Theor Appl Genet (1995) 91:987–993

C. Lanaud · A. M. Risterucci · A. K. J. N'Goran
D. Clement · M. H. Flament · V. Laurent · M. Falque

A genetic linkage map of *Theobroma cacao* L.

Received: 3 January 1995 / Accepted: 12 May 1995

Abstract A linkage map of the cocoa genome comprising 193 loci has been constructed. These loci consist of 5 isozymes, 101 cDNA/RFLPs, 4 loci from genes of known function, 55 genomic DNA/RFLPs and 28 RAPDs. A population of 100 individuals derived from a cross between two heterozygous genotypes was used. Segregation analyses were performed with the JoinMap program. Ten linkage groups, which putatively correspond to the ten gametic chromosomes of cocoa, were identified. The map covers a total length of 759 cM with a 3.9 cM average distance between 2 markers. A small fraction (9%) of the markers deviated significantly from the expected Mendelian ratios.

Key words *Theobroma cacao* · Isozyme · RFLP
RAPD · Linkage map

Introduction

Theobroma cacao L. (2n = 2x = 20) (Carletto 1946) is native to Central and South America. Its genome is small (0.4 pg/1C) (Lanaud et al. 1992; Figueira et al. 1992), and three main genetic groups may be distinguished, Criollo, Forastero and their hybrid form, Trinitario. Present-day breeding schemes are based on crosses between parents belonging to these different groups, but in tree crops such as cocoa, with a long generation time, breeding and classical genetic studies do not lead to rapid progress. Genetic markers may be particularly helpful when used to improve breeding strategies and increase our knowledge of the genetics of important agronomic and quality traits. Restriction fragment length polymorphism (RFLP) markers, first proposed by Botstein et al. (1980) have been extensively used to construct genetic maps in many species, such as tomato (Tanksley et al. 1992), potato (Bonierbale et al. 1988; Gebhart et al. 1989, 1991) and *Brassica* (Landry et al. 1991; Chyi et al. 1992). Random amplified polymorphic DNA (RAPD) markers (Williams et al. 1990) have been more recently used and have also helped in the construction or completion of several linkage maps in species such as *Arabidopsis thaliana* (Reiter et al. 1992), slashpine (Nelson et al. 1993) or *Sorghum* (Chittenden et al. 1994).

This paper presents the first linkage map of the cocoa genome containing 193 RFLP, RAPD and isozyme markers.

Communicated by J. S. Beckmann

C. Lanaud (✉) · A. M. Risterucci · A. K. J. N'Goran · M. H. Flament · V. Laurent
CIRAD-Biotrop, BP 5035, F-34032, Montpellier Cedex, France

D. Clement · M. Falque
IDEFOR/DCC and CIRAD-CP, 01 BP, 1827 Abidjan 01, Ivory Coast

Materials and methods

Plant material

Parental accessions

The two parents of the mapping population were:
UPA402: This is an upper Amazon Forastero clone obtained from a sib-mating involving IMC60 and Na 34, two Forastero genotypes collected in Ecuador. Forastero clones are generally vigorous, contain many interesting agronomic and disease resistance traits but generally give a chocolate of only a medium quality.
UF676: This is a Trinitario selection (hybrid between Forastero and Criollo) made by the United Company in Costa Rica. Criollo clones, which originated from Central America, have unfavorable agronomic traits but give a fine chocolate.

Both parental clones are more or less heterozygous, and pedigree and previous isozyme analyses (Lanaud 1987) indicated that UF676 has a higher level of heterozygosity than UPA402. Thus, a specific analytical approach and software for mapping were required.

Mapping population

A segregating population of 100 individuals from a cross between UPA402 (♀) and UF676 (♂) was used for the genetic linkage analyses. This progeny is still standing 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 malate dehydrogenase (MDH), phosphoglucose isomerase (PGI), phosphoglucomutase (PGM), isocitrate dehydrogenase (IDH) and acid phosphatase (ACP).

RFLP analysis

Two libraries were constructed to provide probes for RFLP analysis: cDNA library: mRNA was isolated from 7-day germinated seeds of the clone IFC5 (lower Amazon Forastero). mRNA isolation and cDNA synthesis were performed using Pharmacia kits. The cDNAs were ligated into PUC18 plasmid vector and used to transform *E. coli* strain DH5 α (Gibco BRL)

Genomic library: Total DNA was isolated from the leaf tissue of clone IFC5. An aliquot of 150 µg of total DNA was digested with 30 U of *Pst*I for 2 h and run on a 0.8% agarose gel. DNA fragments 0.5–4 kb in length were electroeluted using the Biotrap procedure (Schleicher and Schuell) and cloned in DH5 α *E. coli* using the plasmid PUC18 as vector. The genomic library was screened for single-copy sequences. The clones were digested with *Pst*I and run on a 0.8% agarose gel with two controls, a clone for a repetitive sequence of wheat rDNA (PTA71, Gerlach and Bedbrook 1979) and a clone for a single-copy sequence of maize ADH (Gerlach et al. 1982). They were blotted onto a nylon membrane (Hybond N^+) and hybridized with 30 ng of genomic cocoa DNA labeled with [^{32}P] dCT. Inserts from either library were amplified by the polymerase chain reaction (PCR), isolated from low-melting agarose gels and used for hybridization.

In addition, 3 probes of known genes were used in this study: histone (H4C14), isolated from maize (Philipps et al. 1986); alcohol dehydrogenase (*Adh1*), isolated from maize (Gerlach et al. 1982), and a rRNA gene (pTA71) isolated from wheat (Gerlach and Bedbrook 1979).

Finally, 36 DNA fragments obtained from RAPD amplifications (See below) were also used as labeled probes to detect RFLPs: the following primers from Operon Technologies INC were used: rOPD15; rOPK7, 8, 9, 13, 15; rOPL5, 11, 19; rOPM17, 20; rOPO15; rOPP3, 14, 16; rOPQ15. DNA fragments were extracted from low-melting agarose gels and hybridized directly, without cloning, onto total DNA of cocoa digested with several restriction enzymes.

Detection of RFLPs

Approximately 5–6 g fresh green leaves was ground to a fine powder using liquid nitrogen and incubated at 65 °C for 3 h in 21 ml of extraction buffer (0.16 *M* sodium citrate, 62 m*M* Na_2EDTA, 1.2% SDS, 6.2 m*M* mercaptoethanol). Eight milliliters of 5 *M* potassium acetate was added, and the solution was then incubated at 0 °C for 10 min. A volume of 20 ml of chloroform: isoamylalcohol (24:1) was added and mixed gently for 5 min. The subsequent emulsion was centrifugated at 4200 *g* for 1 h. The aqueous phase was removed and mixed with the same volume of isopropanol for 2 h at −20 °C. The resulting DNA precipitate was spooled out and resuspended in 5 ml TE buffer (50 m*M* Tris-HCl, pH 8, 10 m*M* Na_2EDTA) with two other precipitates of the same clone obtained in a similar way. The solution was then incubated with 300 µg of RNase T1 for 30 min at 37 °C and kept at least 1 week at 4 °C. DNA was then purified using an ultracentrifugation in a cesium chloride-ethidium bromide gradient.

Five restriction enzymes were used to screen for RFLPs: *Eco*RI, *Eco*RV, *Hin*dIII, *Xba*I and *Bgl*II.

Five micrograms of total DNA was digested overnight by 3 U/µg of each of the five restriction endonucleases. Restriction fragments were separated by electrophoresis on an 0.8% agarose gel in TAE buffer for 16 h at 1.04 V/cm. After depurination in a 0.25 *N* HCl solution, DNA was denatured in 0.4 *N* NaOH and transferred to a Hybond N^+ membrane for 24 h. Probes were labeled with [^{32}P] dCT by random priming. Prehybridizations and hybridizations were performed overnight at 42 °C in 50% formamide, 6 × SSC, 5 × Denhart, 0.5% SDS, 10% dextran sulfate and 25 µg/ml herring sperm DNA. The blots were washed twice at 65 °C for 30 min in 2 × SSC followed by 30 min at 65 °C in 2 × SSC, 0.1% SDS, and finally for 30 min at 65 °C in 0.4 × SSC, 0.1% SDS. Autoradiographs were exposed for 5 days at −80 °C.

RAPD analysis

A modified version of the protocol of Williams et al. (1990) was applied. Amplifications were performed in 20-µl reactions containing 1.5 units of *Taq* polymerase (Promega), 1 ng cocoa DNA, 0.2 m*M* dNTP mix (Pharmacia), magnesium-free buffer (Promega), 2 m*M* $MgCl_2$ and 0.2 µ*M* primer from Operon Technologies Kits (Alameda, Calif.). These were overlaid with one drop of ultra-pure mineral oil. Amplifications were performed in a Techne PHC2 DNA thermal cycler, and the samples were subjected to an initial incubation at 94 °C for 4 min, and then to 43 repeats of the following cycle: 1 min at 94 °C (denaturation), 1 min at 36 °C (annealing) and 2 min at 72 °C (elongation). After the final cycle, the samples were incubated 8 min at 72 °C, then kept at 4 °C prior to analysis. Amplification products were analyzed by electrophoresis on a 2% TBE agarose gel at 3.12 V/cm for 3 h, stained with ethidium bromide and visualized with UV light. The repeatability of RAPD profiles and band was tested by repeating experiments on 2 different days on both parents and on 6 progeny individuals.

Marker nomenclature

RFLP probes were named "cTcCIRX", "gTcCIRX" or "rTcX" according to the following convention: the first letter c or g stands, respectively, for the cDNA or genomic origin; r is used for isolated RAPD genomic fragments; Tc stands for *Theobroma cacao*; CIR corresponds to CIRAD; X stands for the probe accession number.

RAPD loci not used as probes were named "rOPX#" where OP corresponds to Operon Technologies, X is the primer kit letter and the number of the primer in the kit and # is the approximate molecular weight of the band.

Screening of probes and linkage analysis

The segregation of each marker was first tested for goodness of fit to the expected Mendelian segregation ratio using a chi-square test.

Linkage analyses were performed using the program JOINMAP version 1.4 (Stam 1993). The segregation of 202 markers was studied on 100 individuals using a LOD score of 4.0 to identify the linkage groups. The Kosambi mapping function was used to convert recombination frequencies into map distances (Kosambi 1944).

Results and discussion

Genomic library characterization

A total of 274 genomic clones was surveyed for copy number. Clones were classified into single copy, low copy or repeated sequences based on the relative intensity of the hybridization signal and compared to the intensity of the two control bands (repeated and single-copy controls). Of the clones 2% were classified as repeated sequences, 3% as low-copy sequences and the remaining 95% as single-copy sequences.

The restriction endonuclease *Pst*I used for the construction of the library is methylation-sensitive. In maize and tomato, a significant increase in the proportion of single-copy clones has been observed when the clones

were derived from *Pst*I libraries (Burr et al. 1988; Tanksley et al. 1987). The high level of single-copy sequences obtained could therefore be related to the use of *Pst*I. It could also be related to the small size of the cocoa genome and to the low number of repeated sequences, as has been observed using renaturation kinetics (P. Fritz, personal communication).

Probe screening and polymorphism

The results are given in Table 1. A total of 587 RFLP probes was screened using five restriction enzymes. Polymorphism with at least one of the five restriction enzymes was revealed clearly in the progeny by 28% of the cDNA probes and 27% of the genomic probes. Segregating loci were observed in the progeny with a similar percentage of cDNA or genomic probes. This result seems to vary according to the species under study; in lettuce, for example, cDNA probes detected 2.5 times more polymorphism than genomic DNA probes (Landry et al. 1987).

Of the 36 DNA fragments isolated from RAPD amplifications and used as labeled probes 12 corresponded to unique sequences. With these, 3 segregating loci were detected in the progeny.

As both parents (UPA402 and UF676) of the progeny are heterozygous, the following codominant markers, RFLPs or isozymes segregated according to the three possible Mendelian models and were integrated into the analyses: A × H, H × B or H × H (backcrosses or F_2), where A is a homozygous locus in UPA402, B is a homozygous locus in UF676 and H is a heterozygous locus in UPA402 or UF676. RAPD dominant markers were selected to give a band present in one parent, absent in the other parent and which segregated in the progeny. This corresponds, respectively, to a heterozygous locus in one parent and a homozygous locus in the other parent (backcross model). Using the JoinMap package, we were unable to integrate dominant markers segregating according to an "F2" model.

Table 1 Number and type of markers screened for mapping, classified according to three possible Mendelian ratios: A × H, B × H = backcross models; H × H = F_2 model

Marker type	Number of markers [a] screened	Polymorphic loci detected in the progeny, segregating according to three possible models [b]			
		A × H	H × B	H × H	Total
Isozymes	5	5			5
Genomic RFLPs	205	46	5	5	56
cDNA	379	81	9	13	103
Known genes	3	3	1	0	4
RAPDs	94	27	7	–	34
Total	686	162	22	18	202

[a] RFLP probes, primers or enzymatic systems
[b] A = homozygous locus in UPA402 (♀); B = homozygous locus in UF676 (♂); H = heterozygous locus in UPA402 or in UF676

The segregations observed in each class (A × H, H × A or H × H) are in agreement with the genealogies of both parents. UPA402 (♀) resulted from an inbred cross and UF676 (♂) is a Trinitario hybrid between the two most distant genetic groups of cocoa, Criollo and Forastero.

Segregation distortion

Skewed segregation was shown by 18 loci (or 9% of all loci); most (13/18) were only significant at $P = 0.05$. (Fig. 1). Six loci showing distorted segregations are clustered on linkage group 6 (from rOPR3/0.7 to cTcCIR225). Two other loci with normal Mendelian ratios (cTcCIR223 and cTcCIR241) are also located in this same region of 36.4 cM. These 2 loci segregated according to a H × A model, which can only be due to the heterozygosity of the female parent UPA402; the 6 other loci segregated according to a A × H model, corresponding only to the heterozygosity of the male parent UF676. These results suggest the presence of alleles that induce gametic selection and which are located on linkage group 6 of the male parent.

As the number of chi-square tests is high (193), about 10 loci showing distorted segregation at $P = 0.05$ and 2 loci at $P = 0.01$ could be due to sampling variations. A statistical method integrating the total number of chi-square tests made and the linkage between markers could be used to more objectively evaluate the skewed segregations. However, when such a methodology is not applied, the presence of some clustered markers that display segregation distortion (as in group 5, 6 or 8) may confirm their existence in these regions of the genome.

Distorted segregations have been frequently reported and may be high: 25.5% in potato (Gebhart et al. 1989), 12–59% in *Brassica oleracea* (Kianian and Quiros 1992), 30% in banana (Faure et al. 1993), with clusters of skewed loci often being observed (Prince et al. 1993; Cai et al. 1994). Chromosomal rearrangements (Tanksley 1984; Faure et al. 1993) or gametic selection due to certation as shown in rice (Nakagahra 1986) has been suggested to explain distorted segregations. In *Theobroma cacao*, neither structural rearrangements nor the genes responsible for gametic selection are known. However, an incompatibility system has been described in cocoa (Knight and Rogers 1995; Cope 1962), and the presence of incompatibility alleles could also serve as an hypothesis to explain the distorted segregations in group 6.

Construction of the genetic linkage map

The linkage map is shown in Fig 1.

A total of 193 loci have been assigned to ten linkage groups. These loci correspond to 5 isozymes, 55 genomic RFLPs, 101 cDNA RFLPs, 4 known-function genes and 28 RAPDs. The size of the linkage groups ranged from 48 cM to 105 cM. The total length of the map was 759 cM, with an average distance of 3.9 cM between

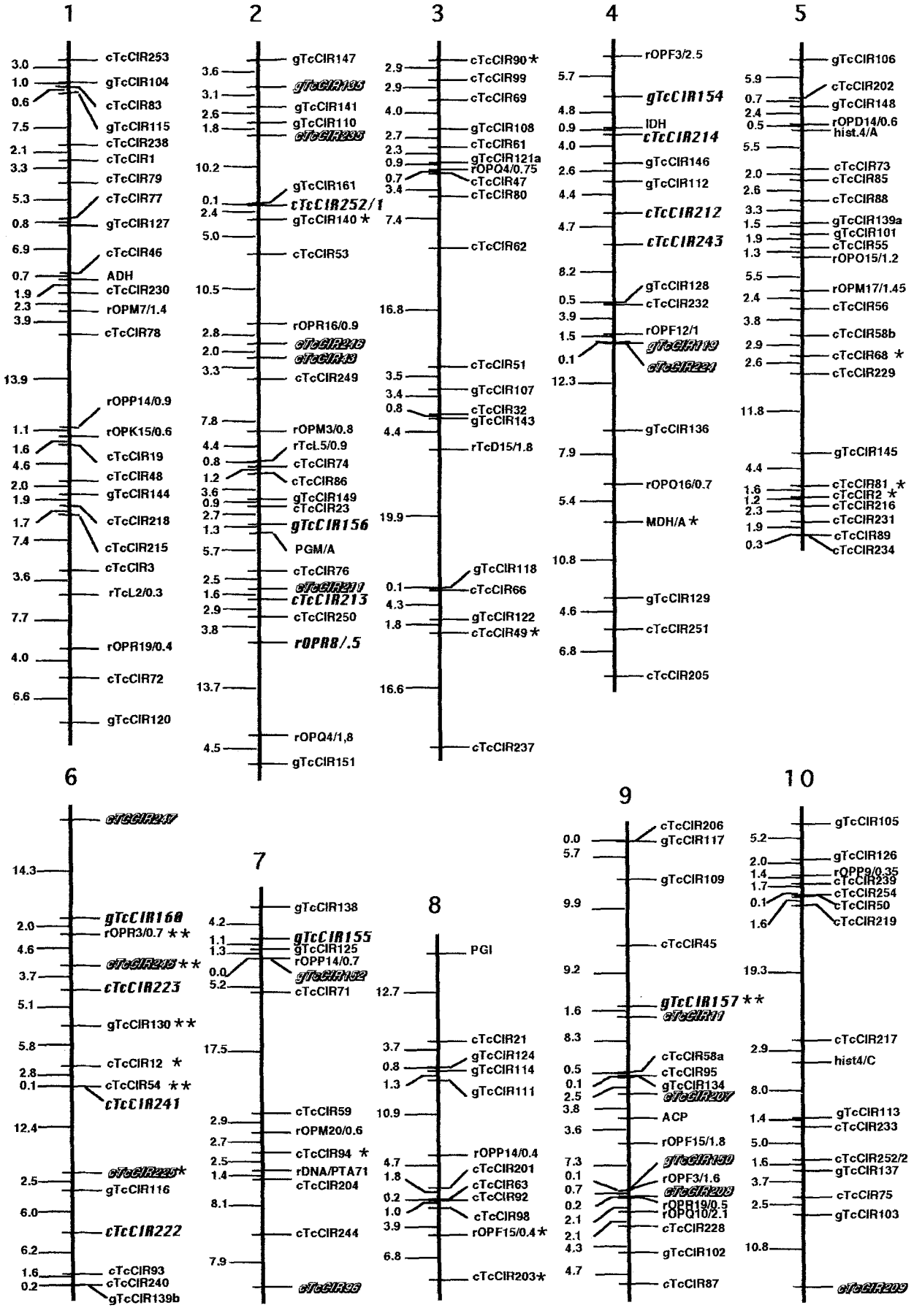
1
3.0 cTcCIR253
1.0 gTcCIR104
0.6 cTcCIR83
7.5 gTcCIR115
2.1 cTcCIR238
3.3 cTcCIR1
cTcCIR79
5.3 cTcCIR77
0.8 gTcCIR127
6.9 cTcCIR46
0.7 ADH
1.9 cTcCIR230
2.3 rOPM7/1.4
3.9 cTcCIR78
13.9 rOPP14/0.9
1.1 rOPK15/0.6
1.6 cTcCIR19
4.6 cTcCIR48
2.0 gTcCIR144
1.9 cTcCIR218
1.7 cTcCIR215
7.4 cTcCIR3
3.6 rTcL2/0.3
7.7 rOPR19/0.4
4.0 cTcCIR72
6.6 gTcCIR120
2
3.6 gTcCIR147
gTcCIR135
3.1 gTcCIR141
2.6 gTcCIR110
1.8 cTcCIR235
10.2 gTcCIR161
0.1 cTcCIR252/1
2.4 gTcCIR140 ★
5.0 cTcCIR53
10.5 rOPR16/0.9
2.8 cTcCIR240
2.0 cTcCIR13
3.3 cTcCIR249
7.8 rOPM3/0.8
4.4 rTcL5/0.9
0.8 cTcCIR74
1.2 cTcCIR86
3.6 gTcCIR149
0.9 cTcCIR23
2.7 gTcCIR156
1.3 PGM/A
5.7 cTcCIR76
2.5 cTcCIR211
1.6 cTcCIR213
2.9 cTcCIR250
3.8 rOPR8/.5
13.7 rOPQ4/1,8
4.5 gTcCIR151
3
2.9 cTcCIR90★
2.9 cTcCIR99
cTcCIR69
4.0 gTcCIR108
2.7 cTcCIR61
2.3 gTcCIR121a
0.9 rOPQ4/0.75
0.7 cTcCIR47
3.4 cTcCIR80
7.4 cTcCIR62
16.8 cTcCIR51
3.5 gTcCIR107
3.4 cTcCIR32
0.8 gTcCIR143
4.4 rTcD15/1.8
19.9 gTcCIR118
0.1 cTcCIR66
4.3 gTcCIR122
1.8 cTcCIR49★
16.6 cTcCIR237
4
rOPF3/2.5
5.7 gTcCIR154
4.8 IDH
0.9 cTcCIR214
4.0 gTcCIR146
2.6 gTcCIR112
4.4 cTcCIR212
4.7 cTcCIR243
8.2 gTcCIR128
0.5 cTcCIR232
3.9 rOPF12/1
1.5 gTcCIR119
0.1 cTcCIR224
12.3 gTcCIR136
7.9 rOPO16/0.7
5.4 MDH/A★
10.8 gTcCIR129
4.6 cTcCIR251
6.8 cTcCIR205
5
gTcCIR106
5.9 cTcCIR202
0.7 gTcCIR148
2.4 rOPD14/0.6
0.5 hist.4/A
5.5 cTcCIR73
2.0 cTcCIR85
2.6 cTcCIR88
3.3 gTcCIR139a
1.5 gTcCIR101
1.9 cTcCIR55
1.3 rOPO15/1.2
5.5 rOPM17/1.45
2.4 cTcCIR56
3.8 cTcCIR58b
2.9 cTcCIR68 ★
2.6 cTcCIR229
11.8 gTcCIR145
4.4 cTcCIR81 ★
1.6 cTcCIR2 ★
1.2 cTcCIR216
2.3 cTcCIR231
1.9 cTcCIR89
0.3 cTcCIR234
6
cTcCIR247
14.3 gTcCIR168
2.0 rOPR3/0.7 ★★
4.6 cTcCIR245 ★★
3.7 cTcCIR223
5.1 gTcCIR130 ★★
5.8 cTcCIR12 ★
2.8 cTcCIR54 ★★
0.1 cTcCIR241
12.4 cTcCIR225★
2.5 gTcCIR116
6.0 cTcCIR222
6.2 cTcCIR93
1.6 cTcCIR240
0.2 gTcCIR139b
7
gTcCIR138
4.2 gTcCIR155
1.1 gTcCIR125
1.3 rOPP14/0.7
0.0 cTcCIR152
5.2 cTcCIR71
17.5 cTcCIR59
2.9 rOPM20/0.6
2.7 cTcCIR94 ★
2.5 rDNA/PTA71
1.4 cTcCIR204
8.1 cTcCIR244
7.9 cTcCIR33
8
PGI
12.7 cTcCIR21
3.7 gTcCIR124
0.8 gTcCIR114
1.3 gTcCIR111
10.9 rOPP14/0.4
4.7 cTcCIR201
1.8 cTcCIR63
0.2 cTcCIR92
1.0 cTcCIR98
3.9 rOPF15/0.4★
6.8 cTcCIR203★
9
0.0 cTcCIR206
gTcCIR117
5.7 gTcCIR109
9.9 cTcCIR45
9.2 gTcCIR157 ★★
1.6 cTcCIR11
8.3 cTcCIR58a
0.5 cTcCIR95
0.1 gTcCIR134
2.5 cTcCIR207
3.8 ACP
3.6 rOPF15/1.8
7.3 gTcCIR150
0.1 rOPF3/1.6
0.7 cTcCIR208
0.2 rOPR19/0.5
2.1 rOPQ10/2.1
2.1 cTcCIR228
4.3 gTcCIR102
4.7 cTcCIR87
10
gTcCIR105
5.2 gTcCIR126
2.0 rOPP9/0.35
1.4 cTcCIR239
1.7 cTcCIR254
0.1 cTcCIR50
1.6 cTcCIR219
19.3 cTcCIR217
2.9 hist4/C
8.0 gTcCIR113
1.4 cTcCIR233
5.0 cTcCIR252/2
1.6 gTcCIR137
3.7 cTcCIR75
2.5 gTcCIR103
10.8 cTcCIR209

Fig. 1 Linkage map of 193 loci in the genome of *Theobroma cacao;* these correspond to 5 isozymes, 52 genomic RFLPs (gTcCIRX), 3 genomic fragments from RAPD amplifications (rTcX), 101 cDNA RFLPs (cTc CIRX), 4 loci from genes of known function and 28 RAPDs (rOPX). The markers segregated according three different models shown in the following examples: A × H Backcross model; H × B Backcross model; H × H F_2 model. *Asterisk* indicates loci showing distorted segregation at $P = 0.05$ (*) or at $P = 0.01$ (**)

markers. Intervals between 2 adjacent markers were no greater than 20 cM. All markers (162) segregating according to the A × H backcross model were integrated into the ten linkage groups. Nine markers could not be assigned to a linkage group: (1) 1 of the 18 markers segregating according to the F_2 (H × H) model; this marker showed a distorted segregation that could explain its non-assignment to the linkage groups; (2) 8 of the 22 markers segregating according to the H × B backcross model.

Due to the heterozygosity of both parents, the markers segregated in the progeny according to three possible models: A × H, H × H and H × A. A × H markers were the most numerous (162/193), and it would be possible, using a more classical linkage analysis package such as MAPMAKER (Lander et al. 1987), which is adapted to the analysis of backcross progenies, to construct a linkage map with only these markers. However, in spite of the interest for quantitative trait loci (QTL) analyses, this would result in the loss of part of the information provided by H × H and H × A markers. Indeed, even if the heterozygosity level of UPA402 is lower, some important traits such as resistance to *Phytophthora* may segregate in this parent, and it is important to be able to localize them using H × H and H × A markers in order to detect linkages between favorable or unfavorable traits. Analyses made separately on both types of markers A × H and H × A, using MAPMAKER did not enable linkage groups to be constructed for each parent, as was the case for alfalfa (Echt et al. 1994), due to the low number of H × A markers. JOINMAP was better adapted to these studies and enabled integration of the three types of marker, since the H × H markers serve as bridges integrating all markers. Similar strategies of linkage map construction in plants resulting from crosses between heterozygous parents have already been developed, notably for potato (Ritter et al. 1990) and Citrus (Cai et al. 1994), where JOINMAP was used to integrate two types of marker (Aa × Aa and Aa × aa) into a core map established with aa × Aa markers. In this case, these markers could be assigned to linkage groups of the core map, but ambiguities still remained in the order of markers within the groups. Relative to map established using MAPMAKER, map length was reduced when using JOINMAP. In our data, few markers were of the type H × H or H × B. This might explain the non-assignment of several markers to linkage groups or the difficulties encountered in localizing them unambiguously. Indeed, four linkage groups (1, 3, 5 and 8) contained no marker segregating according to the H × H model, and in two other linkage groups (4 and 7) markers segregating according to the H × B model could be ordered differently: they could have a symmetric position compared to the unique or very close (H × H) markers used to integrate all the segregations. In our map, in addition to the 162 markers segregating according to a single model (A × H), 27 markers were integrated unambiguously as a result of their linkage with several H × H markers. However, 4 other markers could not be ordered unambiguously due to a lack of H × H markers in those groups. We did not compare our results with a map established with MAPMAKER, but have compared them with a map constructed using JOINMAP with the 162 A × H markers only: map order was conserved and genetic distances were unchanged except for a few small variations never exceeding 1–2 cM.

Conclusion

This is the first report of a complete linkage map of cocoa comprising ten linkage groups that putatively corresponding to the ten gametic chromosomes of *T. cacao* (Carletto 1946). The small genome of cocoa (1C = 0.4 pg) is convenient for rapid, in-depth coverage of the entire genome. In this map, the average distance between 2 adjacent markers was 3.9 cM, with no gap exceeding 20 cM. Additional markers will be placed to complete this map and to define the ends of linkage groups. However, this average distance is already sufficiently small to allow mapping of QTL. On the basis of the deduced haploid genome size ($0.388.10^9$ bp) (Lanaud et al. 1992) and on the total length of the genetic map (759 cM), the physical distance per unit of genetic distance could be estimated at 511 kbp/cM, which is only 2.2 times that of *A. thaliana* (Reiter et al. 1992). This could be favorable to the development of map-based cloning of genes of interest.

This map consists of a majority of RFLP markers, which are suited to the exchange of probes between several laboratories and to the joining of different maps. It will provide efficient tools for the improvement of cocoa breeding strategies. Many traits of interest in cocoa are not easy to manage due to the long generation time of this tree crop and to the complex genetic determinism involved. The mapping of these traits will allow them to be split into their genetic components and thus help identify in the genome the most important regions involved in the variation of characteristics. Among these, resistance to black pod disease due to *Phytophthora palmivora* and *P. megakarya* is of particular interest. This disease is widespread all over the world and is responsible for substantial losses in yield (Gregory and Maddison 1981; Djiekpor et al. 1982). Chemical control of black pod disease is effective but costly, time-consuming and polluting. Genetic control of this disease may be

an alternative solution but requires a good knowledge of the genetic basis of resistance. Resistance in the field may be dependent on several traits, such as intrinsic resistance, time of pod production, duration of pod maturation, etc. Mapping strategies will help identify the most important factors involved in field resistance by locating common QTL and evaluating the contribution of each QTL to the observed variation. Identification of early screening markers would allow a better mastering of the accumulation of resistance genes and a gain of time during the breeding steps. Similar mapping strategies could help increase our knowledge of the biochemical and genetic basis of several quality traits, such as bean size, butter content or aroma. The involvement of particular biochemical compounds in aroma traits could also be confirmed by locating QTL common to biochemical and flavor traits.

The progeny used for mapping have parents of different genetic origins and also show variation in a large number of traits such as quality, pod and flower characteristics or resistance to *Phytophthora*. However, this progeny is divided into two populations (55 and 45 trees) located in two different sites and, because of this, QTL analyses will provide only indicative results for the most important QTL on each sub-population. Other larger populations located in the same site are now available and are being evaluated for QTL identification of resistance and quality traits.

Acknowledgments We acknowledge Dr. M. Lorieux for his computer assistance.

References

Bonierbale MW, Plaisted RL, Tanksley SD (1988) RFLP maps based on a common set of clones reveal modes of chromosomal evolution in potato and tomato. Genetics 120:1095–1103

Botstein D, White RL, Skolnick M, Davis RV (1980) Construction of a genetic linkage map in man using restriction fragment length polymorphisms. Am J Hum Gene 32:314–331

Burr B, Burr FA, Thompson KH, Albertson MC, Stuber CW (1988) Gene mapping with recombinant inbreds in maize. Genetics 188:519–526

Cai Q, Guy CL, Moore GA (1994) Extension of the linkage map in *Citrus* using random amplified polymorphic DNA (RAPD) markers and RFLP mapping of cold-acclimatation-responsive loci. Theor Appl Genet 89:606–614

Chittenden LM, Schertz KF, Lin YR, Wing RA, Paterson AH (1993) A detailed RFLP map of *Sorghum bicolor* × *S. propinquum*, suitable for high-density mapping suggests ancestral duplication of *Sorghum* chomosomes or chromosomal segments. Theor Appl Genet 87:925–933

Chyi YS, Hoenecke ME, Semyk JL (1992) A genetic linkage map of restriction fragment length polymorphism loci for *Brassica rapa* (syn. *campestris*). Genome 35:746–757

Cope FW (1962) The mechanism of pollen incompatibility in *Theobroma cacao*. Heredity 17:157–182

Djiekpor EK, Partiot M, Lucas P (1982) La pourriture brune des cabosses du cacaoyer due à *Phytophthora sp.* au togo : détermination des espèces responsables. Café-Cacao-Thé 16:297–108

Echt CS, Kidwell KK, Knapp SJ, Osborn TC, McCoy TJ (1994) Linkage mapping in diploid alfalfa (*Medicago sativa*). Genome 37:61–71

Faure S, Noyer JL, Horry JP, Bakry F, Lanaud C, Gonzalez de Leon D (1993) A molecular marker-based linkage map of diploïd 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, Bredbrook 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

Gregory PH, Maddison AG (1981) Epidemiology of *Phytophthora* on cocoa in Nigeria. Phytopathological paper, 25

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

Knight R, Rogers H (1955) Incompatibility in *Theobroma cacao*. Heredity 9:67–69

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. Café Cacao Thé 30:271–280

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

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

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

Landry BS, Kesseli RV, Farrara B, Michelmore RW (1987) A genetic map of lettuce (*Lactuca sativa* L.) with restriction fragment length polymorphism, isozyme, disease resistance and morphological markers. Genetics 116:331–337

Landry BS, Hubert N, Etoh T, Harada JJ, Lincoln SE (1991) A genetic map for *Brassica napus* based on restriction fragment length polymorphisms detected with expressed DNA sequences. Genome 34:543–554

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

Nelson CD, Nance WL, Doudrick RL (1993) A partial genetic linkage map of slash pine (*Pinus elliottii* Engelm. var 'elliottii') based on random amplified polymorphic DNAs. Theor Appl Genet 87:145–151

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

Prince J, Pochard E, Tanksley SD (1993) Construction of a molecular linkage map of pepper and a comparison of synteny with tomato. Genome 36:404–417

Reiter RS, Williams JGK, Feldmann KA, Rafalski JA, Tingey SV, Scolnik PA (1992) Global and local genome mapping in *Arabidopsis thaliana* by using recombinant inbred lines and random amplified polymorphic DNAs. Proc Natl Acad Sci USA 89:1477–1481

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

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