Theor Appl Genet (2003) 107:1148–1155
DOI 10.1007/s00122-003-1355-4

I. Coulibaly · B. Revol · M. Noirot · V. Poncet · M. Lorieux · C. Carasco-Lacombe · J. Minier · M. Dufour · P. Hamon

AFLP and SSR polymorphism in a *Coffea* interspecific backcross progeny [(*C. heterocalyx* × *C. canephora*) × *C. canephora*]

Received: 10 March 2003 / Accepted: 25 March 2003 / Published online: 23 July 2003

Abstract An interspecific cross (BC 1) involving a species with one of the largest genomes in the *Coffea* genus [*Coffea heterocalyx* (HET), qDNA = 1.74 pg] and a species with a medium-sized genome [*Coffea canephora* (CAN), qDNA = 1.43 pg] was studied using two types of molecular markers, AFLP and SSR. One hundred and eighty eight AFLP bands and 34 SSR primer pairs were suitable for mapping. The total map length was 1,360 cM with 190 loci distributed in 15 linkage groups. The results were compared to those obtained previously on an interspecific BC 1 progeny involving a species with a medium-sized genome (*Coffea liberica* var dewevrei, DEW) and a species with one of the smallest genomes (*Coffea pseudozanguebariae*, PSE). They are discussed relative to three main points: (1) the relevance of the different marker types, (2) the genomic distribution of AFLP and SSR markers, and (3) the relation between AFLP polymorphism and genome size.

Keywords *Coffea* · AFLP · SSR · Genomic distribution

Introduction

The *Coffea* sub-genus includes about 80 taxa originating from the intertropical forests of Africa, Malagasy and Comoro islands. All but one (*Coffea arabica*) are diploid and share the same basic chromosome number ($x = 11$). However, within diploid species, the 2C nuclear DNA content (qDNA) ranges from about 1 pg in *Coffea racemosa* Lour., an East African species, to 1.8 pg in *Coffea humilis* Chev., a West African species (Cros et al. 1995; Noirot et al., in press).

Interspecific crosses between a large range of African species were performed by Louarn (1992). Studies on F1 hybrid fertility yielded two major conclusions: (1) hybrid fertility is higher for crosses involving species from the same geographical group (East African or West and Central African) than between the two groups; (2) the F1 hybrid fertility increases as the qDNA difference decreases.

These results gave rise to some questions on the speciation mode, the direction of DNA changes (DNA gain/loss) and the genomic distribution of the qDNA difference. Considering only this last point, Barre et al. (1998) focused on a cross involving a species with one of the smallest genomes [*Coffea pseudozanguebariae* (PSE), 1.13 pg per 2C genome] and another with a medium-sized one [*Coffea liberica* var dewevrei (DEW), 1.43 pg]. The genome size-difference between PSE and DEW concerns all chromosomes, as shown by in situ hybridization (GISH) of the F1 hybrids, and observed in *Allium* sp. by Narayan (1988). Staining differences indicated the presence of species-specific repeated sequences. Their location in centromeric and sub-telomeric regions were suggested by stronger stain intensity. In backcross-1 hybrids (BC 1), the genome size of the BC 1 hybrid increased as the chromosome number from DEW increased.

Based on this same progeny, Ky et al. (2000) published the first *Coffea* interspecific linkage map. Because of the strategy applied, this has become a reference map for any potential crosses involving parents from the two species considered. Indeed, it was based on species-specific AFLPs corresponding to bands present in all PSE genotypes (representative of the species) and absent in all DEW genotypes. Moreover, although these markers were PSE species-specific, they were either clustered or

Communicated by H.F. Linskens

I. Coulibaly · B. Revol · M. Noirot · V. Poncet · J. Minier · P. Hamon (✉)
IRD, UMR 1097, 911 av Agropolis, BP 64501,
34394 Montpellier cedex 5, France
e-mail: perla@mpl.ird.fr
Tel.: +33-4-67416248
Fax: +33-4-67416222

M. Lorieux
IRD, UMR genome and plant development, CIAT,
Biotechnology Unit, AA 6713, Cali, Colombia

C. Carasco-Lacombe · M. Dufour
CIRAD, avenue Agropolis, 34 398 Montpellier cedex 5, France

not, and distributed throughout the genome. Amplified fragment length polymorphism (AFLPs) were generated with a non-methylation-sensitive restriction enzyme (*Eco*RI). Consequently, both coding and non-coding regions should be equally accessible, and thus species-specific regions could most-likely be tagged with these AFLPs. These results raised the question as to whether the species-specific regions could be associated with DNA gain/loss, i.e. with the qDNA that distinguishes these species. In this work, we mainly applied the strategy developed by Ky et al. (2000) to study an interspecific cross involving one of the largest genomes [*Coffea heterocalyx* (HET), qDNA = 1.74 pg] and a medium-sized genome [*Coffea canephora* (CAN), qDNA = 1.43 pg]. *C. heterocalyx* is a wild species supposedly originating from Cameroon (Stoffelen et al. 1996). This self-fertile species is only represented by a single currently available genotype, and was previously used to study potential self-fertility introgression (Coulibaly et al. 2002). *C. canephora*, the diploid species that is cultivated worldwide, originated from West and Central Africa. Two types of molecular markers, AFLP and simple sequence repeats (SSRs), were used to analyse the segregating progeny. The bands were characterised according to the polymorphism observed at two levels, i.e. within-species, when possible, and between-species. The results are discussed together with previous results obtained with *Coffea* species with different genome sizes.

Materials and methods

Plant material

The mapping population involved backcross progeny derived from an interspecific cross [*C. canephora* (CAN) × *C. heterocalyx* (HET)] × *C. canephora* (CAN).

Plant material was maintained at the IRD station at Man (Côte d'Ivoire). *C. canephora* (CAN, genotype IF182) was hand-pollinated by the only known genotype of *C. heterocalyx* (HET) in 1976. Among the 17 F1 hybrids produced, only one (2CO17) – the most fertile – was pollinated with a mixture of pollen from a dozen *C. canephora* genotypes. The initial BC1 mapping population consisted of 84 plants. Before analysing the segregating population, polymorphism was tested in 12 CAN genotypes, the sole HET and the F1 hybrid.

Fresh coffee leaves were collected, packed in wet newspaper and then sent to Montpellier (France) in an isotherm box. Until lyophilization, they were stored at –80 °C. After 3-days lyophilization, they were stored at 4 °C and 20% relative humidity.

AFLP markers

Genomic DNA was isolated using the Ky et al. (2000) procedure. The AFLP protocol of Zabeau and Vos (1993) and Vos et al. (1995), and adapted for *Coffea* by Ky et al. (2000), was carried out. AFLP markers were named with the three extra selective nucleotides of the *Eco*RI and *Mse*I primers, respectively, followed by a number giving the fragment position on the gel. The numbers are given in order of decreasing molecular weight.

SSR markers

One hundred and thirteen *Coffea* microsatellite sequences produced from *C. arabica* by Rovelli et al. (2000) were retrieved from the Genbank/EMBL databases and 47 sequences from *C. canephora* were kindly provided by CIRAD (Dufour et al. 2001).

Amplification reactions

PCR amplifications were carried out in a Biometra thermocycler. Each 10-μl reaction contained 25 ng of genomic DNA, 1× reaction buffer (Promega), 2.5 mM of $MgCl_2$, 200 μM of each dNTP, 0.1 μM of each primer and 0.05 U/μl of *Taq* polymerase (Promega). Most of the amplifications were performed using a "touchdown" PCR profile consisting of an initial 2-min denaturation at 94 °C followed by five denaturation cycles at 94 °C for 45 s, 1 min primer-annealing at 60 °C (or 55 °C) with the temperature decreasing by one degree 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 (or 50 °C) and 1 min 30 s at 72 °C were performed and followed by a final 8-min elongation at 72 °C. In a few cases, amplification was performed with 35 denaturation cycles at 94 °C for 1 min, 1 min primer-annealing at 58 °C and 1 min elongation at 72 °C.

Pre-screening

Pre-screening was carried out after migration of PCR products on a 1.2% agarose gel, TBE 1× and visualisation by staining with ethidium bromide. Primers giving good amplifications were selected. New PCRs including labelled M13 [0.06 μM IRD700- or IRD800-labeled M13 Primer (MWG-Biotech AG)], were carried out for analyses on an IR^2 Automated DNA Sequencer (LI-COR, Lincoln, Neb, USA).

Detection of microsatellite polymorphism

PCR products were separated on a gel containing a 6.5% polyacrylamide mix from Sciencetech at 55 W constant power and the images were directly visualised on the IR^2 Automated DNA Sequencer. The bands were scored visually. The primers and annealing temperatures used for the progeny analysis are reported in Table 1.

Data analyses

We considered markers only present both in the HET genotype and the F1 hybrid, but absent in all CAN genotypes. In these conditions, we mapped only specific HET bands. This strategy was applied whatever the marker used. Consequently, both SSRs and AFLPs were considered here as dominant markers. HET bands were scored 1 for presence and 0 for absence independently by two persons in order to minimize reading errors. Marker segregation was checked by the χ^2 test for goodness of fit with the expected 1:1 ratio at the 0.01 significant level.

The linkage analysis was carried out using MAPMAKER/EXP 3.0 (Lander et al. 1987) and Mapdisto (available via http://www.mpl.ird.fr/mapdisto/). Given the small chromosome size, it was assumed that the Kosambi (1944) mapping function would be more suitable. The map construction strategy was described by Ky et al. (2000). In a first step, normal and distorted markers were separately treated to form linkage groups using a LOD score of 5 and a theta value of 0.3. In a second step, the merging of linkage groups belonging to these two sets was tested at a LOD score of 3. We considered a linkage group to be a group of at least four linked markers.

Table 1 Primer sequences were designed from accessions retrieved from databases or clone numbers from the CIRAD library. The Tm and locus codes used in this study are given in each case. M13 = CACGACGTTGTAAAACGAC

Accession/ clone no.	Locus code	Fwd sequence (5′>3′) Rev sequence (5′>3′)	Tm	Accession/ clone no.	Locus code	Fwd sequence (5′>3′) Rev sequence (5′>3′)	Tm
AJ250257	M 257	M13-CGACCATTACATTTCACACAC GCATTTTGTTGCACACTGTA	60–55	AJ308856	M 856	M13-CGCTGGTTTTCCTTTTCTGTGTT GGACTGTTCCTAGCTCCTCGT	60–55
AJ250259	M 259	M13-CATCCGTCATAATCCAGCGTC AGGCCAGGAAGCATGAAAGG	60–55	AJ308860	M 860	M13-CGTTTCCATCCACCCTGTC GCTTCTTTCTGGTTTGTTCA	60–55
AJ250260	M 260	M13-CTGATGGACAGGAGTTGATGG TGCCAATCTACCTACCCCTT	60–55	Sati177[a]	M 177	M13-CAGCTTCCGTCGTTATCTCCT GGCAAAGAGGAGTTTCGAGT	60–55
AJ308738	M 738	M13-CACCATGACCAAGACCAAGTA ACAGGGGCGTTCAGTTAT	60–55	004ZB02[b]	M 306		60–55
AJ308742	M 742	M13-CGGCTTCTTGGGTGTCTGTGT CCATTGGCTTTGTATTTCTGG	60–55	004ZH10[b]	M 310		60–55
AJ308779	M 779	M13-CTCCCCCATCTTTTTCTTTCC GGGAGTGTTTTTGTGTTGCTT	60–55	004ZC07[b]	M 312		60–55
AJ308783	M 783	M13-CCCAACTTCGTATGGTTGTCTG TGATAGGAGGCACTTGACACA	60–55	004ZF06[b]	M 314		58
AJ308784	M 784	M13-CTTGCTTGCTTGTTCTGTTAT TGACACGAGAGTTAGAAATGA	60–55	003ZC10[b]	M 324		60–55
AJ308791	M 791	M13-CTGGAGCACGTTTAGTTGAGG CTGCCCTCCGTGAGAAAG	60–55	003ZE10[b]	M 326		60–55
AJ308804	M 804	M13-CTGGGTTGGTTCTGATTTTGG CCTCCCTCTCTCCCTGACTC	58	003ZG10[b]	M 328		60–55
AJ308809	M 809	M13-CAGCAAGTGGAGCAGAAGAAG CGGTGAATAAGTCGCAGTC	60–55	003ZG02[b]	M 329		60–55
AJ308821	M 821	M13-CGGTACATGCTCCTTCCAAGA TCTCTCTGTATCTCTTTACCTTCACC	55–50	003ZF05[b]	M 331		60–55
AJ308822	M 822	M13-CTGCGAAGAGAAAAGTAACCAA TCACTGCAAGCTGTTAAATGAT	60–55	005ZF11[b]	M 333		60–55
AJ308832	M 832	M13-CGCCCTCTAACCAACTCCGTA GATGGGAGGACGAATGAAGA	60–55	contig 106[b]	M 387		60–55
AJ308834	M 834	M13-CGCAGGTATTTGAAGGATGAACC GTGTAGGTGGTGCGATGTGT	60–55	contig 119[b]	M 389		60–55
AJ308837	M 837	M13-CCTCGCTTTCACGCTCTCTCT CGGTATGTTCCTCGTTCCTC	60–55	contig 11[b]	M 390		60–55
AJ308847	M 847	M13-CGCACACATGAAAAAGATGCTTG TGATGGACAGGAGTTGATGG	60–55	contig 178[b]	M 394		60–55

[a] Locus code based on Combes et al. (2000)

[b] Clone number obtained from Dufour et al. (2001), primer sequences for these PCR markers are available from magali.dufour@cirad.fr

Results

Marker selection

In this study, as only one HET genotype was compared to ten CAN genotypes, whatever the molecular marker, specific HET bands represented both species-specific and genotype-specific HET bands. In the latter case, this meant that bands present in the HET genotype might not be present in all other potential HET genotypes.

In the AFLP assay, 57 primer combinations were tested. Irrespective of the species, i.e. HET or CAN, the total number of bands produced was not significantly dependent on the AT content of the 3+/3+ selective nucleotides ($F = 1.77$, $p = 0.19$).

The total number of AFLP bands per individual observed for CAN was higher than for HET (980 vs 780).

Depending on the AFLP primer pairs, between two and 25 specific HET bands could be scored (mean 13). Out of these, 12 couples gave a total of 188/780 (24.1%) HET bands suitable for the mapping (Table 2).

Each couple revealed a high level of within-CAN species polymorphism (51.8–72.2%). In the SSR assay, out of 160 couples tested, 133 and 125 were amplified on agarose gels with CAN and HET, respectively. Using acrylamide gels for the PCR product migration, 70 SSR couples revealed polymorphism within CAN species and/or between HET and CAN species. Out of these, 45 couples were suitable for detecting specific-HET alleles.

Only 34 couples were easily readable on the BC1 progeny. In two cases, the locus was duplicated both in HET and CAN (M328 and M389, Table 1) and two bands were scored. For all loci but three (M306, M389a and M389b), only one band was detected for HET (heterozygosity rate of 3/36 = 8.3%), while no locus appeared to be homozygous for all CAN genotypes.

At this stage, ten non-legitimate offspring were identified, thus reducing the mapping progeny to 74 plants.

Genetic linkage map

Sixteen percent of the 225 bands (15.9% of the AFLPs and 16.6% of the SSRs) displayed distorted segregation at $\alpha = 0.01$, and 75% of these distortions were in favour of CAN bands. A unimodal distribution of segregation rates was obtained regardless of the type of marker.

Linkage analysis was carried out with 188 AFLPs, 36 SSRs and one qualitative trait (presence/absence of self-compatibility as recorded by Coulibaly et al. (2002).

One hundred and ninety markers (84.4%) were mapped on 15 linkage groups, each containing at least four markers. The total map length was 1,360 cM and the mean distance between markers was 7.2 cM (Fig. 1). Skewed loci were mainly distributed in two linkage groups (LG B and LG G). In both cases, the distortion was in favour of CAN alleles.

The distribution within linkage groups appeared to be uniform, except for three linkage groups (LG A, LG B and LG I) which showed clusters.

Discussion

Relevance of the different marker types

Although the number of species-specific bands in *Coffea* species was 20 to 35% lower than the number of polymorphic bands in intraspecific crosses (11.3% in barley, Becker et al. 1995; 14.8% in sorghum Boivin et al. 1999), these results indicated that AFLPs are very interesting markers for interspecific maps when parental data are not available. Such progenies are probably not that rare in perennial long-cycle plants used in mapping programmes.

SSR markers permitted the identification of ten illegitimate progeny. Despite a few cases where the illegitimate AFLP pattern was completely different from the legitimate pattern, SSRs appeared to be more efficient than AFLPs in error identification.

Table 2 AFLP primer combinations analysed in the BC1 [(CAN × HET) × CAN] progeny. The type and number of amplified bands produced per species are noted. X/Y: species comparison – X relative to Y

Primers	Total CAN bands	Total monomorphic CAN bands	Species-specific CAN/HET bands	Intra-CAN polymorphic bands (%)	Total HET bands	Specific HET/CAN bands
E-AAC/M-CAA	108	38	17	70 (64.8%)	65	14
E-AAC/M-CAG	78	23	9	55(70.5%)	50	19
E-AAC/M-CAT	84	31	9	53 (63.1%)	67	18
E-AAC/M-CTA	85	41	14	44 (51.8%)	53	15
E-AAG/M-CAA	141	54	12	87 (61.7%)	102	21
E-AAG/M-CAC	103	42	16	61 (59.2%)	77	25
E-AAG/M-CAG	95	32	9	63 (66.3%)	66	22
E-AAG/M-CTG	84	27	10	57 (67.8%)	50	21
E-AAG/M-CTT	130	52	20	78 (60%)	86	11
E-ACA/M-CAA	107	37	12	70 (65.4%)	72	13
E-ACC/M-CTT	90	25	10	65 (72.2%)	44	14
E-ACT/M-CAA	75	26	9	49 (65.3%)	48	13
Total	1,180	428 (36.3%)	147 (12.5%)	752 (63.7%)	780	206 (26.5%)

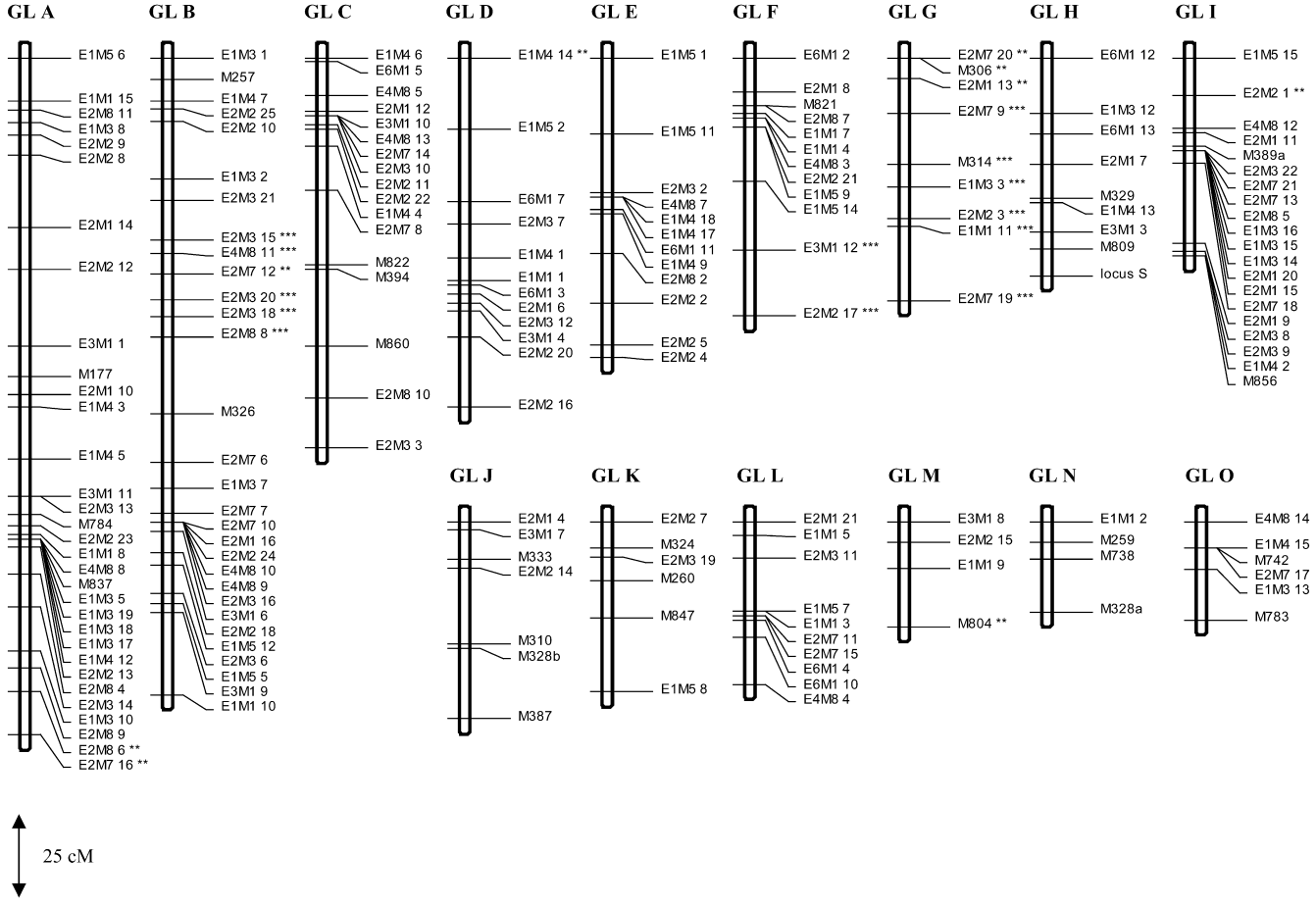

Fig. 1 Genetic linkage map of a [(*C. heterocalyx* × *C. canephora*) × *C. canephora*] interspecific backcross based on combined 160 AFLPs, 29 SSRs and 1 phenotypic marker analysis of 74 BC1 hybrids. Genetic distances are given in centiMorgans (cM) and marker names given to the right of each linkage group (LG). Fifteen LGs covering 1,360 cM are indicated by letters *A* through *O*, ordered by length. Loci marked **, *** and **** deviated significantly from a 1:1 ratio at $P < 0.001$, $P < 0.0001$ and $P < 0.00001$, respectively. AFLP markers are described in terms of the three selective nucleotides used. The *Eco*RI primer is given by the letter *E* and the *Mse*I primer by the letter *M* according to the following correspondance: E1: AAC, E2: AAG, E3: ACA, E4: ACC, E6: ACT, M1: CAA, M2: CAC, M3: CAG, M4: CAT, M5: CTA, M7: CTG, M8: CTT

SSR markers were developed from *C. arabica* and *C. canephora* but they could also amplify genomes of different *Coffea* species, including *C. heterocalyx* (125 couples out of 160 initially tested) and PSE (data not shown). This demonstrates the transferability of SSRs, across different *Coffea* species, even genetically distant species such as CAN, HET and PSE. Our results are in agreement with those of Dirlewanger et al. (2002) who concluded on the cross-species transportability of SSRs derived from *Prunus persica* (peach) to other *Prunus* species, and other genera within the Rosaceae as well as other genera not belonging to the Rosaceae. Similar observations were reported by Guyomarc'h et al. (2002) on the transferability of SSRs from *Triticum tauschii* to bread wheat (*Triticum aestivum*) and diploid related species.

The high level of polymorphism revealed by SSRs has increased their use in genetic mapping. Testolin et al. (2000) obtained 102/251 (40%) primers suitable for mapping a controlled interspecific cross within the *Actinidia* genus. In our conditions, 34/70 (48.6%) could be scored. However, despite the relatively low percentage of SSRs useful for mapping obtained in this study, further research is required in order to produce a saturated SSR *Coffea* map.

Genomic distribution of AFLP and SSR markers

As observed by Ky et al. (2000), AFLPs can be produced in *Coffea* to construct a genetic linkage map. Consequently, they have a dual advantage, since (1) they allow mapping of qualitative morphological traits, as well as QTLs, and (2) they give a core map that can be further used to locate additional loci (SSRs, cloned genes, SCARs); 87.4% of scored AFLPs were mapped and distributed throughout the 15 linkage groups. Similarly, 80.5% of scored SSRs have been mapped and were evenly distributed throughout the map (12 out of the 15 linkage groups). Only one SSR locus was located within

an AFLP cluster on linkage group A. This marker distribution is in line with those reported by Roder et al. (1998), Lespinasse et al. (1999) and Testolin et al. (2000), and in agreement with the assumed random SSR distribution on the genome.

In a previous study, Coulibaly et al. (2002) demonstrated that one locus controlled self-compatibility in HET. The S locus mapped in a terminal position on a short linkage group 14 cM from the nearest locus. The use of SSRs allowed us to reduce this distance since the M809 locus is located between the S locus and the nearest AFLP locus (8.8 cM from the S locus). The above-mentioned SSRs could be used instead of the AFLP marker to facilitate marker-assisted selection and reduce introgression of undesirable traits.

AFLP polymorphism and genome size

AFLPs do not preferentially tag specific genomic regions except for three clusters

In this study, the A/T content of +3/+3 selective nucleotides was neither correlated with the total number of bands obtained nor with the number of polymorphic bands within-CAN species, contrary to reports on soybean (*Glycine max* L. Merr., Keim et al. 1997; Young et al. 1999). Coffee trees differ from soybean by some features which could explain these observations. The soybean genome is large (2.4 pg per 2C genome, Bennett and Leitch 1995) and repeated sequences are abundant (Young et al. 1999). Furthermore, non-coding soybean sequences have a higher AT content than coding sequences (Zhu et al. 1994). *Coffea* has a small genome (one- to two-fold larger than the rice genome) and, although the proportion of repeated sequences has not yet been determined, they most likely represent a lower portion of the genome as in other plants of similar genome size (e.g. about 50% in rice). In such a situation, the production of *Eco*-AFLP bands in coding or non-coding sequences (supposedly AT-rich) should be random, and could explain our results.

The AFLP polymorphism we observed herein was not correlated with the proportion of repeated sequences or with their AT richness.

The number of AFLP loci does not depend on the genome size

CAN has a smaller genome than HET (a qDNA difference of 0.31 pg) but more total bands were amplified (980 vs 780 in a one-by-one comparison). Conversely, despite a qDNA difference of 0.27 pg and considering another 12 polymorphic couples, Ky et al. (2000, original data) observed a similar total number of bands for PSE and DEW (1,103 and 1,087). The same degree of magnitude was noted for the total CAN bands. These results suggest that the total number of amplified bands per species is independent of the genome size (HET has the biggest one, i.e. 1.74 pg, and PSE the smallest, 1.13 pg). This could more likely be explained by the heterozygosity level which mainly depends on the mating system. CAN, PSE and DEW are self-incompatible and actually highly heterozygous (for ten genotypes per species and 73 SSR loci analysed, about 40% of the genotypes were heterozygous, data not shown), while the self-fertile HET is highly homozygous (three out of the 36 scored loci were heterozygous). In these conditions, higher levels of AFLP intraspecific polymorphism are expected and indeed noted within the strictly allogamous species studied (64% for CAN, 63% for PSE and 61% for DEW).

In conclusion, when comparing species with the same mating system but different genome sizes, the total number of bands was similar, suggesting that loci that could be tagged with AFLP markers are not correlated with the genome size differences.

This hypothesis should be further tested by analysing self-incompatible *Coffea* species with the largest genome size (*C. humilis*).

Species-specific AFLP band differentiation is independent of genome-size evolution

Within each CAN, PSE and DEW species, a similar number of total monomorphic AFLP bands was recorded (36.3%, 36.8% and 39%, respectively, Table 3). Moreover, the species-specific band percentages were similar when comparing PSE and DEW genomes (8.9 and 9.5%, respectively, Table 3) while the percentage was higher for CAN in comparison to HET (12.5%). In fact, these discrepancies could be explained by the peculiarity of the cross studied here since one (HET × CAN) hybrid was

Table 3 Characterisation of AFLP bands obtained in different genome comparisons. The total numbers of bands amplified were calculated from the sole HET genotype and ten genotypes for the three species analysed (CAN: *C. canephora*, PSE: *C. pseudozanguebariae* and DEW: *C. liberica* var dewevrei)

Bands	Total monomorphic bands (%)	Species-specific bands (%)	Within-species polymorphic bands (%)
CAN/HET	428/1180 (36.3)	147/1180 (12.5)	752/1180 (63.7)
HET/CAN[a]		*188/780 (24.1)*	
PSE/DEW[b]	463/1259 (36.8)	112/1259 (8.9)	796/1259 (63.2)
DEW/PSE[b]	506/1296 (39)	123/1296 (9.5)	790/1296 (60.9)

[a] Bands present in HET and absent in the ten CAN genotypes tested, representing both species-specific and genotype-specific HET bands
[b] Unpublished data from Ky et al. (2000), calculated from the 12 primer combinations giving the highest species-specific **PSE**/DEW bands

pollinated with a mixture of CAN pollen. Consequently, only one HET genotype was compared to ten CAN genotypes. In these conditions, the number of CAN species-specific bands is inevitably over-estimated as it is for HET species. These results highlighted a reciprocal phenomenon in genome differentiation. Species-specific bands seemed to be similar in number, irrespective of the genome sizes considered and their position on the phylogenetic tree (Lashermes et al. 1997, Cros et al. 1998). These results strongly suggest that radiative speciation has occurred within the *Coffea* genus, a process already proposed by Leroy (1980). Our results suggest that two independent mechanisms are involved: genomic differentiation affecting all species simultaneously, and directional DNA expansion/contraction in genomes over an east-to-west geographical distribution gradient (Noirot et al., submitted).

Genomic distribution of species-specific loci vs genotype-specific loci

Assuming that genome differentiation within the *Coffea* genus is a symmetrical phenomenon involving about 10% of all AFLP bands (as discussed above), about 10% of the 24.1% HET specific bands should correspond to HET species-specific bands. In this case, 14% should be genotype-specific HET bands, which are currently scored for the mapping. In general, their distribution is uniform and covers the entire genome, despite some clusters. Using an intraspecific *Coffea* progeny, Lashermes et al. (2001) observed such a distribution and we hypothesised a similar even-distribution in our map. Thus the question is where are the 10% HET species-specific loci located? In our map, the even-distribution indicated a uniform distribution of these bands since only three clusters were detected but did not represent 10% of the markers used for the mapping. This distribution is in agreement with that of Ky et al. (2000).

Mapping efforts are required to saturate this map, while focusing especially on additional SSRs and SCARs (obtained from bands mapped on other progenies). Once the first complete locus-specific map is available, AFLPs will become less suitable, as already noted in rice, maize and other widely studied plants. However, as species-specific bands could be mapped in (PSE × CAN) × CAN progeny, it would be possible to map PSE species-specific monomorphic bands partially common to those mapped by Ky et al. (2000) in (PSE × DEW) × DEW progeny. In such conditions, recombination rates could be compared for linked couples of loci while gaining insight into the genome organisation of the different *Coffea* species.

Acknowledgements We thank the entire IRD team in Man (Côte d'Ivoire) for valuable collaboration and technical assistance.

References

Barre P, Layssac M, D'Hont A, Louarn J, Charrier A, Hamon S, Noirot M (1998) Relationship between parental chromosomic contribution and nuclear DNA content in the coffee interspecific hybrid *C. pseudozanguebariae* × *C. liberica* var 'dewevrei'. Theor Appl Genet 96:301–305

Becker J, Vos P, Kuiper M, Salamini F, Heun M (1995) Combined mapping of AFLP and RFLP markers in barley. Mol Gen Genet 249:65–73

Bennett MD, Leitch IJ (1995) Nuclear DNA amounts in Angiosperms. Ann Bot 76:113–176

Boivin K, Deu M, Rami JF, Trouche G, Hamon P (1999) Towards a saturated sorghum map using RFLP and AFLP markers. Theor Appl Genet 98:320–328

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

Coulibaly I, Noirot M, Lorieux M, Charrier A, Hamon S, Louarn J (2002) Introgression of self-compatibility from *Coffea heterocalyx* to the cultivated species *Coffea canephora*. Theor Appl Genet 105:994–999

Cros J, Combes MC, Chabrillange N, Duperray C, Monnot des Angles A, Hamon S (1995) Nuclear DNA content in the subgenus *Coffea* (Rubiaceae): inter- and intra-specific variation in African species. Can J Bot 73:14–20

Cros J, Combes MC, Trouslot P, Anthony F, Hamon S, Charrier A, Lashermes P (1998) Phylogenetic analysis of chloroplast DNA variation in *Coffea* L. Mol Phylog Evol 9:109–117

Dirlewanger E, Cosson P, Tavaud M, Aranzana J, Poizat C, Zanetto A, Arus P, Laigret F (2002) Development of microsatellite markers in peach [*Prunus persica* (L.) Batsch] and their use in genetic diversity analysis in peach and sweet cherry (*Prunus avium* L.). Theor Appl Genet 105:127–138

Dufour M, Hamon P, Noirot M, Ristrerucci AM, Brottier P, Vico V, Leroy T (2001) Potential use of SSR markers for *Coffea* spp. genetic mapping. 19th Int Sci Colloquium on Coffee, Trieste, Italy

Guyomarc'h H, Sourdille P, Edwards J, Bernard M (2002) Studies of the transferability of microsatellites derived from *Triticum tauschii* to hexaploid wheat and to diploid related species using amplification, hybridization and sequence comparisons. Theor Appl Genet 105:736–744

Keim P, Schupp JM, Travis SE, Clayton K, Zhu T, Shi L, Ferreira A, Webb DM, Zhu T, Shi LA (1997) A high-density soybean genetic map based on AFLP markers. Crop Sci 37:537–543

Kosambi DD (1944) The estimation of map distance from recombination values. Ann Eugen 12:172–175

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

Lander ES, Green P, Abrahamson J, Barlow A, Daly MJ, Lincoln S, Newburg L (1987) MAPMAKER: an interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. Genomics 1:174–181

Lashermes P, Combes MC, Trouslot P, Charrier A (1997) Phylogenetic relationships of coffee-tree species (*Coffea* L.) as inferred from ITS sequences of nuclear ribosomal DNA. Theor Appl Genet 94:947–955

Lashermes P, Combes MC, Prakash NS, Trouslot P, Lorieux M, 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

Leroy JF (1980) Evolution et taxogenése chez les caféiers: hypothése sur l'origine. CR Acad Sci France 291:593–596

Lespinasse D, Rodier-Goud M, Grivet L, Leconte A, Legnate H, Seguin M (1999) A saturated genetic linkage map of rubber tree (*Hevea* spp.) based on RFLP, AFLP, microsatellite and isozyme markers. Theor Appl Genet 100:127–138

Louarn J (1992) La fertilité des hybrides interspécifiques et les relations génomiques entre caféiers diploïdes d'origine africaine (Genre *Coffea* L. sous-genre *Coffea*). Montpellier II, Sciences et Techniques du Languedoc, Montpellier, France

Narayan RKJ (1988) Constraints upon the organisation and evolution of chromosomes in *Allium*. Theor Appl Genet 75:319–329

Noirot M, Poncet P, Barre P, Hamon P, Hamon S, de Kochko A (2003) Genome size variations in diploid African coffea species. Ann Bot (in press)

Roder MS, Korzun V, Wendehake K, Plaschke J, Tixier MH, Leroy P, Ganal MW (1998) A microsatellite map of wheat. Genetics 149:2007–2023

Rovelli P, Mettulio R, Anthony F, Anzueto F, Lashermes P, Graziosi G (2000) Microsatellites in *Coffea arabica* L. In: Sera T, Soccol CR, Pandey A, Roussos S (eds) Coffee biotechnology and quality. Kluwer Academic Publishers, The Netherlands, pp 123–133

Stoffelen P, Robbrecht E, Smets E (1996) *Coffea* (Rubiaceae) in Cameroon: a new species and a nomen recognized as species. Belgian J Bot 129:71–76

Testolin R, Huang WG, Lain O, Messina R, Vecchione A, Cipriani G (2000) A kiwifruit (*Actinidia* spp.) linkage map based on microsatellites and integrated with AFLP markers. Theor Appl Genet 103:30–36

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

Young WP, Schupp JM, Keim P (1999) DNA methylation and AFLP marker distribution in the soybean genome. Theor Appl Genet 99:785–792

Zabeau M, Vos P (1993) Selective restriction fragment amplification: a general method for DNA fingerprinting. European patent application EP534858A1

Zhu T, Schupp JM, Oliphant A, Keim P (1994) Hypomethylated sequences: characterization of the duplicate soybean genome. Mol Gen Genet 244:638–645