perkin Elmer) software.

RESULTS.

Both genomic libraries contained a large number of clones carrying SSR. Thus the selection procedures proved to have been very effective. Table I reports the numbers of clones analysed in the various successive steps of selection. The clones presently screened constitute approximately 20% of the libraries.

Table I

	N.	%
Clones screened	7.800	
Positive clones	692	9%
Clones sequenced	349	50%
Clones containing SSRs	249	71%
Primer designed	46	18%
Clones containing polymorphic microsatellites	13	25%

Some of the 249 SSR containing clones are still being analysed for primer design, hopefully they will provide more polymorphisms. The 46 primer analysed so far gave the following results: 14 gave no amplification product, 5 were aspecific, 2 gave an unexpected product, 12 were monomorphic and the remaining 13 showed polymorphic bands.

The screening for polymorphism was carried out on a limited number of cultivars and on a F_2 population (12 plants) and accordingly a large number of different alleles was not expected. Nevertheless in our screening population we found from 2 to 5 different alleles. Normally each single plant tested had either one or two alleles, presumably they were either homo or heterozygotes. Only locus E12-3CTG showed 3-4 alleles in the same plant. The result for the locus C2-2CATC can be seen in fig.1 where the parental plant Caturra is heterozygote, plants F_2-1 is apparently homozygote for the 209 bp allele and plant F_2-18 carries two alleles one of which is common to the Caturra progenitor.

DISCUSSION

As clearly shown by the sequencing, the approaches adopted here to generate two genomic libraries enriched in microsatellites were very effective. On the bases of the polymorphisms so far identified we can estimate that the two libraries contain at least 200 polymorphic microsatellites.

The second point of interest is the number of alleles found in each plant. As *C. arabica* is tetraploid and we could expect our primers to recognise the homologous locus on both the two pairs of ancestral homologous chromosomes. This is most probably the case for the locus E12-3CTG, as we obtained 3- 4 alleles in the same plant and the E12-3CTG primer definitely amplifies multiple loci. The remaining primer pairs are apparently specific for only one of the ancestral genomes.

Finally, we wish to point out that, as shown by the analysis of the F_2 segregating population, microsatellites in *Coffea arabica* behave as in any other species: they are codominant and can be easily traced through a family together with possible useful associated traits.

ACKNOWLEDGEMENTS.

This research project has been supported by the European Community grant, INCO-DC Contract n. ERBIC18CT970181

REFERENCES.

Bertraud J., Charrier A. Genetic Resources of Coffea. In "Coffee", Vol. 4, Agronomy, R.J. Clarke and R. Macrae Ed.s, Elsevier Applied Science, London 1988 pp1-41

Lashermes P., Andrzejewski S., Bertrand B., Combes M.C., Dussert S., Graziosi G., Trouslot P., Anthony F.. 1999 Molecular analysis of introgressive breeding in coffee (*Coffea arabica* L.) Theor. Appl. Genet. (in print).

Lashermes P., Cros J., Combes M.C., Trouslot P., Anthony F., Hamon S., Charrier A. 1996a Inheritance and restriction fragment lenght polymorphism of chloroplast DNA in the genus *Coffea* L. *Theor. Appl. Genet.* 93: 626-632.

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

Morgante M, Pfeiffer A, Jurman I, Paglia G, Olivieri AM. 1998. Microsatellite markers in plants. In: Karp A, Isaac PG, Ingram DS (eds) Molecular tools for screening biodiversity. Plants and animals. Chapman and Hall, pp. 206-208, 288-296.

Orozco-Castillo, C., Chalmers K.J., Waugh R., Powell W. 1994 Detection of genetic diversity and selective gene introgression in coffee using RAPD markers. Theor. Appl. Genet. 87: 934-940.

Paillard M, Duchateau N, Petiard V. 1993. Diversité génétique de quelques groupes de caféiers: utilisation des outils moléculaires, RFLP et RAPD. ASIC, 15^e Colloque, Montpellier, pp.33-40.

Paillard M, Lashermes P, Petiard V. 1996 Construction of a molecular linkage map in coffee. Theor Appl Genet 93: 41-47.

Rafalski JA, Vogel JM, Morgante M, Powell W, Andre C, Tingey SV (1996) Generating and using DNA markers in plants. In: Birren B, Lai E (eds) Nonmammalian genomic analysis: a practical guide. Academic Press, London New York, pp.75-134.

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. III SIBAC, Londrina (in print)

SUMMARY.

We are currently screening two genomic libraries enriched for the sequences $(TG)_n$ and $(ATC)_n$. The identification of microsatellites is based on the following strategy: a) sequencing of the DNA inserts; b) design of specific primers for those clones containing Simple Sequence Repeats (SSRs); c) primer test on a panel of cultivars and on an F_2 population. Until now we sequenced about 350 clones; 71% contained a microsatellite. Of the 46 primer pairs designed, 13 gave polymorphic bands. As far as the allele distribution is concerned, we found loci with 2, 3 or 4 alleles. Each single plant carried only one or two alleles, presumably they were either homozygotes or heterozygotes. One single locus showed a peculiar behaviour: all the samples so far analysed carried either 3 or 4 alleles.

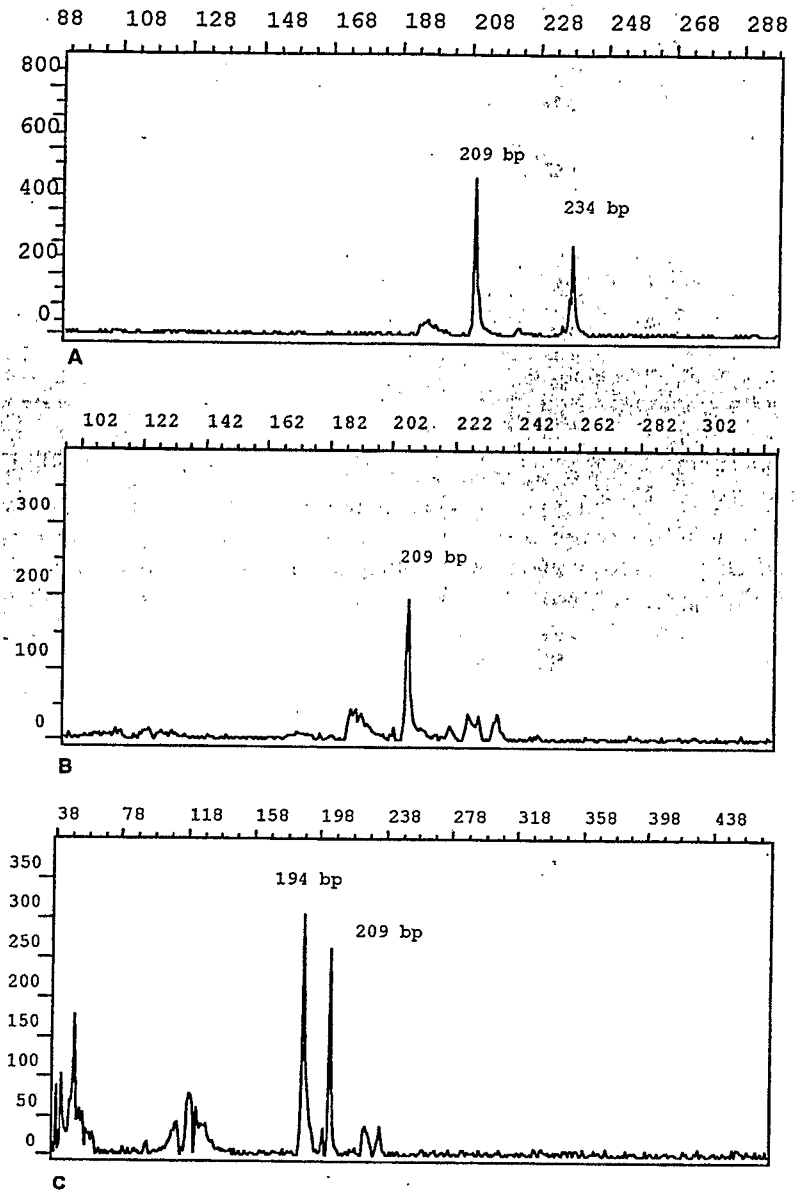

Fig. 1 Electropherograms of the amplification products of locus C2-2CATC: Pane (**A**) Caturra progenitor plant. Two examples of the F_2 segregating population are reported in the panes (**B**) and (**C**). The numbers close to the peaks refer to the allele expressed in bp.

ASIC

Association Scientifique
Internationale
du Café

18ème
COLLOQUE SCIENTIFIQUE INTERNATIONAL SUR LE CAFÉ

18th
INTERNATIONAL SCIENTIFIC COLLOQUIUM ON COFFEE

18.
INTERNATIONALES WISSENSCHAFTLICHES KOLLOQUIUM ÜBER KAFFEE

18°
COLÓQUIO CIENTÍFICO INTERNACIONAL SOBRE EL CAFÉ

Helsinki, 2-6 août 1999

ISBN 2-900212-17-0

DIX-HUITIEME COLLOQUE SCIENTIFIQUE INTERNATIONAL SUR LE CAFE

Helsinki, 2 - 8 août 1999

Association Scientifique Internationale du Café (ASIC)
Siège social : 18, rue de la Pépinière, 75008 PARIS

DIX-HUITIÈME COLLOQUE SCIENTIFIQUE INTERNATIONAL SUR LE CAFÉ

Helsinki, 2-6 août 1999

Introduction

Communications et posters

Chimie

Communications

Posters

Génie alimentaire

Communications

Posters

Effets physiologiques

Communications

Workshop I

Workshop II

Agronomie

Biotechnologie

Communications

Poster

Sélection et amélioration génétique

Communications

Posters

Pratiques agronomiques

Communications

Posters

Technologie après-récolte

Maladies et insectes

Communications

Posters