Journal of Heredity 2005:96(1):59–65
doi:10.1093/jhered/esi013
Advance Access publication December 15, 2004

Base Composition of *Coffea* AFLP Sequences and Their Conservation Within the Genus

V. Poncet, P. Hamon, M.-B. Sauvage de Saint Marc, T. Bernard, S. Hamon, and M. Noirot

From UMR 1097 Diversité et Génome des Plantes Cultivées, Centre IRD, 911 avenue Agropolis, BP 64501, 34394 Montpellier Cedex 5, France.

Address correspondence to V. Poncet at the address above, or e-mail: poncet@mpl.ird.fr.

Amplied fragment length polymorphism (AFLP) is often used for genetic mapping and diversity analysis, but very little information is currently available on their sequence characteristics. Species-specic sequences were analyzed from a single *Coffea* genome (*Coffea pseudozanguebariae*) associated with clustered or nonclustered AFLP loci of known genetic position. Compared with the expressed sequence tag (EST) sequence composition, their AT content exhibited a bimodal distribution with AT-poor sequences corresponding mainly to putative coding sequences. AT-rich sequences, apart from the EST distribution, were usually clustered on the genetic map and might correspond to noncoding sequences. Conversion of these AFLP markers into sequence-characterized amplied region (SCAR) anchor markers allowed us to assess sequence conservation within *Coffea* species with respect to species relatedness.

Polymerase chain reaction (PCR)-based amplified fragment length polymorphism (AFLP) is an efficient tool for detecting reproducible and reliable polymorphism. It has been used both to construct high-density genetic maps and to saturate specific regions in plants such as alfalfa (*Medicago truncatula*) (Thoquet et al. 2002), cowpea (*Vigna unguiculata*) (Ouedraogo et al. 2002), sunflower (*Helianthus annuus*) (Mokrani et al. 2002), and sorghum (*Sorghum bicolor*) (Menz et al. 2002).

Amplified fragment length polymorphism is based on the ligation of adapters to genomic restriction fragments followed by PCR-based amplification with adapter-specific primers. Like restriction fragment length polymorphisms (RFLPs), the majority of AFLP fragments define single loci (Qi and Lindhout 1997; Vos et al. 1995). However, AFLP markers generated with the methylation-insensitive enzyme *Eco*RI in combination with *Mse*I are often not uniformly distributed over chromosomes. Strong clustering around the centromere region has been reported in tomato (*Lycopersicon* ssp.), through the analysis of an integrated RFLP/AFLP map (Haanstra et al. 1999), and also in barley (*Hordeum vulgare*) (Qi et al. 1998) and maize (*Zea mays*) (Vuylsteke et al. 1999). Consequently it has been suggested that the targeted regions could differ from the rest of the genome with respect to their sequence, methylation, and transcription level, A + T/C + G ratio, and/or recombination rate (Young et al. 1999).

Very little information is currently available on the sequence nature and characteristics of these AFLP bands. Indeed, for most applications, the AFLP protocol does not require preliminary genomic sequencing or cloning. Sequencing of AFLP bands has mainly been carried out in studies requiring conversion of AFLP markers into sequence-characterized amplified regions (SCARs); for example, in *Brassica* (Negi et al. 2000), soybean (*Glycine* ssp.) (Hayashi et al. 2000), carrot (*Daucus carota*) (Bradeen and Simon 1998), apple (*Malus floribunda*) (Xu et al. 2001), and willow (*Salix viminalis*) (Gunter et al. 2003). Most of these studies were conducted with the goal of enhancing marker-assisted selection, and involved identification and conversion of AFLP markers tightly linked to genes of interest. To our knowledge, no published studies have involved an analysis of AFLP sequences distributed all over a single genome.

The first *Coffea* interspecific linkage map was based on a backcross progeny [(*C. pseudozanguebariae* (PSE) × *C. liberica* var. Dewevrei (DEW)) × DEW] (Ky et al. 2000). The two parental species are diploid with the same chromosome number ($2n = 2x = 22$). A special strategy was applied, which involved mapping of AFLP bands specific to PSE species relative to DEW. This linkage map comprised 181 loci distributed over 14 linkage groups covering 1144 cM. Although these markers were PSE specific, they were distributed throughout the genome, which is a common feature on AFLP plant linkage maps.

In this article we analyze specific sequences of the PSE genome associated with different types of clustered and nonclustered AFLP bands. SCAR markers were derived from these PSE AFLP sequences in order to evaluate their conservation across *Coffea* species. The corresponding primer pairs were tested on other species to determine whether they could amplify homologous sequences.

Materials and Methods

Plant DNA

Amplified fragment length polymorphism profiles were obtained for PSE and PSE × DEW F_1 hybrids. Specific SCAR primers were tested on genomic DNA of *Coffea arabica* (ARA), *Coffea canephora* (CAN), *Coffea congensis* (CON), *Coffea eugenioides* (EUG), *Coffea heterocalyx* (HET), *Coffea liberica* var. Dewevrei (DEW), *Coffea millotii* (MIL), *Coffea* sp. Moloundou (MOL), *C. pseudozanguebariae* (PSE), and *Arabidopsis thaliana* (THA). All coffee trees were maintained in tropical conditions in a greenhouse in Montpellier, France. Genomic DNA was extracted from lyophilized leaves as described by Ky et al. (2000).

Isolating and Cloning AFLP Fragments

Amplified fragment length polymorphism reactions, electrophoresis, and detection were performed as described in Ky et al. (2000). Selective amplification was performed using selective 3+/3+ primers (Table 1). Selected bands were excised from the dried polyacrylamide gel, eluted in 100 µl sterile ddH_2O, and kept on boiling water for 5 min to release DNA from the gel. After the gel debris was spun down, the DNA-containing supernatant was transferred into a new tube and ethanol precipitated. AFLP fragments were resuspended in 50 µl sterile ddH_2O and used as a template for reamplification with the appropriate preselective or selective AFLP primer combination. The standard AFLP PCR program was supplemented with an elongation step of 30 min at 72°C in order to add the 3′A overhang to the PCR products. The products were purified using E.Z.N.A. Cycle-Pure (OMEGA Bio-Tek) and cloned onto pCR4-Topo from the TOPO TA cloning kit (Invitrogen) according to the manufacturer's instructions. Plasmid DNA was extracted using the Plasmid Mini Preparation kit (Qiagen).

The general strategy applied to select specific sequences underlying AFLP bands was based on Xu et al. (2001) and proved to be reliable and robust. In particular, after PCR amplification using the corresponding AFLP primers, inserts of five colonies were run on polyacrylamide gels along with the initial AFLP amplification to check that the cloned fragments and the original band shared the same position. The selected clones were then prepared for sequencing using the Plasmid Mini Preparation kit (Qiagen).

Sequence Analysis

The inserts were sequenced by MWG-Biotech. Sequences were edited and analyzed using the DNASTAR software package (Lasergene, Madison, WI). Similarities between sequences were tested with the DNASTAR SeqMan module using the "Assembling" option.

To compare these sequences to public database sequences, similarity searches were performed using BLAST algorithms (http://www.ncbi.nlm.nih.gov:80/BLAST/) (Altschul et al. 1997). BLASTx was used to compare sequences to protein sequences, while BLASTn was used to compare sequences to the nonredundant DNA database sequences and EST database sequences. The search was conducted in June 2004 using the default parameters. Similarity searches to *C. canephora* EST sequences (unpublished data) derived from fruit copy DNA (cDNA) (5706 sequences) or leaf cDNA (2709 sequences) were also performed. Similarities to known sequences were considered significant when the probability values were less than 10^{-4}.

The base compositions of AFLP sequences were then compared to those of 103 *Coffea* EST sequences. The EST sequences were chosen from a *C. canephora* EST library derived from fruit cDNA and they showed a similarity with known protein sequences with an *E*-score value less than 8e-29 (BLASTx). No EST was predicted to be from plastids.

SCAR Primer Pairs

Sequence-characterized amplified region primers were designed with the public Primer 3 software package (Rozen and Skaletsky 2000) and synthesized by MWG-Biotech. When possible, primers were designed to retain the original restriction enzyme flanking site, the three selective base pairs, and the next internal sequences. The primer sizes ranged from 17 to 27 bases. One primer of every pair was 5′-tailed with the M13 sequence 5′-CACGACGTTGTAAAACGAC-3′ for subsequent use of labeled M13 primer.

Each 15 µl PCR contained 25 ng of genomic DNA, 1× reaction buffer (Promega), 2.0 mM $MgCl_2$, 200 µM of each dNTP, 0.2 µM of each primer, and 0.05 U/µl of *Taq* polymerase (Promega). The amplifications were performed following a "touchdown" PCR profile consisting of an initial 2 min denaturation at 94°C for 45 s, 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 1°C 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. The primers and annealing temperatures are shown in Table 2.

For visualization of the PCR product on an IR^2 Automated DNA Sequencer (LI-COR, model 4200L-2, Lincoln, NE), an M13 primer coupled to an IRD700 or IRD800 infrared tag (MWG-Biotech AG) at 0.2 µM was included in the PCR and the concentration of the corresponding M13-tailed primer was decreased to 0.02 µM (Oetting et al. 1995). PCR products were electrophoresed on a 6.5% KBPlus (LI-COR, 827-05607) 25 cm gel. The gel image was processed by SAGA GT to estimate the size of the sample bands according to a 50–350 bp size standard (LI-COR, 829-05343, 829-05344).

Results

Selection and Sequencing of AFLP Bands

Twenty-three AFLP loci were chosen on the reference map (Ky et al. 2000) to be distributed either inside (9) or outside

Table 1. Sequenced AFLP fragments: designation on the (PSE × DEW) × DEW linkage map (Ky et al. 2000), AT composition, and similarity searches with different BLAST algorithms

AFLP selective primer combinations	No. of the excised band	Linkage group	Cluster	Sequence	Size (bp)	A–T (%)	Blast		
							Blastn[a]	Blastx[a]	tBlastx[a]
*Eco*RI-AAC *Mse*I-CAT	1-1	G	No	+	289	67.8	Nh	Nh	Nh
	1-4	L	Yes	−			—	—	—
	1-5	D	No	+	155	61.9	Nh	Nh	Nh
	1-9	C	No	−			—	—	—
	1-10	K	Yes	+	66	71.2	Nh	Nh	Nh
*Eco*RI-AAC *Mse*I-CTA	2-3	H	No	+	171	57.9	Nh	Nh	Nh
	2-4	A	No	+	125	75.2	Nh	Nh	Nh
	2-6	G	Yes	+	109	73.4	Nh	Nh	Nh
	2-7	F	No	−			—	—	—
*Eco*RI-AAC *Mse*I-CAC	3-1	F	No	+	233	54.1	Nh	*E*-value = 1e-07 (NP_563868) *A. thaliana*, expressed protein	*E*-value = 3e-05 (AY116950) *A. thaliana*, mRNA
	3-4	A	No	+	184	61.4	*E*-value = 1e-10 (AJ275893) *Ceratitis capitata*, unknown repetitive DNA fragment	Nh	Nh
	3-5	H	Yes	+	173	62.4	Nh	Nh	Nh
	3-6	F	No	+	99	71.7	Nh	Nh	Nh
	3-7	J	Yes	+	99	73.7	Nh	Nh	Nh
	3-8	A	No	+	63	68.2	Nh	Nh	Nh
	3-9	F	No	+	61	49.2	*E*-value = 1e-19 (AE005652) *E. coli*, chromosomal region	Nh	*E*-value = 8e-04 (AP002568) *E. coli*, complete genome
*Eco*RI-AAC *Mse*I-CAG	4-1	E	No	+	135	56.3	Nh	*E*-value = 5e-06 (BAB10716) *A. thaliana*, unnamed protein	*E*-value = 2e-08 (BX830455) *A. thaliana*, cDNA complete sequence
	4-2	A	No	+	100	58.0	Nh	Nh	Nh
	4-3	B	Yes	+	90	58.9	Nh	Nh	Nh
*Eco*RI-AAC *Mse*I-CTG	6-5	I	Yes	+	155	64.5	Nh	Nh	Nh
*Eco*RI-AAG *Mse*I-CTC	7-2	?		+	274	62.4	Nh	Nh	Nh
*Eco*RI-AAG *Mse*I-CTG	8-2	B	No	+	190	66.8	Nh	Nh	Nh
	8-4	?		+	51	54.9	Nh	Nh	Nh
*Eco*RI-AAG *Mse*I-CTT	9-2	?		+	156	62.2	Nh	Nh	Nh
	9-4	G	Yes	+	74	74.3	Nh	Nh	Nh
*Eco*RI-ACA *Mse*I-CTA	14-8	G	Yes	+	135	56.3	*E*-value = 1e-09 (AY102435) *C. arabica*, microsatellite sequence	Nh	*E*-value = 5e-07 (AY102435) *C. arabica*, microsatellite sequence

[a] Results of the similarity searches with the corresponding *E*-value; Nh = no hit, no significant similarity.

(14) the clusters over the first 11 linkage groups (Table 1). A locus was considered as being in a cluster when the mean distance to its two neighboring markers was less than 4 cM. Three additional AFLP segregating markers had an unknown location, that is, they did not link to any linkage group. After amplification with the corresponding AFLP primer combination, AFLP bands were excised, reamplified, and cloned. Twenty-three were all successfully sequenced.

Table 2. Amplification with SCAR primers across *Coffea* species and *A. thaliana* and PCR product size

Locus	Primer sequences (5′-3′)	Anneal temp. (°C)	PSE expected PCR product size	PSE[a]	DEW	CAN	CON	EUG	HET	MIL	MOL	ARA	THA	Total +
Scar-1-5	ACATTTTGTTGCTCAATTATTCTCA	60–55	151	151		151	151	151	151	151	151	151		
	M13-AATTCAACTCCTACCCAAACACA			+	−	+	+/−	+/−	+	+/−	+/−	+	−	4
Scar-1-10	GAATTCAACTATTTGTTACTTTACCTG	60–55	62	62	54	54	54	54	54	54	54	54+62	54	
	M13-CATCAACATGTTTTTCCAATG			+	+/−	+	+	+	+	+	+	+	+	9
Scar-2-3	M13-TTCAACGAATATAAGAGCAA	55–50	168	168	168									
	TTAACTATGGTAATTGGGTTAG			+/−	+/−	−	−	−	−	−	−	−	−	0
Scar-2-4	AACTAACATCGGATAATAAATAAAC	55–50	121	121					121+141					
	M13-ATTCAACGTTTGCTTGAAA			+	−	−	−	−	+	−	−	−	−	2
Scar-2-6	M13-ATGCACTACATTTACATCCTA	55–50	85	85		85	85	85		85	85	85		
	AATTCAACATATGAAAAGAAAG			+	−	+	+	+	−	+	+	+/−	−	6
Scar-3-1	AACACCATTTTGACAAACACAAA	55–50	230	230		230	230	230	230		230			
	M13-AATTCAACCTGCATGAACTAAGC			+	−	+/−	+/−	+/−	+	−	+	−	−	3
Scar-3-7	CAACTTGTATTGGTGGTAGAATTT	60–55	78	78		78	78/80	78	78	78	78	78		
	M13-AGCCACCAAAATAACAAATGTA			+	−	+	+	+	+	+	+	+	−	8
Scar-3-8	M13-ATTCAACTGTCAAAGAGTTTT	55–50	61	61	61	61	61	61	61	61	61	61		
	TTAACACGAGATAGATCATACTT			+	+	+	+	+	+	+	+	+	−	9
Scar-4-1	M13-CAACGACTAAAATGTCGAG	55–50	128	128		128	128	128	128	128	128	128		
	AACAGGGAAATCCAAGA			+	−	+	+	+	+	+	+	+	−	8
Scar-4-2	GAATTCAACCAGTAGGATACTCA	60–55	96	96	96	96	96	96	96	96	96	96		
	M13-CAGCAGTGATGCACTTGAC			+	+/−	+/−	+	+	+	+	+	+	−	7
Scar-4-3	M13-GAATTCAACACTGGCAGAAG	60–55	90	90	90	90	90	90	90	90	90	90		
	TTAACAGACGAGCAGGAGAAG			+	+	+	+	+	+	+/−	+	+	−	8
Scar-6-5	M13-AACTGCCCATGTAACCAAAGA	60–55	148	148	149	149	149/150	149+156	149	150		149+156		
	CAACCAAATCCTCCTCTACCC			+	+	+	+	+/−	+	+	−	+	−	7
			Total +	11	3	8	8	7	10	7	8	8	1	

[a] PCR product size (number of bp) and amplification efficiency, "−": no PCR product; "+" clear amplification bands; "+/−" weak amplification.

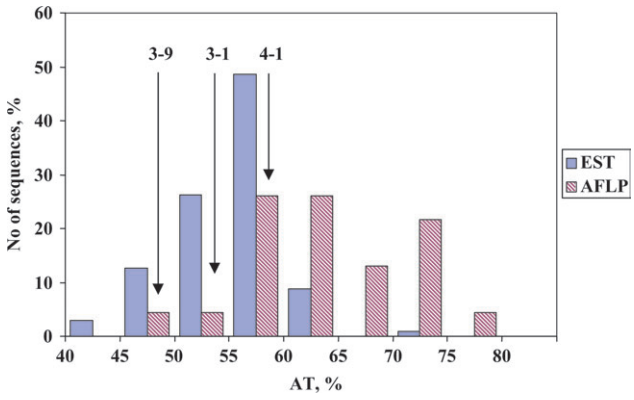

Figure 1. Compositional distribution of AFLP and EST sequences.

Similarity Searches

No similarities between PSE sequences or with *Coffea* EST sequences were noted. Comparison of PSE sequences to the databases maintained at GenBank and EMBL revealed significant similarity with only five sequences: 3-1, 3-4, 3-9, 4-1, and 14-8 (Table 1). In particular, some similarity (*E*-value = 5e-06) between AFLP sequence 4-1 and a protein from *A. thaliana* (accession no. BAB10716) was observed. Likewise, AFLP sequence 3-1 showed similarity (*E*-value = 1e-07) to an *A. thaliana*-expressed protein (accession no. NP_563868). AFLP sequence 3-9 was found to be similar (*E*-value = 1e-19) to a portion of an *Escherichia coli* chromosome (accession no. AE005652), but without significant similarity to any known protein.

Compositional Distribution of AFLP Sequences

The base compositions of the AFLP sequences were calculated and compared to those of *Coffea* EST sequences (Figure 1). The EST AT content distribution was unimodal, with a mean of 55.0%. The AFLP sequences showed an AT content distribution ranging from 49.2% to 75.2% with two modes. Taking 65% AT as the cutting point, the first mode, with a mean AT content of 58.6%, fit with the EST sequence distribution. The second mode showed significantly higher AT content, with a mean of 71.4%.

ALFP sequences 3-1 and 4-1, which showed some similarity with *A. thaliana* proteins, had AT contents of 54.1% and 56.3%, respectively; that is, close to the mean AT content of EST sequences (Figure 1). When considering the nine AFLP sequences associated to clusters, seven of them had an AT content greater than 60%. AFLP sequence 3-4, which showed a similarity to repetitive DNA, had an AT content of 61.4%.

SCAR Primer Transferability Across *Coffea* Species

The specific primers designed from the PSE sequences were tested on genomic DNA of each coffee species (ARA, CAN, CON, EUG, HET, DEW, MIL, MOL, PSE) and *A. thaliana* (THA). Eleven primer pairs efficiently amplified at least one *Coffea* species, and SCAR primers such as Scar-3-8 or Scar-4-3 gave good amplification with the nine *Coffea* species tested. *A. thaliana* was only amplified with Scar-1-10 primers.

For PSE, the PCRs resulted in amplification of a single fragment of the expected size. For species other than EUG and ARA, the PCR also gave a single band with a size similar to that of the PSE product. When ARA was amplified with the Scar-1-10 or Scar-6-5 primer pairs, two PCR products were revealed that differed by about seven nucleotides. The same situation was observed for EUG when amplified with the Scar-6-5 primers.

In some species, blanks were found that might have indicated the presence of null alleles. However, amplification success was independent of the species considered.

Discussion

GC Composition of AFLP Sequences

The distribution of the AFLP sequence composition was found to be bimodal in this study (Figure 1). One mode, CG-rich AFLP sequences, could mainly correspond to coding sequences for two reasons: (1) it fits well with the distribution of *Coffea* EST sequences, and (2) 3-1 and 4-1 AFLP sequences showing similarity with *A. thaliana* protein sequences had a GC content of 45.9% and 43.7%, respectively, included in the EST distribution.

Tested EST sequences (regulatory sequences + exons) could be considered as a good representation of *Coffea* coding sequences. Indeed, their compositional distribution showed a clear unimodal distribution that would correspond to the distribution reported for dicot genes, whereas Poaceae coding sequences are characterized by distributions that are at least bimodal (Carels and Bernardi 2000). Moreover, the *Coffea* genus, with an EST composition mean of 45.0% GC, was very similar to Solanaceae, the phylogenetically closest family (see, e.g., Fulton et al. 2002).

The other distribution mode of the AFLP sequence composition consisted of sequences with AT content greater than 65%, which is significantly different from the distribution of *Coffea* EST sequences. Consequently these sequences could mainly correspond to noncoding sequences. For example, the 3-4 AFLP sequence, with an AT content of 61.4%, higher than the mean AT content of EST sequences, showed a great degree of similarity to an unknown repetitive DNA fragment from the Mediterranean fruit fly (*Ceratitis capitata*).

Clustering of AFLP Markers

The analyzed AFLP sequences were selected according to their location on the *Coffea* interspecific linkage map of Ky et al. (2000), where eight *Eco*RI/*Mse*I AFLP clusters had been observed. Such clustering was also noted in other plant AFLP linkage maps; for example, in potato (*Solanum tuberosum)* (van Eck et al. 1995), barley (Powell et al. 1997),

soybean (Keim et al. 1997), *Arabidopsis* (Alonso-Blanco et al. 1998), and maize (Vuylsteke et al. 1999).

As reported by Young et al. (1999), AFLP markers derived from *Eco*RI + *Mse*I selective primers are expected to form clusters in some specific genomic regions. Since *Eco*RI is insensitive to CpNpG methylation, AFLP markers developed with this restriction enzyme may identify polymorphism in hypermethylated regions of the genome with low recombination rates, such as in centromeres. These regions are not covered by traditional RFLP maps (Haanstra et al. 1999) and were found to correspond to centromeric regions in maize (Vuylsteke et al. 1999) and barley (Qi et al. 1998) genomic maps. In *Coffea*, clustering was observed on a genetic map (Ky et al. 2000) without information on the physical location of the AFLP markers. It was impossible to determine whether this resulted from a reduced recombination rate around centromeres, a higher frequency of AFLP markers in these regions, or a combination of the two.

To further explain the observed clustering, Peters et al. (2001) suggested that such polymorphism was more often encountered in centromeric regions because they are rich in noncoding DNA with theoretically less constrained sequences. In addition, repetitive sequences typical of centromeric regions may produce a number of AFLP markers representing variants of a single repeat. In the present study, we selected AFLP loci throughout the genome and none of them were in the same cluster. No sequence similarity between AFLP bands was observed, which could suggest that the AFLP sequences, even though they are located in different clusters, could correspond to repetitive sequences with a common origin.

However, AFLP fragments associated with clustered loci tended to have an AT content higher than the mean coding sequence AT content. These preliminary results are in accordance with the correspondence between centromeric AT-rich noncoding sequences and AFLP clusters.

Sequence Conservation Across *Coffea* Species

The 14 SCAR primer sets derived from PSE AFLP bands were assessed for their ability to amplify homologous sequences in nine *Coffea* species and in *A. thaliana*. SCAR primers would be transferable to related species, mainly when conserved sequences are used for primer design. In contrast to SCAR markers derived from coding sequences such as cDNA RFLP clones, primer binding sites are less likely to be conserved across species for SCAR markers derived from AFLP. In this context, our transferability results obtained are relatively favorable, since any *Coffea* species was amplified by a minimum of 58% primer pairs.

Twelve primer pairs gave amplification across *Coffea* species and one even amplified *A. thaliana*. Primer pairs amplified a single product (only one band) in all species, except Scar-6-5 in *C. arabica* and *C. eugenioides* and Scar-1-10 in *C. arabica*. While *C. eugenioides* is a diploid species, *C. arabica* is known to be an amphidiploid formed by hybridization between *C. eugenioides* and *C. canephora*, or ecotypes related to these diploid species (Berthou et al. 1983; Lashermes et al. 1999). In this study, the amplification patterns obtained for *C. arabica* were a juxtaposition of patterns of the putative diploid parental species.

The lack of amplification could have been due to mutations in the primer binding sites or to any mutational event that could prevent amplification, such as major insertion/deletion between primer sites or an absence of the locus in the genome. The initial polymorphism observed for the AFLP pattern—the presence of a band for PSE, the absence for DEW—was not conserved when we considered some of the SCAR amplifications (Table 2). Indeed, although the primers were designed to amplify the largest part of each AFLP sequence, they might not include the site of variation.

The amplification success was found to be independent of the species considered and did not reflect the between-species relatedness as defined by Lashermes et al. (1997). A similar situation has been reported for microsatellites across *Coffea* species (Poncet et al., in press) and is in agreement with the recent radiative speciation scenario noted within the *Coffea* genus (Cros et al. 1998; Leroy 1980).

Available linkage maps for *Coffea* species are based on random amplified polymorphic DNA (RAPD), RFLP, AFLP, and microsatellite markers (Coulibaly et al. 2003; Ky et al. 2000; Lashermes et al. 2001; Paillard et al. 1996). However, no RFLP markers were common to the different maps and only a few microsatellites have been mapped so far.

The set of SCAR primers obtained in this study might be suitable for use across the *Coffea* genus. During the development of SCARs, the loss of original polymorphism often leads to experimental challenges (Shan et al. 1999). Indeed, the original polymorphism was found to be mainly located at the restriction sites, and it was not always possible to design primers covering the corresponding region. Nevertheless, the amount of variation they could reveal should be further studied; for example, through enzymatic digestion of PCR products (cleaved amplified polymorphic sequences [CAPS]).

In conclusion, these SCAR primers, together with the currently developed microsatellites (Poncet et al., in press), will facilitate assembly of a panel of "anchor" PCR-based markers for comparative mapping studies in coffee trees and for marker-assisted selection.

Acknowledgments

This work was partly supported by a grant from EEC ICA4-CT2001-10068.

References

Alonso-Blanco C, Peeters AJ, Koornneef M, Lister C, Dean C, van den Bosch N, Pot J, and Kuiper MT, 1998. Development of an AFLP based linkage map of Ler, Col and Cvi *Arabidopsis thaliana* ecotypes and construction of a Ler/Cvi recombinant inbred line population. Plant J 14:259–271.

Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, and Lipman DJ, 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 25:3389–3402.

Berthou F, Mathieu C, and Vedel F, 1983. Chloroplast and mitochondrial DNA variation as indicator of phylogenetic relationships in the genus *Coffea*. L. Theor Appl Genet 65:77–84.

Bradeen JM and Simon PW, 1998. Conversion of an AFLP fragment linked to the carrot Y2 locus to a simple, codominant, PCR-based marker form. Theor Appl Genet 97:960–967.

Carels N and Bernardi G, 2000. Two classes of genes in plants. Genetics 154:1819–1825.

Coulibaly I, Revol B, Noirot M, Poncet V, Lorieux M, Carasco-Lacombe C, Minier J, Dufour M, and Hamon P, 2003. AFLP and SSR polymorphism in a *Coffea* interspecific backcross progeny, (*C. heterocalyx* × *C. canephora*) × *C. canephora*. Theor Appl Genet 107:1148–1155.

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

Fulton TM, van der Hoeven R, Eannetta NT, and Tanksley SD, 2002. Identification, analysis, and utilization of conserved ortholog set markers for comparative genomics in higher plants. Plant Cell 14:1457–1467.

Gunter LE, Roberts GT, Lee K, Larimer FW, and Tuskan GA, 2003. The development of two flanking SCAR markers linked to a sex determination locus in *Salix viminalis*. L. J Hered 94:185–189.

Haanstra JPW, Wye C, Verbakel H, Meijer-Dekens F, van den Berg P, Odinot P, van Heusden AW, Tanksley S, Lindhout P, and Peleman J, 1999. An integrated high-density RFLP-AFLP map of tomato based on two *Lycopersicon esculentum* × *L. pennellii* F2 populations. Theor Appl Genet 99:254–271.

Hayashi M, Nishioka M, Kitamura K, and Harada K, 2000. Identification of AFLP markers tightly linked to the gene for deficiency of the 7S globulin in soybean seed and characterization of abnormal phenotypes involved in the mutation. Breed Sci 50:123–129.

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

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

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

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

Lashermes P, Combes MC, Trouslot P, and 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.

Leroy JF, 1980. Evolution et taxogenèse chez les caféiers: hypothèse sur l'origine. Comptes Rendus Acad Sci 291:593–596.

Menz MA, Klein RR, Mullet JE, Obert JA, Unruh NC, and Klein PE, 2002. A high-density genetic map of *Sorghum bicolor* (L.) Moench based on 2926 AFLP, RFLP and SSR markers. Plant Mol Biol 48:483–499.

Mokrani L, Gentzbittel L, Azanza F, Fitamant L, Al-Chaarani G, and Sarrafi A, 2002. Mapping and analysis of quantitative trait loci for grain oil content and agronomic traits using AFLP and SSR in sunflower (*Helianthus annuus* L.). Theor Appl Genet 106:149–156.

Negi MS, Devic M, Delseny M, and Lakshmikumaran M, 2000. Identification of AFLP fragments linked to seed coat colour in *Brassica juncea* and conversion to a SCAR marker for rapid selection. Theor Appl Genet 101:146–152.

Oetting WS, Lee HK, Flanders DJ, Wiesner TA, Sellers TA, and King RA, 1995. Linkage analysis with multiplexed short tandem repeat polymorphisms using infrared fluorescence and M13 tailed primers. Genomics 30:450–458.

Ouedraogo JT, Gowda BS, Jean M, Close TJ, Ehlers JD, Hall AE, Gillaspie AG, Roberts PA, Ismail AM, Bruening G, Gepts P, Timko MP, and Belzile FJ, 2002. An improved genetic linkage map for cowpea (*Vigna unguiculata* L.) combining AFLP, RFLP, RAPD, biochemical markers, and biological resistance traits. Genome 45:175–188.

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

Peters JL, Constandt H, Neyt P, Cnops G, Zethof J, Zabeau M, and Gerats T, 2001. A physical amplified fragment-length polymorphism map of *Arabidopsis*. Plant Physiol 127:1579–1589.

Poncet V, Hamon P, Minier J, Carasco-Lacombe C, Hamon S, and Noirot M, in press. SSR cross-amplification and variation within coffee trees (*Coffea* spp.). Genome.

Powell W, Thomas WTB, Baird E, Lawrence P, Booth A, Harrower B, McNicol JW, and Waugh R, 1997. Analysis of quantitative traits in barley by the use of amplified fragment length polymorphisms. Heredity 79:48–59.

Qi X and Lindhout P, 1997. Development of AFLP markers in barley. Mol Gen Genet 254:330–336.

Qi X, Stam P, and Lindhout P, 1998. Use of locus-specific AFLP markers to construct a high-density molecular map in barley. Theor Appl Genet 96:376–384.

Rozen S and Skaletsky HJ, 2000. Primer3 on the WWW for general users and for biologist programmers. In: Bioinformatics methods and protocols: methods in molecular biology (Krawetz S and Misener S, eds). Totowa, NJ: Humana Press; 365–386.

Shan X, Blake TK, and Talbert LE, 1999. Conversion of AFLP markers to sequence-specific PCR markers in barley and wheat. Theor Appl Genet 98:1072–1078.

Thoquet P, Gherardi M, Journet EP, Kereszt A, Ane JM, Prosperi JM, and Huguet T, 2002. The molecular genetic linkage map of the model legume *Medicago truncatula*: an essential tool for comparative legume genomics and the isolation of agronomically important genes. BMC Plant Biol 2:1.

van Eck HJ, van der Voort JR, Draaistra J, van Zandvoort P, van Enckevort E, Segers B, Peleman J, Jacobsen E, Helder J, and Bakker J, 1995. The inheritance and chromosomal localization of AFLP markers in a non-inbred potato offspring. Mol Breed 1:397–410.

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

Vuylsteke M, Mank R, Antonise R, Bastiaans E, Senior ML, Stuber CW, Melchinger AE, Lubberstedt T, Xia XC, Stam P, Zabeau M, and Kuiper M, 1999. Two high-density AFLP R linkage maps of *Zea mays* L.: analysis of distribution of AFLP markers. Theor Appl Genet 99:921–935.

Xu M, Huaracha E, and Korban SS, 2001. Development of sequence-characterized amplified regions (SCARs) from amplified fragment length polymorphism (AFLP) markers tightly linked to the Vf gene in apple. Genome 44:63–70.

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

Received May 14, 2004
Accepted July 2, 2004

Corresponding Editor: J. Perry Gustafson