Theor Appl Genet (1994) 88:193–198

V. Laurent · A. M. Risterucci · C. Lanaud

Genetic diversity in cocoa revealed by cDNA probes

Received: 28 April 1993 / Accepted: 9 September 1993

Abstract The variability of the cocoa (*Theobroma cacao*) nuclear genome was investigated. A total of 203 cocoa clones was surveyed for restriction fragment length polymorphisms (RFLPs) using four restriction endonuclease and 31 seed cDNA probes. A high level of polymorphism has been found. This study points to a structuring of the species that fits with the distinction between the Criollo and Forastero populations. These results combined with previously obtained nuclear rDNA and mtDNA data allow us to propose new hypotheses on the origin and evolution of the different cocoa populations.

Key words *Theobroma cacao* · RFLP · diversity study seed cDNA

Introduction

Cocoa trees, *Theobroma cacoa*, are native to humid tropical regions of the American continent. They are classified into three different groups depending on geographical location and morphological characters (Cheesman 1944). The three morpho-geographical groups are designated Criollo, Forastero and Trinitario.

The Forastero group is subdivided into Lower Amazon Forastero and Upper Amazon Forastero according to the geographical origin of the clones. Lower Amazon Forastero trees were initially cultivated in the Amazon basin and were the first to be introduced into Africa. Although Upper Amazon Forastero trees were sampled more recently, they are often used in breeding programs due to their strength, precocity, and resistance to disease. This group clusters a series of highly-diverse populations.

The Criollo group is composed of trees with thick, white or rosy beans yielding the most flavored and finest chocolate. They were the first cocoa trees to be domesticated. They have been cultivated in Central America for 2000 years, but at present are infrequently grown because of their weaknesss and their susceptibility to disease and pests. Almost all the known Criollo clones are either cultivated or subspontaneous, very few truly wild-type forms exist.

The Trinitario group is made up of hybrid forms of the first two groups.

The high morphological variability found among the different populations in the Upper Amazon region has resulted in this region being considered the center of origin of the species (Cheesman 1944).

Until now, all breeding strategies have relied on crosses between clones of various groups to obtain fine tasting chocolate from vigorous and productive trees. There is, therefore, a general need to improve our understanding of the relationships between cocoa populations. A previous enzymatic study has shown that the variability of the Upper Amazon Forastero group embraces the global variability of the species, confirming that the Upper Aamazom region could correspond to the first center of diversity (Lanaud 1987). The presence of rare alleles in individual populations suggests that different modes of diversification may exist. Previous RFLP studies of mitochondrial DNA (Laurent et al. 1993b) and nuclear ribosomal DNA (Laurent et al. 1993a) have confirmed the originality of some populations but, contrary to the isozyme study, they have pointed up a general distinction between Forastero and Criollo types.

In order to obtain a more accurate description of the diversity of the species, nuclear cDNA probes have been used. Although coding regions comprise the smallest part of nuclear DNA, cDNA probes have often been used in RFLP studies of a number of plants, including potato (Debener et al. 1991) and alfafa (Brummer et al. 1991). Moreover, for tomato (Miller and Tanksley 1990), lentil (Havey and Muehlbauer 1989), peanuts (Paik-Ro et al. 1992), and lettuce (Landry et al. 1987), cDNA probes have revealed more polymorphism than random genomic probes.

Communicated by P. M. A. Tigerstdt

V. Laurent · A. M. Risterucci · C. Lanaud (✉)
AGETROP/CIRAD BP5035, 34032 Montpellier cedex, France

Table 1 Cocoa clones studied for restriction fragment length polymorphism with cDNA probes (Clones are clustered according to their presumed classification (*C* Criollo, *T* Trinitario, *L* Lower Amazonian Forastero and *U* Upper Amazonian Forastero) and country of origin or selection)

Group	Origin	Clones
T	Trinidad	ICS (6, 16, 46, 53, 75, 89, 95, 98), ACT2-11
T	Grenada	GS (29, 26)
T	Venezuela	CHUAO (24, 120), CNS (22, 23)
T	Ivory Coast	IFC (6, 7, 11, 19, 413, 420, 422)
T	Ghana	IFZ (304, 305), ACU85, W41
T	Indonesia	DR1, G23, WA40
T	Samoa	LAFI7
T	Ecuador	EQX (27, 94, 100, 107), MOQ (122, 216, 413, 647, 663)
T	Honduras	MT1, TJ1
T	Nigeria	N38
T	Mexico	RIM (8, 15, 19, 76, 105, 113)
T	Colombia	SPEC (54-2, 138-8, 160-9), SC (5, 6)
T	Guatemala	SGU3
T	Cameroon	SNK (12, 109)
T	Costa Rica	UF (10, 221, 296, 667, 676), CC (10,39)
T	Panama	UF168
C	Venezuela	BO204, BOC210, CHO31, CHUAO (49, 211), Providencia 201, Hernandez 212 CATA (201, 211), MTC201, CUM (209, 214) OC (61, 63, 73, 77), PV (2, 4, 6), ZEA (1, 206), JS (202, 206, 210), Porcelana (3, rojo) POB, POC, POR, POR (210, 211)
C	Nicaragua	ICS (39, 40, 48, 60, 100), IS201
C	Trinidad	ICS84
C	Mexico	La Esmida
C	Indonesia	G8
C	Nigeria	CD8/6
C	Costa Rica	LAF (1, 2, 3)
C	Peru	PA35
C	Ghana	Q7
C	Colombia	SEPC185-4
L	Brazil	Comun tipico, Para, SIC864, IFC361, SIAL (42, 70, 325) ERJOH(1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15)
L	Guiana	GU (144, 154, 346, 349)
L	Ecuador	ECNR, EET59, Nacional
L	Ivory Coast	IFC (1, 2, 4, 5, 15), SF23
L	Ghana	IFC (303, 307, 414)
L	Costa Rica	MAT (1-6, 1-9)
L	Venezuela	VEN (1, C4, 5, 11, 15, 20, 31)
U	?	Amazon15-15
U	Colombia	EBC (5, 6, 10), SPA (5, 11, 17)
U	Peru	PA (4, 13, 120, 150), P (1, 2, 16, 32A), MO (9, 81, 98), NA (32, 79)
U	Ecuador	LCT-EEN (37, 84, 109, 127, 67, 202, 295, 325, 326, 355, 371) IMC (5, 31, 67, 78), SCA (6, 9, 12)
U	Ghana	IFC312, T (60/887, 79/501, 85/799)
U	Ivory Coast	UPA (401, 402, 410, 413, 608, 620)

Materials and methods

Plant material

A sample of 203 cocoa clones from various areas belonging to the three different morpho-geographical groups was screened (Table 1). Both cultivated and wild clones were represented. Dried leaves of the different clones were supplied by IDEFOR-DCC (Ivory Coast), IRCC (France), CRU (Trinidad), CENIAP (Venezuela), CEPEC (Brazil) and CATIE (Costa Rica).

DNA probes

Thirty-one cDNA probes obtained from germinated seed mRNA were used to screen individuals. For the diversity studies, the most polymorphic enzyme/probe combination was applied for each cDNA. Sixteen cDNA probes were hybridized on *Eco*RI-digested clones, five on *Eco*RV digests, five on *Hin*dIII digest and six on *Xba*I digests.

RFLP procedures

cDNA library. Total RNA was obtained from fresh seed tissue in guanidine (MacDonald et al. 1987) followed by purification on a cesium-chloride gradient. Isolation of mRNA and cDNA synthesis were both performed using Pharmacia kits. The cDNAs were ligated onto the pUC18 plasmid vector and recombinants used to transform the DH5α strain of *E. coli*. After isolation of cDNA plasmids by minipreparation (Chowdhury 1991), inserts were amplified by PCR and the resulting amplification products were isolated from low-melting-point agarose gels.

DNA extraction. DNA extractions were performed as previously described in Laurent et al. (1993a). Five micrograms of total DNA were digested overnight by 3 UE/μg of the restriction endonucleases *Eco*RV, *Xba*I, *Hin*dIII or *Eco*RI. Restriction fragments were separated by electrophoresis in 0.7% agarose gels in TAE buffer (Sambrook et al. 1989) for 16 h at 1.04 V/cm. Fragments were fixed on nylon Hybond N + membranes in 0.4 N NaOH by Southern blotting. Probes were labeled with α-^{32}PdCTP by random priming (Feinberg and Vogelstein 1983). Prehybridizations and hybridizations were performed at 68 °C overnight in 6 × SSC, 5 × Denhart, 0.5% SDS, 25 μg/ml of herring sperm DNA. Blots were washed twice at 68 °C for 30 min in 2 × SSC followed by 30 min in 2 × SSC, 0.1% SDS and finally 30 min in 0.1 × SSC, 0.1% SDS. Autoradiographs were exposed for 1 weak at −80 °C one intensifying screen.

Data and statistical analyses

Each polymorphic band was scored for presence and absence. A factorial analysis of correspondences, FAC (Benzecri 1973), was performed with these new variables. A cluster analysis dendogram was constructed by the UPGMA method on the coordinates of the first seven factors of the FAC. All computations were performed with ADDAD (1983) software.

Results

Most of the cDNA probes revealed two or three variable bands (Fig. 1) that behave as allelic bands (i.e., each individual bears only one or two bands). A few probes revealed more bands. Only one probe revealed eight polymorphic bands which could correspond to a repeated gene. Thus 87 variable bands were scored from 31 probes. Variable bands were used in a factorial analysis of correspondences and in a cluster analysis.

The first plane of the FAC covers 25% of the total variability with a major contribution of 56 variable

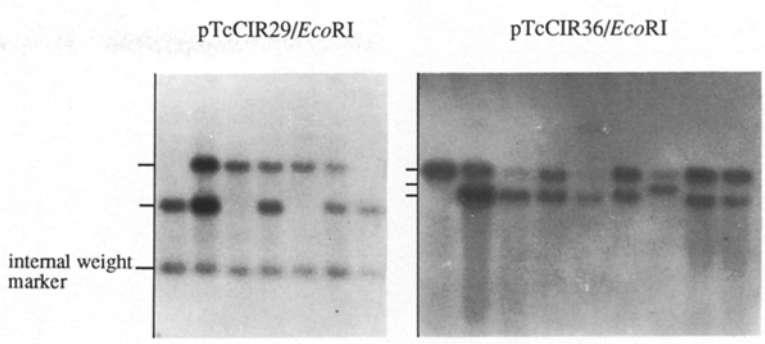

Fig. 1 Example of hybridization patterns obtained with two cDNA probes. Bands revealed by cDNA probes are indicated on the left, the internal molecular weight marker is 1.5 kb

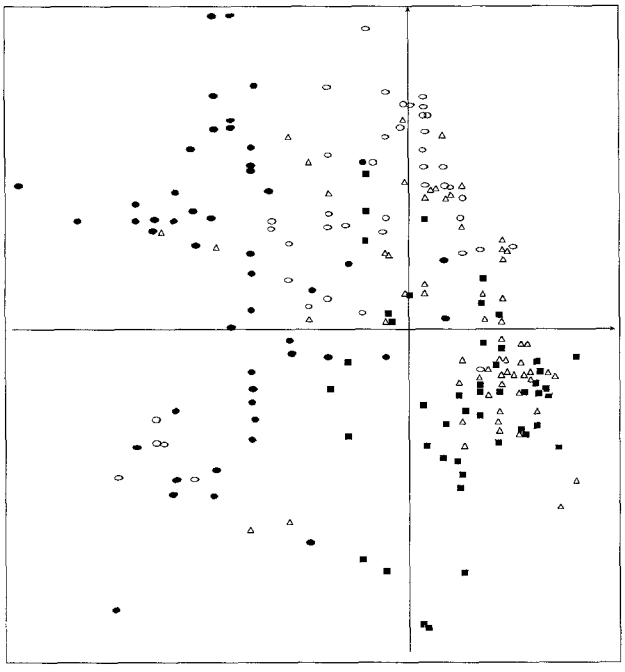

Fig. 2 Distribution of the various clones on the first FAC plane. Morpho-geographic group are indicated: ■, Criollo; ●, Upper Amazon Forastero; ○ = Lower Amazon Forastero and △, Trinitario

bands from the total of 87 (Fig. 2). Although the FAC does not reveal a striking discrimination between Criollo and Forastero, the clones of the two groups are characterized by different combinations of variables. Upper Amazon Forastero show the widest range of variability occupying all the left part of the plane. Criollo show a different range of variability organized around a cluster of quite homogeneous clones. A similar organization of the variability occurs for Lower Amazon Forastero. Indeed, most of the cultivated Lower Amazon Forastero are clustered in a reduced part of the plane whereas the wild Lower Amazon Forastero show a wide range of variability. Most of the wild Lower Amazon Forastero (VEN, ERJOH) occupy a median position between cultivated Lower Amazon Forastero and Upper Amazon Forastero. A few Forastero from Guiana are isolated from the others and fall within the range of Upper Amazon Forastero variability. Trinitario show a large range of variability covering the variability of both Criollo and Lower Amazon Forastero. Very few of them are represented among the Upper Amazon Forastero. The African Trinitario are clustered with the cultivated Lower Amazon Forastero whereas the American Trinitario are clustered with Criollo. Most of Trinitario that are located with Upper Amazon Forastero were collected in farms in the Upper Amazon region (MOQ and EQX).

In order to obtain a more synthetic representation of the variability of the sample, a cluster analysis was performed using the first seven factors of the FAC (Fig. 3). These seven factors represent 55% of the total variability with a major contribution of 76 variables. The dendogram confirms the general distinction between Upper Amazon Forastero, on the one hand, and Lower Amazon Forastero, Trinitario and Criollo, on the other hand, that was revealed by the FAC analysis. At a lower level, the Lower Amazon Forastero and some Trinitario and discriminated from Criollo and other Trinitario.

Discussion

cDNAs have proved to be an excellent source of low-copy probes for a number of studies. Indeed, for lentil, 88% of the probes from a cDNA library were low copy in contrast to 41% for a *Pst*I genomic library (Havey and Muehlbauer 1989). For cacao, out of the 252 cDNAs screened, 99% were single-copy probes. A similar high level of single-copy probes (98%) was obtained from 450 probes screened for a cocoa genomic *Pst*I library (unpublished data). A number of different studies have pointed up a positive correlation between the amount of repeated DNA sequences in a genome and its DNA content (reviewed by Lapitan 1992). The high level of single-copy cocoa probes found in both the genomic and the cDNA library could be related to the small size of the cocoa genome, 1C = 0.4 pg (Lanaud et al. 1992), and its corresponding low level of repetitive DNA. Indeed, tomato (1C = 0.7 pg), *Gossypium raimondy* (1C = 0.68 pg) and *Citrus* (1C = 0.6 pg) are all characterized by a small genome size and a low number of repetitive

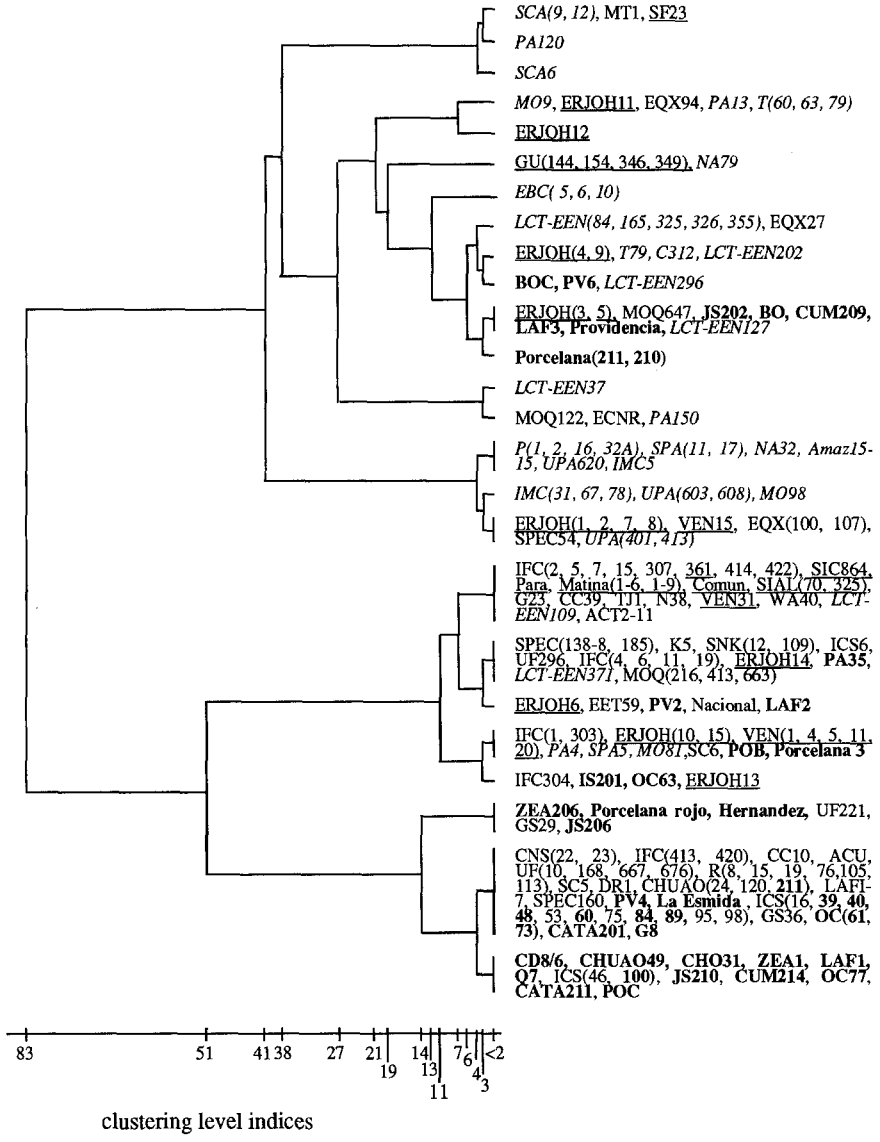

Fig. 3 Dendogram of the cluster analysis. Morpho-geographical groups are indicated: Criollo (*bold*), Trinitario (*plain text*), Lower Amazon Forastero (*underline*) and Upper Amazon Forastero (*italic*)

sequences (Zamir and Tanksley 1988; Geever et al. 1989; Jarell et al. 1992).

For a number of studies, cDNA probes have revealed a higher level of polymorphism than random genomic probes (Landry et al. 1987; Havey and Muehlbauer 1989; Miller and Tanksley 1990; Paik-Ro et al. 1992) or isozymes (McGrath and Quiros 1992). This high level of polymorphism found with cDNA probes is believed to be provided by introns and flanking sequences (Miller and Tanksley 1990; McGrath and Quiros 1992) rather than by translated sequences. In contrast to the observations of McGrath and Quiros, cocoa cDNA probes have revealed a similar number of alleles or bands to isozymes (Lanaud 1987). Further, cocoa RFLP studies with genomic probes should help define the origin of the polymorphism revealed by cDNA probes. Especially, it could indicate whether the polymorphism originates from non-coding regions or from translated sequences.

Restriction fragment length polymorphism may arise by point mutations or large scale rearrangements. Point mutations are considered to affect only a single restriction site whereas rearrangements may modify several restriction sites for different enzymes. Thus, most of the studies that have revealed simultaneous polymorphism for several restriction enzymes have assumed them to be caused by rearrangements (McCouch et al. 1988; Wang and Tanksley 1989; Graner et al. 1990). On the other hand, when a probe revealed polymorphism with only one enzyme, this polymorphism was assumed to be due to point mutations (Chase et al. 1991; Jarell et al. 1992; Neuhausen 1992; Webb et al. 1992). Since half of the probes used in this study revealed polymorphism in cocoa in association with only one restriction enzyme, whereas the other half revealed polymorphism with more than one enzyme (data not shown), this suggests that polymorphism in cocoa is due to both point mutations and rearrangements.

The structuring of the nuclear variability is in general accordance with the distinction between the morpho-geographical Criollo and Forastero types. Such a structur-

ing of the species into groups corresponding to these two morpho-geographical types has also been found for nuclear rDNA (Laurent et al. 1993a) and mitochondrial DNA (Laurent et al. 1993b). Moreover, the present study has pointed up a distinction between Criollo and the white beaned Forastero. In contrast, the isozyme analysis of the species showed that the variability of Upper Amazon Forastero embraces the global variability of the species (Lanaud 1987) which agrees with the Upper Amazon region being considered as the center of origin of the species. According to Cheesman (1944), the origin of Criollo would have relied on human selection of some white-beaned genotypes in the Upper Amazon region. The general discrimination of Criollo from Forastero at both nuclear and mitochondrial levels rather suggests that these two types have differentiated independently on both sides of the Andean barrier as suspected by Cuatrecasas (1964).

In addition to this global distinction between the Forastero and Criollo types, cDNA probes point up a difference between Lower Amazon Forastero and Upper Amazon Forastero that differs from their clustering for both rDNA and mtDNA. As rDNA is reported to maintain homogeneity among populations, because of its particular mode of evolution (Delseny et al. 1986), while mtDNA is usually considered to reflect older changes than cDNA, this result could reflect a recent differentiation of Lower Amazon Forastero and Upper Amazon Forastero.

One Lower Amazon population (GU) is well separated from the Lower Amazon pool. A separation of the Guyanese population from the Lower Amazon pool has also been found for isozyme (Lanaud 1987) and rDNA analyses (Laurent et al. 1993a). As a possible explanation for its particular isozyme pattern, Lanaud had considered that this population had become differentiated in a distinct forest refuge area during the last glaciation periods of the Pleistocene. During the aridity period of the Quaternary, the tropical forest shrank to small areas receiving enough rainfall for their persistence, thus causing speciation and secondary differentiation of many of the modern deep-forest-inhabiting plants and animals (Simpson and Haffer 1978). Common refuge areas identified for several plant and animal species (Fig. 4) correspond to high rainfall regions (Ratisbona 1976; Snow 1976) and include the Upper Amazon and Guiana. When populations such as GU were isolated in refuge areas, their rDNA could rapidly have become specific whereas they would have retained some mtDNA and nuclear single-copy DNA characteristics of individuals adapted to high rainfall regions. On the contrary, the other Lower Amazon Forastero would have differentiated when propagating into the Amazonian basin.

The clustering of Trinitario with Criollo and Lower Amazon Forastero could reflect the history of their creation. Indeed, most of the Trinitario were obtained first from crossings between Lower Amazon Forastero and Criollo and then from crossings between Trinitario and Criollo or Trinitario and Lower Amazon Forastero. These successions of backcrosses could explain the return of the hybrid form to the parental types. The clustering of the African Trinitario with Lower Amazon Forastero reflects the backcrosses with Lower Amazon Forastero from which the African Trinitario were descended. Since Criollo and Upper Amazon Forastero were introduced into Africa only recently, the first crosses were between the Lower Amazon Forastero and Trinitario. At present, backcrosses with Lower Amazon Forastero could allow us to conserve the adaptation of Lower Amazon Forastero to African agricultural and climatical conditions. A similar hypothesis could be made for the clustering of American Trinitario with Criollo. Indeed, we can assume that backcrosses in America are preferentially performed with Criollo in order to maintain the Criollo quality characters.

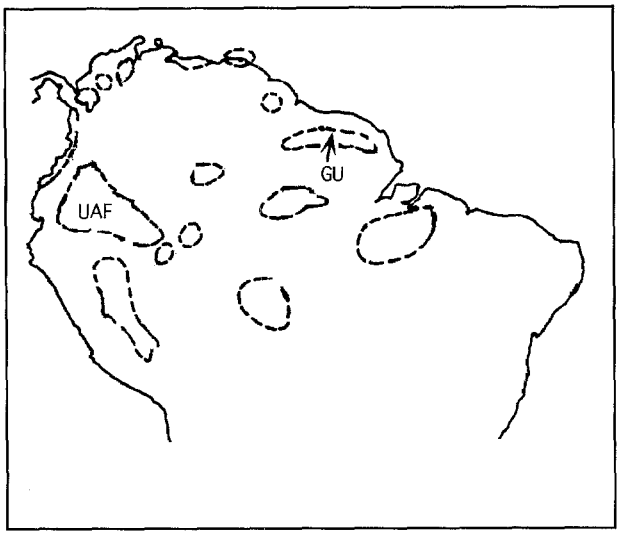

Fig. 4 Location of Pleistocene forest refugia based on the distribution of four families of tropical forest trees (from Prance 1973). The location of Upper Amazon Forastero populations (*UAF*) and Guyanese population (*GU*) are also indicated

This study of nuclear single-copy sequences, in combination with both rDNA and mtDNA studies, allow us to propose new differentation schemes for cocoa populations. The general distinction of Criollo and Forastero supports the hypothesis of Cuatrecasas (1964) for two distinct subspecies, *T. cacao* subsp. *cacao* and *T. cacao* subsp. *sphaerocarpum*. Forastero seem to have differentiated from the Upper Amazon region into Upper Amazon Forastero and Lower Aamazon Forastero whereas Criollo would have differentiated independently on the other side of the Andean barrier. Since very little is known about the wild forms of Criollo, several questions on their evolution remain unsolved. The major one involves the original distribution of Criollo in Central America and north of South America. New collections currently being undertaken in north South America and Central America could provide new insights into the evolution of Criollo.

Acknowledgements We thank Dr. Reyes, Dr. Fonseca, Dr. Morera, Professor Spence, M. Sounigo and Ms. Johnson for providing us with plant material. We are grateful to Dr. Bartley and Dr. Lockwood for information about the origins of clones.

References

ADDAD (1983) Manuel de référence. Association pour le Développement et la diffusion de l'Analyse des Données., Paris

Benzecri JP (1973) L'analyse des correspondances. In: Dunod (ed) L'analyse des données, tome 2 Paris, pp 616

Brummer EC, Kochert G, Bouton JH (1991) RFLP variation in diploid and tetraploid alfalfa. Theor Appl Genet 83:89–96

Chase CD, Ortega VM, Vallejos CE (1991) DNA restriction fragment length polymorphisms correlate with isozyme diversity in *Phaseolus vulgaris* L. Theor Appl Genet 81:806–811

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

Chowdhury (1991) One step "miniprep" method for the isolation of plasmid DNA. Nucleic Acids Res 19:2792

Cuatrecasas J (1964) Cacao and its allies: a taxonomic revision of the genus *Theobroma*. Bull US National Museum, Smithsonian Institution, Washington 35:379–614

Debener T, Salamini F, Gebhardt C (1991) The use of RFLP (Restriction Fragment Length Polymorphisms) detects germplasm introgressions from wild species into potato (*Solanum tuberosum* ssp. *tuberosum*) breeding lines. Plant Breed 106:173–181

Delseny M, Grellet F, Tremoussaygue D, Raynal M, Panabieres F (1986) Structure, evolution et expression de l'ADN nucléaire des plantes supérieures. In: Colloque de la Société de Botanique, Orsay, pp 21

Feinberg AP, Vogelstein B (1983) A technique for radiolabelling DNA restriction endonuclease fragments to high specific activity. Anal Biochem 132:6–13

Geever RF, Katterman FRH, Endrizzi JE (1989) DNA hybridization analyses of a *Gossypium* allotetraploid and two closely related diploid species. Theor Appl Genet 77:553–559

Graner A, Siedler H, Jahoor A, Hermann RG, Wenzel G (1990) Assesment of the degree and the type of restriction fragment length polymorphism in barley (*Hordeum vulgare*). Theor Appl Genet 80:826–832

Havey MJ, Muehlbauer FJ (1989) Linkages between restriction fragment length, isozyme, and morphological markers in lentil. Theor Appl Genet 77:395–401

Jarrell DC, Roose ML, Traugh SN, Kupper RS (1992) A genetic map of *Citrus* based on the segregation of isozymes and RFLPs in an intergeneric cross. Theor Appl Genet 84:49–56

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 espèce. Doctorat d'état, Paris IX

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

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

Lapitan NLV (1992) Organization and evolution of higher plant nuclear genomes. Genome 35:171–181

Laurent V, Risterucci AM, Lanaud C (1993a) Variability for nuclear ribosomal genes within *Theobroma cacao*. Heredity 71:96–103

Laurent V, Risterucci AM, Lanaud C (1993b) Chloroplast and mitochondrial DNA diversity in *Theobroma cacao*. Theor Appl Genet 87:81–88

McCouch SR, Kochert G, Yu ZH, Wang ZY, Khush GS, Coffman WR, Tanksley SD (1988) Molecular mapping of rice chromosomes. Theor Appl Genet 76:815–829

McDonald RJ, Swift GH, Przybyla AE, Chirgwin JM (1987) Isolation RNA using guanidium salts. Methods Enzymol 152:219–227

McGrath JM, Quiros CF (1992) Genetic diversity at isozyme and RFLP loci in *Brassica compestris* as related to crop type and geographical origin. Theor Appl Gnenet 83:783–790

Miller JC, Tanksley TD (1993) Effect of different restriction enzymes, probe source, and probe length on detecting restriction fragment length polymorphism in tomato. Theor Appl Genet 80:385–389

Neuhausen SL (1992) Evaluation of restriction fragment length polymorphism in *Cucumis melo*. Theor Appl Genet 83:379–384

Paik-Ro OG, Smith RL, Knauft DA (1992) Restriction fragment length polymorphism evaluation of six peanut species within the *Arachis* section. Theor Appl Genet 84:201–208

Prance GT (1973) Phytogeographic support for the theory of Pleistocene forest refuges in the Amazon Basin, based on evidence from distribution patterns in Caryocaraceae, Chrysobalanaceae, Dichapetalaceae and Lecythidaceae. Acta Amazonica 3:5–28

Ratisbona LR (1976) The climate of Brazil. In: W. Scherdtfeger (eds) Climates of Central and South America. Elsevier, New York, pp 219–294

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning: a laboratory manual (second edition). Cold Spring Harbor Laboratory, Cold Spring Harbor, New York

Simpson BB, Haffer J (1978) Speciation patterns in the Amazonian forest biota. Annu Rev Ecol Syst 9:497–518

Snow JW (1976) The climate of northern South America. In: W. Scherdtfeger (eds) Climates of Central and South America. Elsevier, New York, pp 295–404

Wang ZY, Tanksley SD (1989) Restriction fragment length polymorphism in *Oryza sativa* L. Genome 32:1113–1118

Webb DM, Knapp SJ, Tagliani LA (1992) Restriction fragment length polymorphism and allozyme linkage map of *Cuphea lanceolata*. Theor Appl Genet 83:528–532

Zamir D, Tanksley SD (1988) Tomato genome is composed largely of fast-evolving, low-copy-number sequences. Mol Gen Genet 213:254–261