Theor Appl Genet (2000) 101:948–955

A.M. Risterucci · L. Grivet · J.A.K. N'Goran
I. Pieretti · M.H. Flament · C. Lanaud

A high-density linkage map of *Theobroma cacao* L.

Received: 2 January 2000 / Accepted: 12 February 2000

Abstract The first linkage map established by Lanaud et al. (1995) was used as a starting point to produce a high-density molecular linkage map. A mapping population of 181 progenies resulting from a cross between two heterozygous genotypes, a Forastero and a Trinitario (hybrid between Forastero and Criollo), was used for the linkage analysis. A new DNA isolation protocol was established, which allows enough good quality DNA to construct a genetic map with PCR-based markers. The map comprises 424 markers with an average spacing between markers of 2.1 cM. The marker types used were five isozymes, six loci from known function genes, 65 genomic RFLPs, 104 cDNA RFLPs, three telomeric probes, 30 RAPDs, 191 AFLPs and 20 microsatellites. The use of new marker types, AFLP and microsatellites, did not disturb the original order of the RFLP loci used on the previous map. The genetic markers were distributed over ten linkage groups and cover 885.4 cM. The maximum distance observed between adjacent markers was 16.2 cM, and 9.4% of all loci showed skewed segregation.

Key words *Theobroma cacao* L. · AFLP · Microsatellites · RFLP · High-density genetic map

Introduction

Theobroma cacao L. (2n=20) is native to Central and South America and is one of the most economically important perennial crops in the tropics. Its genome is small (0.4 pg/1 C) (Lanaud et al. 1992; Figueira et al. 1992), and its germplasm is usually classified into three distinct groups based on geographical and morphological characters (Cheesman 1944): Forastero, Criollo and Trinitario. The last group corresponds to hybrid forms between the first two groups. Populations of the Forastero group are highly diversified and are often used in breeding programs, due to their vigor, precocity and resistance to disease. Criollo types are generally less vigorous but yield a fine chocolate. Until now, the main breeding strategies have involved crosses between individuals of distinct genetic groups to obtain productive trees combining resistance and quality traits.

The genetic determinism of most of the traits of interest is generally not known, and mapping strategies are crucial to increasing our knowledge and mastering of favorable traits during the different steps of selection. Molecular mapping of the cocoa genome was recently initiated (Lanaud et al. 1995; Crouzillat et al. 1996; Risterucci et al.1996). The first cocoa genetic map comprised 195 markers, mostly restriction fragment length polymorphisms (RFLPs) (Lanaud et al. 1995). RFLP markers are laborious to use, especially in cocoa, for quantitative trait loci (QTL) analyses on large numbers of progeny because of the difficulty in purifying of large amounts of good quality DNA, which requires ultracentrifugation. Other highly reproducible markers, such as those based on polymerase chain reaction (PCR) for example, AFLPs and microsatellites, could advantageously replace RFLP in cocoa mapping.

Microsatellites, also called simple-sequence repeats (SSRs) (Tautz and Rentz 1984), are small repeats of two tandemly arranged nucleotides. They are highly polymorphic and can be amplified by PCR (Mullis et al. 1986) with unique flanking primers (Beckman and Soller 1990). They have been used in mapping several mammalian (Love et al. 1990; Weissenbach et al. 1992; Serikawa et al. 1992) and plant genomes such as soybean (Akkaya et al. 1992), rice (Zhao et al. 1993) and maize (Senior and Heun 1993). Amplified fragment

Communicated by P.M.A. Tigerstedt

A.M. Risterucci (✉) · L. Grivet · I. Pieretti · M.H. Flament
C. Lanaud
CIRAD, Centre de Cooperation Internationale en Recherche Agronomique pour le Développement, avenue Agropolis, BP 5035, 34032 Montpellier cedex, France
e-mail: risterucci@cirad.fr
Fax: +33467615605

J.A.K. N'Goran
IDEFOR-DCC, BP 1827, Abidjan 01, Côte d'Ivoire

length polymorphism (AFLP) markers (Vos et al. 1995) are based on the detection of genomic selective restriction fragments by PCR amplification (Zabeau and Vos 1992) and have also been used more recently for mapping annual plant genomes (Meksem et al. 1995; Thomas et al. 1995; Cho et al. 1996; Keim et al. 1997) or tree genomes (Marques et al. 1998; Remington et al. 1999). This method generates a large number of restriction fragment bands, thereby facilitating the detection of polymorphisms and improving the prospects for a high-density cocoa genetic map. However, some authors have noted an irregular distribution of these markers along chromosomes (Maheswaran et al. 1997; Vuylsteke et al. 1999).

In the investigation described in this paper we used the first linkage map established by Lanaud et al. (1995) as a starting point to produce a more saturated linkage map that includes isozyme, RFLP, AFLP, microsatellite and RAPD markers and to evaluate the potential of AFLP and microsatellites markers in cocoa mapping.

Materials and methods

Mapping population

The progeny analyzed resulted from a cross between a UPA402 female parent and a UF676 male parent. UPA 402 is an upper amazon Forastero clone obtained from a sib mating involving IMC 60 and Na 34, two Forastero genotypes collected in Peru. UF 676 is a Trinitario selection derived from the United Company's breeding program in Costa Rica. UPA 402 has a higher level of homozygosity than UF 676, based on pedigree and previous molecular analyses (Lanaud et al. 1995). A mapping population of 181 progenies was used for the linkage analysis, planted in the Ivory Coast (IDEFOR/DCC).

Isozyme analysis

Isozyme analysis was performed using five enzyme systems according to protocols described by Lanaud (1986, 1987). These included acid phosphatase (ACP), isocitrate dehydrogenase (IDH), malate dehydrogenase (MDH), phosphoglucose isomerase (PGI), phosphoglucomutase (PGM).

DNA isolation

A new DNA isolation protocol was used to obtain the small quantities of DNA needed for microsatellite and AFLP analysis. DNA was isolated from fresh adult leaves. One gram of tissue frozen in liquid nitrogen and powdered with a mortar was mixed with 5 ml of extraction buffer (1.4 *M* NaCl, 100 m*M* Tris HCl pH 8.0, 20 m*M* EDTA, 10 m*M* Na_2SO_3, 1% PEG 6000, 2% MATAB) preheated to 75°C. The extract was then homogenized for 10 s with a vortex and incubated 30 min at 75°C; after being cooled to 20°C, an equal volume of chloroform-isoamyl alcohol (24:1 v/v) was added, followed by emulsification. The tube was then centrifuged at 7000 *g* for 30 min and the supernatant was precipitated at –20°C overnight after the addition of an equal volume of isopropanol. The DNA was removed with a glass hook and re-suspended in 1 ml of 0.7 *M* NaCl, 50 m*M* TRIS-HCl, 10 m*M* EDTA, pH 7.0 buffer. After total re-suspension, a QIAGEN genomic-tip was performed as recommended by the supplier.

RFLP analysis

cDNA clones and genomic *Pst*I clones from libraries constructed at CIRAD (Lanaud et al. 1995) were screened, along with ten genomic *Pst*I clones provided by Nestlé (Crouzillat et al. 1996). A cloned telomeric probe library, obtained according to Kilian and Kleinhofs (1992), was screened on Southern blots of parental DNA restricted with *Eco*RI, *Eco*RV, *Bgl*II, *Hind*III and *Xba*I. In addition three cloned genes were used: histone (H4Cl4), isolated from maize (Philipps et al. 1986), alcohol dehydrogenase (Adh1), isolated from maize (Gerlach et al. 1982), and the rRNA gene (pTA71), isolated from wheat (Gerlach and Bedbrook 1979). Three bands obtained from RAPD patterns were also used as labeled RFLP probes. DNA fragments were extracted from low-melting agarose gel and hybridized, without cloning, onto the restricted DNA. RFLP procedures were performed according to the protocol described by Lanaud et al. (1995).

AFLP analysis

AFLP were performed using the Life Technologies AFLP analysis system I (Gibco BRL, Gathersburg, Md., USA) as recommended by the supplier. Genomic DNA (250 ng) was restricted with 2.5 U *Eco*RI and 2.5 U *Mse*I in a 25-µl reaction mixture containing 5 m*M* TRIS-HCl (pH 7.5), 5 m*M* magnesium acetate, 25 m*M* potassium acetate for 2 h at 37°C; the enzymes were then inactived at 70°C for 15 min. Twenty-five microliters of a mixture containing 5 pmol *Eco*RI adapter, 50 pmol *Mse*I adapter, 1 U T4 DNA ligase, 0.4 m*M* ATP in 10 m*M* TRIS-HCl (pH 7.5), 10 m*M* magnesium acetate, and 50 m*M* potassium acetate was added. The ligation mixture was incubated at 20°C for 2 h. As 5-µl aliquot of a tenfold-diluted ligation was amplified for a preselective amplification using a pair of primers based on the sequences of the *Eco*RI and the *Mse*I adapters, including one additional selective nucleotide at the 3´ end (*Eco*RI+1 and *Mse*I+1). The selective amplification reaction was performed with *Eco*RI+3 and *Mse*I+3 primers. *Eco*RI+3 primers were labeled by phosphorylating the 5´ end with γ-[^{33}P]-ATP using a T4 polynucleotide kinase (Gibco BRL). This primer was mixed with 1 or 2 *Mse*I+3 primers, dNTPs, 10×PCR buffer (100 m*M* TRIS-HCl pH 8.3, 15 m*M* $MgCl_2$, 500 m*M* KCl) and 0.5 U *Taq* polymerase (Eurobio, Les Ulis, France), then 15 µl of the mix was added to 5 µl of 50-fold-diluted preamplified DNA. Amplifications were performed using an MJ Research (Water Town, Mass., USA) PTC 100 thermal cycler. One cycle was performed at 94°C for 30 s, 65°C for 30 s and 72°C for 60 s, followed by 12 cycles with a 0.7°C lower annealing temperature for each cycle and 23 cycles at 94°C for 30 s, 56°C for 30 s and 72°c for 60 s. After the addition of 20 µl of loading buffer (98% formamide, 10 m*M* EDTA, bromophenol blue, xylene cyanol), the mixes were denatured at 92°C for 3 min, and 3 µl of each sample was loaded onto a 5% polyacrylamide gel with 7.5 *M* urea and electrophoresed in 0.5% TBE buffer at 55 W for 1 h 40 min. The gel was dried for 30 min at 80°C and exposed overnight to X-ray film (Fuji RX).

Microsatellite analysis

A genomic library enriched in SSRs was constructed for microsatellites analysis using a modified version of the protocol of Karagyozov et al. 1993 (Lanaud et al. 1999). The primers were end-labeled with γ-[^{33}P]-ATP, and amplification was performed in a MJ Research PTC 100 Thermal cycler on 20-µl reaction mixtures containing 10 ng of cocoa DNA, 0.2 m*M* dNTP mix, 2 m*M* MgCl2, 50 m*M* KCl, 10 m*M* TRIS-HCl (pH 8.3), 0.2 µ*M* primer (5´ end labeled with γ-[^{33}P]-ATP) and 1 U *Taq* polymerase (Eurobio). The samples were denatured at 94°C for 4 min and subjected to 32 repeats of the following cycle: 94°C for 30 s, 46°C or 51°C for 1 min and 72°C for 1 min. The samples were analyzed under the same electrophoresis and developing conditions as those used for AFLP. After screening on the parents, 20 microsatellites were

analyzed on the mapping population. EMBL accession numbers of the microsatellites are listed in Table 1.

5 S ribosomal DNA analysis

The spacer of the 5S ribosomal DNA was mapped by amplification of cocoa DNA with consensus primers (D'Hont et al. 1995) followed by an *Hpa*II restriction, and analyzed in 2% TBE agarose gel.

Random amplified polymorphic DNA (RAPD) analysis

A modified version of the protocol of Williams et al. (1990) was applied using primers provided by Operon Technologies. Amplification products were analyzed in 2% TBE agarose gel (Lanaud et al. 1995).

Table 1 EMBL accession number and linkage group localization of microsatellite markers

Marker name	Linkage group	EMBL accession no.
mTcCIR1	8	Y16883
mTcCIR2	5	Y16978
mTcCIR3	2	Y16977
mTcCIR6	6	Y16980
mTcCIR7	7	Y16981
mTcCIR8	9	Y16982
mTcCIR9	6	Y16983
mTcCIR10	5	Y16984
mTcCIR11	2	Y16985
mTcCIR12	4	Y16986
mTcCIR15	1	Y16988
mTcCIR16	6	Y16989
mTcCIR17	4	Y16990
mTcCIR18	4	Y16991
mTcCIR19	2	Y16992
mTcCIR21	3	Y16994
mTcCIR22	1	Y16995
mTcCIR24	9	Y16996
mTcCIR25	6	Y16997
mTcCIR26	8	Y16998

Linkage analysis

Segregation of markers was revealed on 100 individuals for RFLPs, RAPDs and isozymes, and on 181 individuals for AFLPs and microsatellites. Linkage analyses were performed using the program JOINMAP version 1.4 (Stam 1993). As both parents of the progeny were heterozygous, the markers segregated according to the two possible Mendelian arrangements, heterozygous for one parent and homozygous for the other parent, and heterozygous in both parents. The markers heterozygous for both parents enabled an integrated map to be established between the two individual parental maps. The segregation of 424 markers was studied using LOD scores of 5.0 and 4.0 to identify the linkage groups. The Kosambi mapping function was used to convert recombination frequencies into map distances (Kosambi 1944).

Marker nomenclature

RFLP probes were named cTcCIR, gTcCIR, and rTcCIR, with the first letter, c, g or r, corresponding to cDNA, genomic and isolated RAPD genomic fragments, respectively. Tc corresponds to *Theobroma cacao* and CIR to CIRAD. Microsatellite markers were named mTcCIR. TEL corresponds to telomeric markers; N corresponds to Nestlé genomic probes. RAPD loci were named rOPX#, with OP corresponding to OPERON Technologies, X the primer kit letter and # the approximate molecular weight of the band. AFLP loci were named AFLP X/Y; X is the number of the primers' combination and Y the number of the polymorphic band (Table 2).

Results and discussion

Screening of polymorphic markers

The screening of polymorphic markers was performed by testing for the marker on the two parents and on 6 individuals of the progeny. The results are given in Table 3. A total of 614 RFLP probes was screened using five restriction enzymes. Segregation in the progeny was observed for 28% of the cDNA probes and 30% of the genomic probes. Twenty telomeric probes were screened,

Table 2. AFLP primer combinations used for cocoa genome mapping based on the cross UPA402×UF676.The number of polymorphic markers used for mapping and the ratio of polymorphism are given. In 12 cases, two *Mse*I primers were used simultaneously

AFLP number	Primer combination	Detectable fragments	Polymorphic markers	Polymorphic markers (%)
1	EAAG/MCAA	66	17	25.7
2	EAAG/MCTT	71	8	11.2
3	EAAC/MCAA	77	7	9.1
4	EACG/MCTT	53	6	11.3
5	EAAG/MCTC-MCTA	85	21	24.7
6	EAAC/MCAC-MCAT	72	16	22.2
7	EAAC/MCAG	40	8	20.0
8	EAAC/MCTA-MCTG	72	14	19.4
9	EAGC/MCAT-MCTA	58	9	15.5
10	EAGC/MCAA	56	15	26.8
11	EAGG/MCTC-MCTT	58	10	17.2
12	EACG/MCAA-MCTG	48	7	14.6
13	EAGG/MCAA-MCAC	47	2	4.3
14	EACT/MCTC-MCTG	56	11	19.6
15	EAGG/MCAT-MCTA	46	6	13.0
16	EACT/MCTT	61	11	18.0
17	EACA/MCTA-MCTT	77	9	11.7
18	EACC/MCAT-MCTG	43	8	18.6
19	EACC/MCAC-MCAG	50	4	8.0
20	EAAC/MCTC	49	2	4.1

Table 3 Level of marker polymorphism on both parents studied with respect to the various molecular techniques used. The number of heterozygous markers in UF676 only, in UPA402 only and in both parents is indicated

	Markers screened	UF676	UPA402	UPA402 and UF676	Total loci mapped	Mapped loci ratio
Isozymes	5	5	0	0	5	100.0
Genomic RFLPs	215	53	7	5	65	30.2
cDNA RFLPs	379	81	12	11	104	27.4
Telomeric RFLPs	20	3	0	0	3	15.0
Known genes	6	5	1	0	6	100.0
RAPDs	394	24	6	0	30	6.1
AFLPs	1185	118	50	23	191	16.1
Microsatellites	26	6	2	12	20	76.9
Total	2230	295	78	51	424	

but only 3 gave a non-repetitive, clear-cut signal scorable in the progeny for mapping. One hundred and twenty random primers, revealing 394 fragments, were screened for RAPD polymorphism. Among them, 30 markers provided by 23 primers segregated in the progeny. Twenty-six microsatellites were screened and 20 of them were segregated in the progeny (76.9%). Sixty-four AFLP primer combinations resulting from 8 *Eco*RI+3 primers and 8 *Mse*I+3 primers were screened. Thirty-two of them, which revealed the greatest number of polymorphic fragments, were selected, 8 pairs of primers corresponding to a single *Eco*RI+3 primer associated with a single MseI+3 primer and 12 combinations of primers corresponding to a single *Eco*RI+3 primer associated with 2 *Mse*I+3 primers. The chosen AFLP combinations revealed a total of 1185 fragments (Table 2). The number of detectable fragments obtained for each combination ranged from 43 to 85, with band sizes ranging from 50 to 500 bp. The number of polymorphic bands ranged from 2 to 21 with a mean of 9.5. Only clearly scorable fragments were used. In total, 191 clearly scorable bands (16.1%) segregated in the progeny and were mapped. This level of polymorphism was in the range of those reported in other species. For example Spada et al. (1998) reported 9.1% in asparagus, Becker et al. (1995) 11.3% in barley, and Mashewaran et al. (1997) 22.0% in rice. The numbers of markers segregating in the gametes of UPA402 only, in the gametes of UF676 only and in the gametes of the both parents are reported in Table 3 for each type of marker. This clearly shows the higher level of polymorphism revealed by microsatellites compared to the other types of markers. Among 26 microsatellites tested, 20 segregated in the progeny, 12 of them being heterozygous in both parents and revealing generally more than two alleles. This is of particular interest to constructing integrated maps and detecting QTL in both parents of such test-cross progenies. A majority of AFLP polymorphic fragments was provided with only 32 primer pairs. The percentage of AFLP markers, heterozygous in both parents and segregating with a theoretical ratio of 3:1, was 12%. This result is similar to what was obtained with RFLPs and lower than that obtained with microsatellites.

Segregation distortion

Each segregating marker was first tested with an χ^2 test for goodness-of-fit to the expected Mendelian segregating ratio. This was 1:1 for markers heterozygous in only one of the two parents; 3:1 for dominant markers like AFLPs and RAPDs, heterozygous in both parents; and 1:2:1 and 1:1:1:1 for codominant multiallelic markers (isozymes, RFLPs, microsatellites) when two, three or four alleles, respectively, were segregating among the two parents. Skewed segregation was detected for 40 loci (9.4% of the total), with 23 being significant at P=0.05, and 17 at P=0.01. Loci showing a skewed segregation were mostly located on four regions of the linkage map and involved several types of markers (AFLPs, RFLPs etc.); 8 loci were linked on group 6, 6 were on group 5, 4 were on group 2 and 8 were on group 9. The 8 loci of group 6 and the 4 loci of group 2 segregated in the gametes of UF676 only. The 8 loci of group 9 segregated either in the gametes of UPA402 only or in the gametes of both parents. Some imprecision in the order and the gap between markers is induced by these distorted segregations, especially in group 6 where 3 of the 6 markers heterozygous in both parents were affected, which is highly uninformative.

Note that by using the Bonferoni procedure, the threshold for a global type-I error of P=0.05 would necessitate an individual threshold of P=0.0001 if all markers were considered independent. At this threshold no skewed segregation would appear as significant.

This skewed segregation ratio was relatively low compared to those found for other species such as potato (25.5%) (Gebhart et al. 1989) or *Brassica oleracea* (12–59%) (Kianian and Quiros 1992). Chromosomal rearrangements (Tanksley et al. 1984, 1987; Faure et al. 1992) or gametic selection (Nakagahra 1986) have been suggested to explain distorted segregation. In *Theobroma cacao*, structural rearrangements have never been reported, and the genes responsible for gametic selection are unknown.

Construction of the genetic linkage map

The 424 polymorphic loci were assigned to ten linkage groups (Fig. 1) that putatively correspond to the ten gametic chromosomes of cocoa (Munoz 1948). These loci

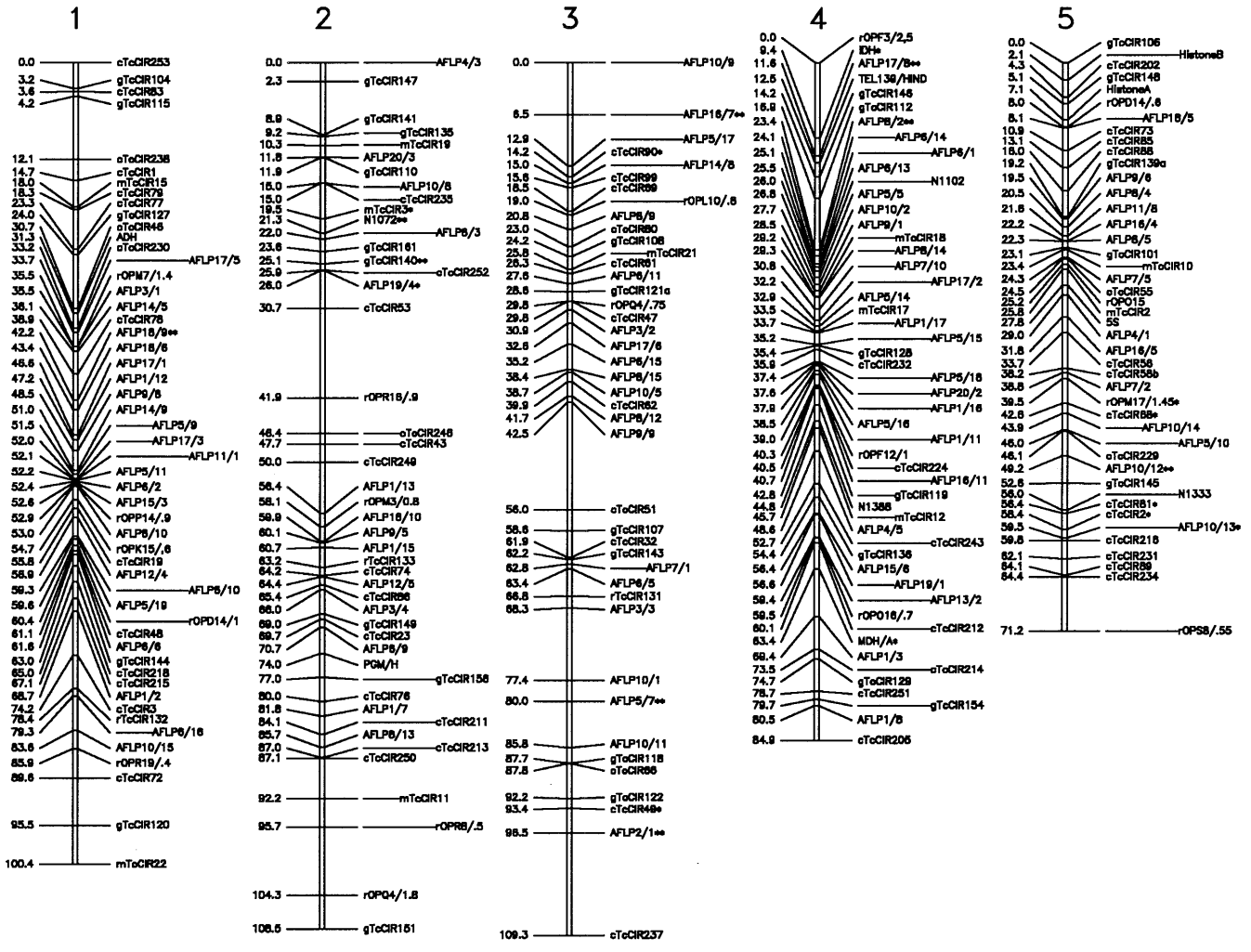

Fig. 1 Linkage map of cocoa, based on the cross UPA402× UF676. Markers polymorphic in the gametes of UPA402, in the gametes of UF676, and in the gametes of the both parents are positioned on the *right*, *left* and *middle*, respectively

correspond to five isozymes, 65 genomic RFLPs, 104 cDNA RFLPs, three telomeric RFLP, six genes of known function, 30 RAPDs, 191 AFLPs and 20 microsatellites. The total length of the map is 885.4 cM. The lengths of individual linkage groups vary between 68.4 cM and 111.3 cM, with a mean of 88.5 cM. The average distance between 2 markers is 2.09 cM.

All markers heterozygous in UF676 only and all markers heterozygous in both parents were integrated into a linkage groups with a two-point LOD score ≥5. In spite of the more limited number of markers heterozygous in UPA 402 and homozygous in UF 676, all of these 76 markers could also be assigned to a linkage group with a LOD score ≥5. The LOD score threshold was, however, lowered to 4.0 for 5 markers, AFLP5/17, AFLP10/9, AFLP14/8, AFLP16/7 and rOPL10/.8. which were assigned to group 3. The presence in group 10 of only 1 marker segregating in the gametes of both parents gives an ambiguous position for the map of UPA402 relative to the map of UF676 for this group. This study established the usefulness of microsatellite markers in cocoa mapping and genetic analysis due to their high level of polymorphism. The addition of new AFLP, RFLP and microsatellite markers did not disturb the original order of the RFLP loci obtained in the first map (Lanaud et al. 1995). The only exception was chromosome 4, where the order of the RFLP markers ctcCIR224 and gtcCIR119 was inverted. These 2 markers, heterozygous in both parents, were very closely linked on the first map. They are also responsible for the bad orientation of a cluster of 4 RFLP markers segregating in the gametes of UPA402 only, gtcCIR154, ctcCIR214, ctcCIR212 and ctcCIR243 on the first map. The addition of 8 new markers heterozygous in both parents in this group enabled us to recover the correct relative positions of the individual maps of both parents.

AFLPs filled important gaps in the RFLP map, and the length of the map was slightly increased. A stretching of 126 cM was observed (plus 16.6%) corresponding essentially to an increase in five linkage groups (LG), LG 9 plus 44.6 cM, LG 8 plus 31.5 cM, LG 7 plus 21.0 cM, LG 3 plus 10.5 cM and LG 6 plus 9.0 cM. This stretching was observed at one end of linkage groups 3,

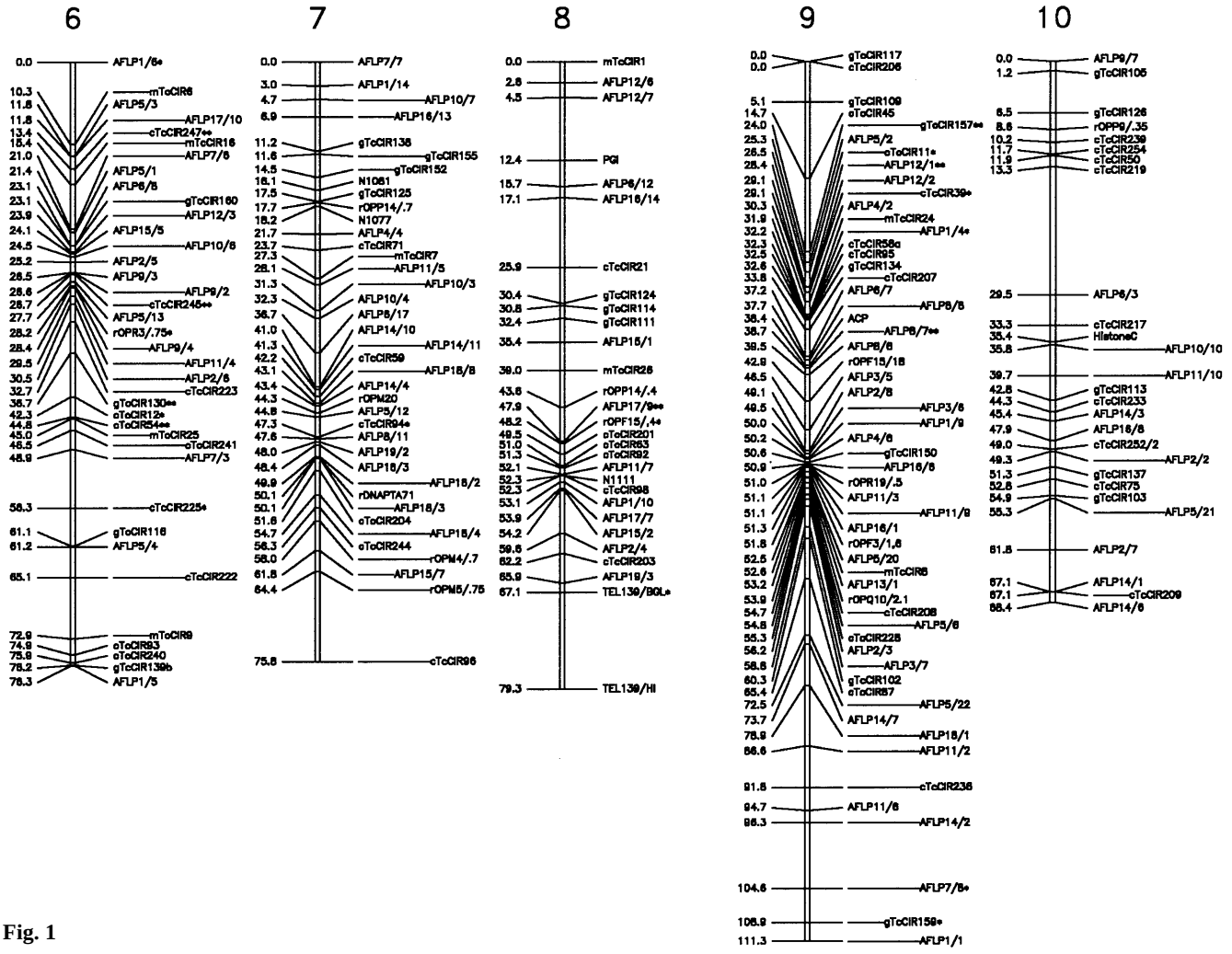

Fig. 1

6, 7 and 9 and at both ends of linkage group 8. For this group the stretching of 17.1 cM was obtained on one arm by incorporating 1 AFLP marker and 2 telomeric RFLPs. It is noteworthy that these telomeric probes map to the linkage group ends and are not interstitial. It seems that the increase in map size was generally lower than that observed in other studies where RFLP and AFLP mixed maps have been integrated. Becker et al. (1995) reported an increase of 70.9% on the barley map, Maheswaran et al. (1997) an increase of 68.8% on the rice map and Boivin et al. (1999) observed an increase of 40.5% on the sorghum map. This may be related to a good saturation of the original RFLP (Lanaud et al. 1995).

A higher density of markers was observed in some regions of the map. This may correspond to centromeric regions where an inhibition of crossing-over have been reported (Tanksley et al. 1992). A localized high density of AFLP markers has been reported in other species like sugar beet (Schondelmeier et al. 1996) and soybean (Keim et al. 1997). However, the increase and filling of important gaps by the inclusion of AFLP markers was observed in specific regions of the genome which were poorly covered by RFLP markers. One possible explanation is the use of *Eco*RI-based AFLPs in contrast to the use of RFLPs provided by cDNA or *Pst*I (sensitive to GC methylation) genomic probes (Boivin et al. 1999). If we consider the map covered only with AFLP markers, its size is 720.9 cM, which corresponds to 81.4% of the mixed map size. The least reliable coverage (50%) was obtained on linkage group 1, where one important AFLP cluster of 13 markers was detected between 38.9 cM and 53.0 cM. Heterogeneity in marker density along genetic maps is commonly found in plants (Heslop-Harrison et al. 1993). The present map offers a relative homogeneity of marker distribution. The average distance between 2 markers varied from 1.6 cM for group 5 to 2.7 cM for group 8. Only six gaps were larger than 10 cM, with the largest gap being 16.2 cM. Such a distribution and density will be useful for supporting marker-assisted selection in plant breeding. This map is the first report of a high-density linkage map constructed for a tree crop. A few high-density linkage maps have been published previously for annual plants such as tomato, with a mean marker spacing of 1.2 cM

(Tanksley et al. 1992), maize, with a mean marker spacing of 2.5 cM (Coe and Gardiner 1994), sugar beet, with a mean marker spacing of 1.5 cM (Hallden et al. 1996), and rice, with a mean marker spacing of 0.67 cM (Harushima et al. 1998). The combined use of AFLPs and microsatellites seems to be the most convenient tool for future genetic studies in cocoa. Due to their high locus specificity and level of polymorphism, microsatellites allow early identification of the linkage group and integration or comparison of maps between several progenies. A special effort is presently being made to produce more microsatellite markers in cocoa, thereby allowing a better coverage of the whole genome. These two marker technologies do not require a great amount of DNA and are convenient for genetic analysis on plants at an early stage of growth. The high multiplex ratio achievable with AFLP offers a great advantage for the analysis of a large number of loci with a good distribution of markers along the genome and enables dense linkage maps to be developed more rapidly. The ease of use of such tools will subsequently enable more progenies to be mapped in order to study traits of interest. This will also help identify the most important regions involved in the variation of the characters of interest to breeding programs. Moreover, microsatellite markers will be easily transferable to laboratories in tropical regions and applied to marker-assisted selection.

Declaration The experiments described herein comply with the current laws of France in which the experiments were performed.

References

Akkaya MS, Bhagwat AA, Cregan PB (1992) Length polymorphisms of simple sequence repeat DNA in soybean. Genetics 132:1131–1139

Becker J, Vos P, Kuiper M, Salamini F, Heun M (1995) Combined mapping of AFLP and RFLP markers in barley. Mol Gen Genet 349:65–73

Beckman JS, Soller M (1990) Toward a unified approach to genetic mapping of eukaryotes based on sequence tagged microsatellite sites. Bio/Technology 8:930–932

Boivin K, Deu M, Rami JF, Trouche G, Hamon P (1998) Towards a saturated sorghum map using RFLP and AFLP markers. Theor Appl Genet 98:320–328

Cheesman EE (1944) Notes on the nomenclature, classification and possible relationships of cacao populations. Trop Agric 21:144–159

Cho YG, Blair MW, Panaud O, McCouch S (1996) Cloning and mapping of variety-specifi**c** rice genomic DNA sequences: Amplified fragment length polymorphisms (AFLP) from silver-stained polyacrylamide gels. Genome 39:373–378

Coe EH, Gardiner JM (1994) RFLP maps of maize. In: Philips RL, Vasil IK (eds) DNA-based markers in plants. Kluwer Academic Publ, Dordrecht, The Netherlands, pp 240–245

Crouzillat D, Lerceteau E, Petiard V, Morera J, Rodriguez H, Walker D, Phillips W, Ronning C, Schnell R, Osei J, Fritz P (1996) *Theobroma cacao* L.: a genetic linkage map and quantitative trait loci analysis. Theor Appl Genet 93:205–214

D'Hont A, Rao PS, Feldmann P, Grivet L, Islam-Faridi N, Taylor P, Glaszmann JC (1995) Identification and characterisation of sugarcane intergeneric hybrids, *Saccharum officinarum*×*Erianthus arundinaceus*, with molecular markers and DNA in situ hybridisation. Theor Appl Genet 91:320–326

Faure S, Noyer JL, Horry JP, Bakry F, Lanaud C, Gonzalez de Leon D (1992) A molecular marker based linkage map of diploid bananas (*Musa acuminata*). Theor Appl Genet 87:517–526

Figueira A, Janick J, Goldsbrough P (1992) Genome size and DNA polymorphism in *Theobroma cacao*. J Am Soc Hortic Sci 17:673–677

Gebhardt C, Ritter E, Debener T, Schachtschabel U, Walkemeier B, Uhrig H, Salamini F (1989) RFLP analysis and linkage mapping in *Solanum tuberosum*. Theor Appl Genet 78:65–75

Gerlach WL, Bedbrook JR (1979) Cloning and characterization of ribosomal RNA genes from wheat and barley. Nucleic Acids Res 7:1869–1885

Gerlach WL, Pryor AJ, Dennis ES, Ferl RJ, Sachs MM, Peacock WJ (1982) cDNA cloning and induction of the alcohol dehydrogenase gene (*Adh1*) of maize. Proc Natl Acad Sci USA 79:2981–2985

Hallden C, Hjerdin A, Rading IM, Sall T, Fridlundh B, Johannisdottir G, Tuvesson S, Akesson C, Nilsson NO (1996) A high density RFLP linkage map of sugar beet. Genome 39:634–645

Harushima Y, Yano M, Shomura A, Sato M, Shimano T, Kuboki Y, Yamamoto T, Yang-Lin S, Antonio BA, Parco A, Kajiya H, Huang N, Yamamoto K, Nagamura Y, Kurata N, Khush GS, Sasaki T (1998) A high density rice genetic linkage map with 2275 markers using a single F_2 population. Genetics 148:479–494

Heslop-Harrison JS, Leitch JS, Schwarsaker T (1993) The physical organization of the interphase nuclei. In: BIOS Scientific Publishers Ltd. The chromosome. John Innes review, John Innes Institutes, Norwich, UK, pp 221–232

Karagyozov L, Kalcheva ID, Chapman VM (1993) Construction of random small-insert genomic libraries highly enriched for simple sequence repeats. Nucleic Acids Res 21:3911–3912

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

Kianian SF, Quiros CF (1992) Generation of a *Brassica oleracea* composite RFLP map : linkage arrangements among various populations and evolutionary implications. Theor Appl Genet 87:544–554

Kilian A, Kleinhofs A (1992) Cloning and mapping of telomere-associated sequences from *Hordeum vulgare* L. Mol Gen Genet 235:153–156

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

Lanaud C (1986) Utilisation des marqueurs enzymatiques pour l'étude génétique du cacaoyer *Theobroma cacao* L. II. Etude du polymorphisme de six systèmes enzymatiques. Cafe, Cacao The 30:271–280

Lanaud C (1987) Nouvelles données sur la biologie du cacaoyer (*Theobroma cacao* L.) : diversité des populations, système d'incompatibilité, haplo spontanés. Leurs conséquences pour l'amélioration génétique de cette espèces. PhD thesis, Université Paris XI, Paris, France

Lanaud C, Hamon P, Duperray C (1992) Estimation of the nuclear DNA content of *Theobroma cacao* L. by flow cytometry. Cafe, Cacao, The 36:3–8

Lanaud C, Risterucci AM, N'Goran JAK, Clement D, Flament MH, Laurent V, Falque M (1995) A genetic linkage map of *Theobroma cacao* L. Theor Appl Genet 91:987–993

Lanaud C, Risterucci AM, Pieretti I, Falque M, Bouet A, Lagoda PJL (1999) Isolation and characterization of microsatellites in *Theobroma cacao* L. Mol Ecol 8:2141–2152

Love JM, Knight A, Mc Aleer M, Todd JA (1990) Towards construction of a high resolution map of the mouse genome using PCR-analysed microsatellites. Nucleic Acids Res 18:4123–4130

Maheswaran M, Subudhi PK, Nandi S, Xu Jc, Parco A, Yang DC, Huang N (1997) Polymorphism, distribution, and segregation of AFLP markers in a double haploid rice population. Theor Appl Genet 94:39–45

Marques CM, Araujo JA, Ferreira JG, Whetten R, O'Malley DM, Liu BH, Sederoff R (1998) AFLP genetic maps of *Eucalyptus globulus* and *E tereticornis*. Theor Appl Genet 96:727–737

Meksem K, Leister D, Peleman J, Zabeau M, Salamini F, Gebhardt C (1995) A high-resolution map of the vicinity of the *R1* locus on chromosome V of potato based on RFLP and AFLP markers. Mol Gen Genet 249:74–81

Mullis K, Faloona S, Scharf R, Saiki R, Horn G, Erlich H (1986) Specific enzymatic amplification of DNA *in vitro*: the polymerase chain reaction. Cold Spring Harbor Symp Quant Biol 51:263–273

Munoz JMO (1948) Estudios cromosomicos en el genero *Theobroma* L. Thesis, CATIE/IICA, Turrialba, Costa Rica

Nakagahra M (1986) Geographic distribution of gametophyte genes in wide crosses of rice cultivars. In: Proc Int Rice Genet Symp. IRRI, Manila, Philippines, pp 73–82

Philipps G, Chaubet N, Chaboute ME, Ehling M, Gigot C (1986) Genomic organization and nucleotide sequences of two corn histone H4 genes. Gene 42:225–229

Remington DL, Whetten RW, Liu BH, O'Malley (1999) Construction of an AFLP genetic map with nearly complete genome coverage in *Pinus taeda*. Theor Appl Genet 98:1279–1292

Risterucci AM, Lanaud C, N'Goran JAK, Pieretti I (1996) A saturated linkage map of *Theobroma cacao* L. In: Cocoa Producers Alliance. Proc 12th Int Cocoa Res Conf. Lagos, Nigeria, pp 417–424

Schondelmaier J, Steinrucken G, Jung C (1996) Integration of AFLP markers into a linkage map of sugar beet (*Beta vulgaris* L.). Plant Breed 115:231–237

Senior L, Heun M (1993) Mapping maize microsatellites and polymerase chain reaction confirmation of targeted repeats using a CT primer. Genome 36:884–889

Serikawa T, Kuramoto T, Hilbert P, Mori M, Yamada J, Dubay CJ, Lindpainter K, Gantezn D, Guenet JL, Lathrop GM, Beckmann JS (1992) Rat gene mapping using PCR-analyzed microsatellites. Genetics 131:701–721

Spada A, Caporali E, Marzani G, Portaluppi P, Restivo FM, Tassi F, Falvigna A (1998) A genetic map of *Asparagus officinalis* based on integrated RFLP, RAPD and AFLP molecular markers. Theor Appl Genet 97:1083–1089

Stam P (1993) Construction of integrated genetic maps by means of a new computer package. JoinMap. Plant J 5:739–744

Tanksley SD (1984) Linkage relationships and chromosomal locations of enzyme-coding genes in pepper, *Capsicum annuum*. Chromosoma 89:352–360

Tanksley SD, Miller J, Paterson AH, Bernatzky R (1987) Molecular mapping of plant chromosomes. In: Gustafson JP, Apples RA (eds) Chromosome structure and function. Plenum Press, New York, pp 157–173

Tanksley SD, Ganal MW, Prince JP, De Vicente MC, Bonierbale MW, Broun P, Fulton TM, Giovannoniä JJ, Grandillo S, Martin GB, Messeguer R, Miller JC, Miller L, Paterson AH, Pineda O, Rôder MS, Wingä RA, Wu W Young ND (1992) High density molecular linkage maps of the tomato and potato genomes. Genetics 132:1141–1160

Tautz D, Rentz P (1984) Simple sequences are ubiquitous repetitive components of eucaryotic genome. Nucleic Acids Res 12:4127–4138

Thomas CM, Vos P, Zabeau M, Jones DA, Norcott KA, Chadwick BP, Jones JDG (1995) Identification of amplified restriction fragment polymorphism (AFLP) markers tightly linked to the tomato *Cf-9* gene for resistance to *Cladosporium fulvum*. Plant J 8:785–794

Vos P, Hogers R, Bleeker M, Reijans M, Van de Lee T, Hornes M, Frijters A, Pot J, Peleman J, Kuiper M, Zabeau 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, Kuiper M (1999) Two high-density AFLP linkage maps of *Zea mays* L. analysis of distribution of AFLP markers. Theor Appl Genet 99:921–935

Williams JGK, Kubelik AR, Livak K J, Rafalski JA, Tingey SV (1990) DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. Nucleic Acids Res 18:6531–6535

Weissenbach J, Gyapay G, Dib C, Vignal A, Morissette J, Millasseau P, Vaysseix G, Lathrop M (1992) A second generation map of the human genome. Nature 359:794–801

Zhao X, Kochert G (1993) Phylogenetic distribution and genetic mapping of a (GGC)n microsatellite from rice (*Orysa sativa* L.). Plant Mol Biol 21:607–614

Zabeau M, Vos P (1992) European patent application, publication no: EP 0534858