Cacao

Claire Lanaud, Juan-Carlos Motamayor and Olivier Sounigo

Cacao has been domesticated from ancient times. It was cultivated about 2000 to 3000 years ago by the pre-Colombian civilizations, particularly the Mayas and the Aztecs (Paradis, 1979). Cacao beans were used in a beverage called 'chocolatl' for which cacao, maize, pimento, and other aromatic plants were mixed and ground. Cacao beans were so highly valued that they constituted a currency of exchange. Even before the arrival of the Spanish, cacao travelled through the trade routes of the Mayas and Aztecs, but also with the help of the Pipil-Nicaraos (Young, 1994; Coe and Coe, 1996). In 1585, the Spanish exported cacao to Spain, where the secret of chocolate—a sugared version of the beverage—was guarded for about 40 years. The popularity of this chocolate drink then spread throughout Europe. To respond to the growing demand, cacao plantations were extended into Central America and new plantations were established in several of the Caribbean islands, such as Trinidad, in 1525, and Jamaica. Cacao trees from Central America, particularly Costa Rica, were introduced in Venezuela by the Spanish, but it is possible that cacao trees were already cultivated in the southwestern part of the country before Spanish colonization (Pittier, 1933; Bergman, 1969; Motamayor, 2001). Around 1750, the French planted cacao in Martinique and in Haiti, and the Portuguese planted it in Belem and Bahia.

Cacao trees have grown in Asia and the Pacific islands since 1560 (Wood, 1991; Young, 1994). During that time, Criollo of Venezuela were introduced in the island of Celebes and then in Java, by the Dutch. In 1614, Criollo from Mexico were transplanted in the Philippines by the Spanish. In 1798, cacao cultivation was introduced in Madras, India, from the island of Amboine and was brought from Trinidad to Ceylon by the British. From Ceylon, the cacao tree was introduced to Singapore and Fiji in 1880, to Samoa in 1883, to Queensland in 1886, and to Bombay and Zanzibar in 1887. Cacao was cultivated in Malaysia from 1778 and in Hawaii in 1831.

The introduction of cacao to Africa is more recent. Spanish or Portuguese navigators imported it to the island of Sao Tome in 1822, then to the island of Fernando Poo in 1855 (Burle, 1952). Other introductions were then made by

Swiss missionaries from Surinam. The first cacao seeds were sown on the African continent in 1857. The first cacao that spread through West Africa from Ghana originated from the Amazon basin (Forastero Amelonado) and then, in 1920, hybrid forms Trinitario and Criollo were imported and were hybridized with the local Amelonado (Toxopeus, 1985). Each of these transfers was made with a small number of genotypes. As a consequence, the genetic basis of the cacao plants initially cultivated in West Africa was very narrow and their origin uncertain.

The technique of chocolate manufacture was perfected with the invention of presses to extract cacao butter in 1828 (Enriquez, 1985). The technique of fermenting and roasting beans made it possible to develop the chocolate aroma from cultivars of cacao other than the varieties of Criollo traditionally cultivated, which require little fermentation.

At the beginning of the 20th century, nearly 80% of the world production of cacao, about 115,000 tonnes, was produced in Central and South America (Braudeau, 1969). In 1997, the production reached nearly 2.7 million tonnes. Côte d'Ivoire is presently the primary world producer with 1.12 million tonnes per year, followed by Ghana and Indonesia, with 330,000 tonnes each.

TAXONOMY AND GENETIC RESOURCES

Taxonomy

Cacao belongs to the family Sterculiaceae and the genus *Theobroma*. *Theobroma cacao* L. (2n = 2x = 20) is a tree originating from humid tropical regions of the northern part of South America and Central America. Even though the first attempts at domestication and cultivation were made in Central America, Cheesman (1944) considers the centre of origin of cacao to be the upper course of the Amazon, near the equatorial Andes. It is in this region that Pound (1938) observed the greatest morphological variation. Cacao has a small genome, estimated at 0.4 pg per haploid genome (Figueira et al., 1992; Lanaud et al., 1992).

The taxonomy of the genus *Theobroma* has been studied by botanists since the end of the 19th century (Bernoulli, 1869; Schumann, 1886; Pittier, 1930; Chevalier, 1946). A later and more complete study is that of Cuatrecasas (1964), which divided the genus *Theobroma* into 6 sections and 22 species, which are spread across the American continent between 18°N and 15°S. One of these sections is made up of the single species *T. cacao*. The classification proposed by this author is based on the mode of germination, the architecture of the trees, and the characters of the fruits and flowers. Among all these species, only *T. cacao* is economically important. However, *T. grandiflora*, the cupuassu, is also exploited in Brazil, the pulp of its fruits being used in the manufacture of drinks and sorbets.

The species *T. cacao* is composed of a large number of interfertile populations, highly variable morphologically. The plants are autogamous or allogamous depending on their genetic origin. A system of gametosporophytic self-incompatibility enforces the allogamy in certain populations (Knight and Rogers, 1955; Bouharmont, 1960; Cope, 1962; Glendinning, 1966).

Classifications proposed for the species *T. cacao* are never entirely satisfactory. They are very often schematic, given the intense genetic exchange between populations of varying genetic origin. Morris (1882) was the first to propose a classification of cacao trees into two groups: Criollo and Forastero[1]. Pittier (1930) placed each species into a group: *T. leiocarpum* for the Forastero and *T. cacao* for the Criollo. Cuatrecasas (1964) considered them two subspecies: *T. cacao* ssp. *cacao* for the Criollo and *T. cacao* ssp. *sphaerocarpum* for the Forastero. At present, a classification into two morphogeographic groups, the Criollo and the Forastero, is generally used. A third group, the Trinitario, combines the hybrid forms of the two former groups.

The Forastero

The Forastero group combines a large number of wild populations and cultivated varieties originating from South America. The cacao trees of this group are found from Ecuador to the Guyanas. It is the vigorous trees that have several types of disease resistance. Their beans, generally purple and flat, but sometimes white and rounded in certain populations, give a medium to good quality cacao. The varieties of the Nacional type, which originated from Ecuador, and the classification of which in this group is now under question (Enriquez, 1992), produce a fine, aromatic cocoa much in demand by chocolate manufacturers.

The Forastero of Brazil and Venezuela

The Forastero of the Amazon basin and the Orinoco valley are widely cultivated throughout the Amazon basin. The most typical form is the Amelonado. It is this type of cacao tree, self-compatible, that was first introduced in Africa. Many collections of wild material in the Amazon region in Brazil have demonstrated high morphological diversity (Barriga et al., 1984).

The Forastero of Peru, Ecuador, and Colombia

During the course of several research expeditions on genotypes resistant to witches' broom disease, a disease caused by the fungus *Crinipellis perniciosa*, Pound (1938, 1943) described the populations of Peru and Ecuador for the

[1] The terms Criollo and Forastero come from Venezuela and distinguish the traditionally cultivated local cacao trees, the Criollo, from foreign cacao trees, the Forastero, also called Trinitario, introduced later from Trinidad. Thus, the term Forastero originally related to all the different varieties of Criollo. By extension, the term Criollo, associated with cacao of high quality resulting from thick, light-coloured, aromatic beans, has been used to designate cacaos of the same quality coming from other countries such as Nicaragua and Mexico.

first time. These have high morphological diversity, with pods of variable shape and colour, green or pale green or slightly pigmented in the zones close to Colombia. Later studies in the upper Amazon and, in particular, eastern Ecuador (Allen and Lass, 1983) indicated indigenous and wild populations of cacao. Contrary to the observations of Pound at Peru, all the populations of eastern Ecuador seem morphologically quite similar. They are characterized by large trees, young leaves of a pale green colour, and green and rough pods containing a high portion of white beans. These cacaos have characters in common with the Criollo varieties of Central or South America, but the pods of Criollo may be more elongated and completely red, contrary to those of Ecuador cacao, which are only slightly pigmented.

Hybrid forms are also found in Ecuador, particularly the Refractario. These are descendants of an introduction of Trinitario made in Ecuador around 1890 from some pods from Trinidad. These Trinitario were then naturally crossed with the Nacional variety originally cultivated in Ecuador. This particularly vigorous and early-producing material, even on poor soils, has been used by most planters. There has thus been considerable genetic exchange between cacao types of different origins. During the occurrence of witches' broom disease, which spread throughout almost all the districts and caused serious losses in production in the 1920s to 1930s, some trees, called Refractario, appeared more tolerant. A small percentage of seeds from these surviving trees were then planted.

The Forastero of Guyana

The first observations of wild cacao in Surinam revealed populations highly homogeneous in terms of pod shape, of the Amelonado type, and distinct from the cultivated types in the region (Myers, 1930). Spontaneous cacao were observed in southwestern French Guyana from 1729 (Lecomte and Challot, 1897). Studies organized in French Guyana revealed certain populations, particularly those of the upper Camopi, that have a significant phenotypic variability (Sallee, 1987). These populations are distinguished from Forastero of the lower Amazon by their pods, the shape of which varies from that of the Amelonado to that of the Angoleta, with more or less marked verrucosity and medium to large size (Lachenaud and Sallee, 1993). The origin of the types cultivated in Guyana is uncertain. According to Guisan (1825), they come from the natural cacao forests located above the Camopi.

The Criollo

The Criollo were the first cacao domesticated by the Mayan and Aztec civilizations. These are the only varieties that were cultivated in all of Central America and northern South America up to the 17th century.

The Criollo, self-compatible, are found from Mexico to Colombia and in Venezuela in the cultivated or subspontaneous form (Soria, 1973; Reyes, 1992). They are slow-growing, are more sensitive to disease and insects than the

Forastero, and have a high morphological diversity (Soria, 1970a, 1970b). The fruits are elongated, with an acuminate point, a smooth or rough surface, red or green colour before maturity, and a poorly lignified cortex. The beans are of variable size but, most often, large and thick and white or pink in colour. Criollo beans are generally rounded, unlike the flat Forastero beans of the Amazon basin and Peru. These bean characteristics, associated with poor lignification of the cortex, are normally used to classify the cacao clones in the Criollo group. The cultivated forms of Criollo vary from a smooth type of pod, such as Porcelana of Venezuela, to Cundeamor type, with highly verrucose pods, found in Mexico, Colombia, and Venezuela. A peculiar form, the Pentagona, is found in the old plantations of Mexico, Guatemala, Nicaragua, and Venezuela.

Apart from some traditional varieties of Criollo present in the collections, more vigorous and productive Criollo types have been collected in the plantations of the last decades. It is these clones of Criollo, generally representing this genetic group in the collections, that are called 'modern Criollo' in the rest of this study.

The Trinitario

The Trinitario combine all the hybrid forms between Criollo and Forastero of the lower Amazon or the Orinoco valley, which are the source of these varieties. The Trinitario were cultivated first of all in Trinidad, then in Venezuela and Central America, where they were mixed with the Criollo in the traditional plantations.

Genetic Resources

About 40 collections of wild or cultivated material have been put together since 1930, mainly on the Forastero of Peru, Brazil, Ecuador, Colombia, Venezuela, and Guyana and, more recently, on the Criollo of Central America, Venezuela, and Colombia.

The most important collections from these studies are conserved at CRU, Cocoa Research Unit, at Trinidad (2500 genotypes), at CEPLAC (Comissão Executiva do Plano de Lavoura Cacaueira, in Brazil (2000 genotypes), and at CATIE, the Centro Agronómico Tropical de Investigación y Enseñanza, in Costa Rica (700 genotypes). Each of these collections has its speciality: The one in Trinidad is rich in Forastero of the upper Amazon. That of Brazil includes material from collection expeditions of the Amazonian region in Brazil and the varieties cultivated at Bahia. The CATIE collection has mainly Trinitario and Criollo varieties. The CIRAD collection in French Guyana has descendants of about 200 spontaneous mother plants collected in that country.

International exchanges of plant material are subjected to a quarantine of two years in order to prevent viral and fungal diseases. Three quarantine

stations are in operation: that of the University of Reading in the United Kingdom, that of CIRAD in France, and that of CRU in Barbados.

An international database was formed during the 1990s. It includes information on more than 27,000 genotypes conserved in 43 collections and is updated regularly (Wadsworth and Harwood, 2000). Another database, Tropgene-db, presently established in CIRAD, contains molecular data of more than 400 genotypes.

The genetic improvement of cacao, targeted on different objectives depending on the country—quality, production, resistance—relies always on the creation of hybrids between progenitors belonging to different populations. Since it is difficult to establish a classification solely on the basis of morphological characters, biochemical and molecular markers are used to refine the genetic organization of species and to study the processes of domestication.

ORGANIZATION OF GENETIC DIVERSITY

Morphological Diversity

To describe the morphological diversity, several authors have attempted to define the most efficient morphological descriptors that take into account the intra- and interclonal variance (Enriquez and Soria, 1968; Engels et al., 1980; Soria and Enriquez, 1981; Engels, 1983a, b; Bekele, 1992; Bekele et al., 1994; Raboin and Paulin, 1993). The IBPGR (1981) published a list of 65 morphological descriptors to characterize the genetic resources of cacao. However, in order to reduce the time required for characterization in studies on a large number of genotypes, this list has often been shortened.

Engels (1986) studied the diversity of 294 clones of the CATIE collection using 39 qualitative and quantitative descriptors pertaining to the morphology of flowers, leaves, and fruits. This analysis indicates a structuration between the Criollo and Forastero types as well as a significant diversity in each of these groups. In this analysis, the Trinitario are grouped with the Criollo.

N'goran (194) analysed the diversity of 52 clones belonging to Forastero, Criollo, and Trinitario groups for 9 characters of beans and pods. This analysis confirmed the structuration observed by Engels, with a differentiation between Forastero on the one hand, and Criollo and Trinitario on the other.

A later study (Bekele and Bekele, 1998) on 100 clones resulting from 24 populations indicated a structuration of populations depending on geographic origin.

These studies together show that morphological markers allow an overall structuring of the diversity of different populations of cacao in collections. They are accessible to all, but they are time-consuming, and they are difficult to use because they vary according to the environment.

Enzymatic Diversity

Enzymatic diversity of clones or of cacao populations has been evaluated by several authors (Lanaud and Berthaud, 1984; Amefia, 1986; Atkinson et al., 1986; Lanaud, 1986a, b, 1987; Yidana et al., 1987; Ronning and Schnell, 1994; Warren, 1994; Sounigo et al., 1996).

Lanaud (1987) analysed the diversity of 296 genotypes using 6 polymorphic enzymatic systems representing 9 loci. On the set of the samples analysed, a total of 30 alleles were identified. With the exception of *PAC1*, all the alleles are found in the populations originating in the upper Amazon. *PAC1* is on the other hand frequent among the Criollo and Trinitario. Another allele, *MDHA1*, is also frequent among the Crillo and the Trinitario and rare among the Forastero.

The populations that have the largest number of alleles per locus are those of the upper Amazon, with 1.8 to 2.2 alleles per locus. The populations of Venezuela and Guyana are the least variable with 1 to 1.5 alleles per locus. The percentage of polymorphic loci is the highest in populations of the upper Amazon, where it is higher than 50%. The mean heterozygosity is highest among the American Trinitario (0.36), the modern Criollo (0.29), and the Refractario EQX of Ecuador (0.35). The heterozygosity is lowest for the African Amelonado (0.04), the Forastero of Guyana (0.06), and the two populations of Venezuela that were analysed (0.1 and 0).

This variability in rates of heterozygosity within the populations of Forastero could result from the system of reproduction of the trees and their system of self-incompatibility. The cacaos of the upper Amazon are highly self-incompatible and preferentially allogamous, while the Forastero of the lower Amazonian basin are self-compatible and thus preferentially autogamous.

A correspondence analysis (CA) was done using 31 variables observed on 286 individuals. No clear structuration between Forastero and Criollo or Trinitario appeared on the first CA axis, while a slight structuration was visible on the 3-4 plane. The Forastero from the upper Amazon present the largest genetic diversity. Those of Ecuador are as variable as those of Peru, contrary to the morphological observations of Allen and Lass (1983). Comparatively, the other populations of Forastero analysed are much less diverse. The Forastero cultivated in Guyana are clearly differentiated from wild Forastero collected in the south of Guyana on the first CA plane (Fig. 1). The direct use of spontaneous local cacao seems thus excluded in the explanation for the origin of cultivated cacao, the provenance of which remains unknown. A Venezuelan origin is, however, probable (Lachenaud and Sallee, 1993). The diversity of the wild population of Guyana is relatively low, taking into account the large number of individuals analysed (92). The two populations of Guyana, cultivated and wild, are clearly differentiated from the lower Amazonian Forastero (Amelonado) cultivated in West Africa

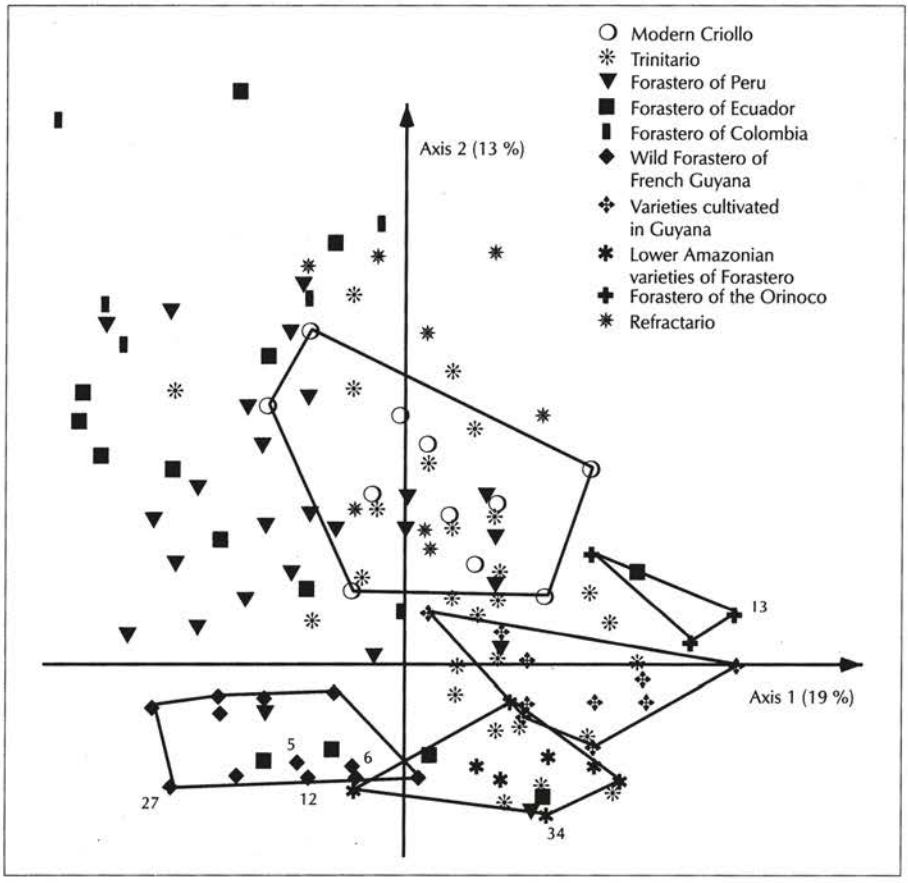

Fig. 1. Primary CA plane of 296 individuals analysed by isozymes, according to Lanaud (1987). The numbers shown on the figure indicate the number of individuals superimposed on a given point.

and from populations of the Orinoco valley in Venezuela. The cultivated Guyanese cacao, however, are more closely related to the Forastero of Venezuela. On the first axis of the CA, the wild Guyanese cacao are closer to certain Forastero of Peru and Ecuador than to Forastero of Venezuela and the cultivated forms of Guyana.

There is close genetic proximity between the modern cultivated varieties of Criollo and the Trinitario cultivated in America and Africa, which is easily explained by the significant genetic exchange that occurs on plantations. The variability of Trinitario cultivated in Africa is very high and extends from the Forastero Amelonado types to the Criollo. This situation is linked to the history of the introduction of cacao in Africa, with a first wave of Forastero Amelonado and a second wave of Trinitario. The genetic exchange that

followed resulted in certain hybrid types closer to the Forastero Amelonado type.

Using four enzymatic systems, Warren (1994) analysed the diversity of seven populations of Forastero of Colombia, Ecuador, and Peru and two populations of Trinitario, at the rate of 10 individuals per population. He observed a large number of alleles for the PGI system (10 alleles against 5 in the study of Lanaud, 1987). The Shannon diversity indexes indicate a greater diversity in the populations of Forastero of Ecuador than in those of Peru. The hybrid Trinitario forms also appear variable. The author concludes that, if there really is a centre of diversity of wild cacao, it must be located not in Peru, as suggested by Cheesman (1944), but further north, in Ecuador and Colombia.

Ronning and Schnell (1994) studied the enzymatic diversity of 86 clones of cacao resulting from Forastero, modern Criollo, Trinitario, and an undefined hybrid group using six enzymatic systems corresponding to eight loci. The allelic frequencies and the genetic distances obtained confirm the differentiation between Forastero and Criollo. The values of parameters of genetic diversity are overall similar to those of other perennial ligneous species: $H_T = 0.295$, $H_S = 0.266$, $G_{ST} = 0.096$ (Hamrick and Godt, 1990).

Sounigo et al. (1996) studied the enzymatic diversity of 487 clones of cacao belonging to 28 populations or groups of accessions present in the CRU collection at Trinidad, using five polymorphic enzyme systems. The analysis of parameters of genetic diversity showed highly variable allelic richness and heterozygosity. Certain populations such as the Forastero of Guyana and the Trinitario of the Dominican Republic or Martinique are poorly diversified, with an average of 1.3 to 1.8 alleles per locus and 6% to 16% of mean heterozygosity per locus and per individual. The reverse is observed for certain hybrid populations of Trinitario and Refractario, which have 2.2 to 2.5 alleles per locus and 42% to 50% of mean heterozygosity per locus and per individual. Other populations of Trinitario show high heterozygosity but a low Shannon index of diversity or allelic richness. On the other hand, certain populations of Forastero and of Refractario of Ecuador have low heterozygosity but great allelic richness.

The low diversity of certain populations, such as those of Martinique and the Dominican Republic, reflect the homogeneity of the material resulting probably from a small number of introductions into these Caribbean islands. On the other hand, the high diversity values of Refractario clearly indicate their multiple origins.

Estimated from the Nei diversity index, the mean intrapopulation diversity, H_{pop}, is 0.77 and the interpopulation diversity is 0.23. Thus, three quarters of the total diversity is explained by the intrapopulation diversity. These results agree with those of Ronning and Schnell (1994) and Russel et al. (1993). They are characteristic of allogamous perennial species in which significant genetic exchange has occurred (Hamrick et al., 1992). The

hierarchical clustering (HC) based on the genetic distances of Nei is represented in Fig. 2. Two major groups are observed at $d = 0.2$. One comprises the populations of Refractario and Trinitario as well as the population of IMC Forastero of Peru. This group is in turn structured into two subgroups: the first subgroup comprises the Trinitario differentiated according to their geographic origin, and the second subgroup includes all the Refractario as well as the populations of IMC (Forastero), ICS, and CC (Trinitario). The other major group combines all the other populations of Forastero, including the wild clones of Guyana, which seem closely related to the MO and PA populations of Peru. In this analysis, the population of Scavina is close to that of Forastero of Ecuador LCTEEN.

This analysis, despite the small number of loci analysed, thus allows us to reveal an overall structuration corresponding to the different genetic groups, and a geographic structuration of populations within certain groups.

Molecular Diversity

The level of heterozygosity of clones

The percentage of heterozygous loci in 300 clones has been evaluated using RFLP after hybridization of 33 genomic probes and cDNA. These probes have been mapped on a reference map (Lanaud et al., 1995), which can be used to find the genetic determinism of the markers used.

The clones that have the highest rate of heterozygosity are the Trinitario (86% for the UIT clones). High heterozygosity is also observed for certain modern varieties of Criollo (73% for CHO42), as well as in certain Forastero of the upper Amazonian region, such as IMC105 (42%). The lowest levels of heterozygosity are found in the old varieties of Criollo (Porcelana, 3%) or lower-Amazonian Forastero (Pará, 4%). Some Forastero of Guyana (GU154), Venezuela (VENC20), Colombia (EBC5), Peru (P2), and the lower Amazon (Catongo) are totally homozygous.

Molecular diversity revealed by RFLP

The molecular diversity of cacao populations, revealed by RFLP at the nuclear or cytoplasmic level, was evaluated by several authors (Laurent et al., 1993a, b, 1994; Figueira et al., 1994; Lerceteau et al., 1997; Motamayor et al., 1997).

Nuclear diversity

Laurent et al. (1994) analysed 201 genotypes belonging to different morphogeographic groups. The diversity of total nuclear DNA was analysed using 31 cDNA probes that enabled the identification of 87 polymorphic bands. Despite continuous variation between the groups, due particularly to the large number of Trinitario hybrids, a fairly clear structuration appeared on axis 1 of a CA (Fig. 3) between Forastero on the one hand, and modern

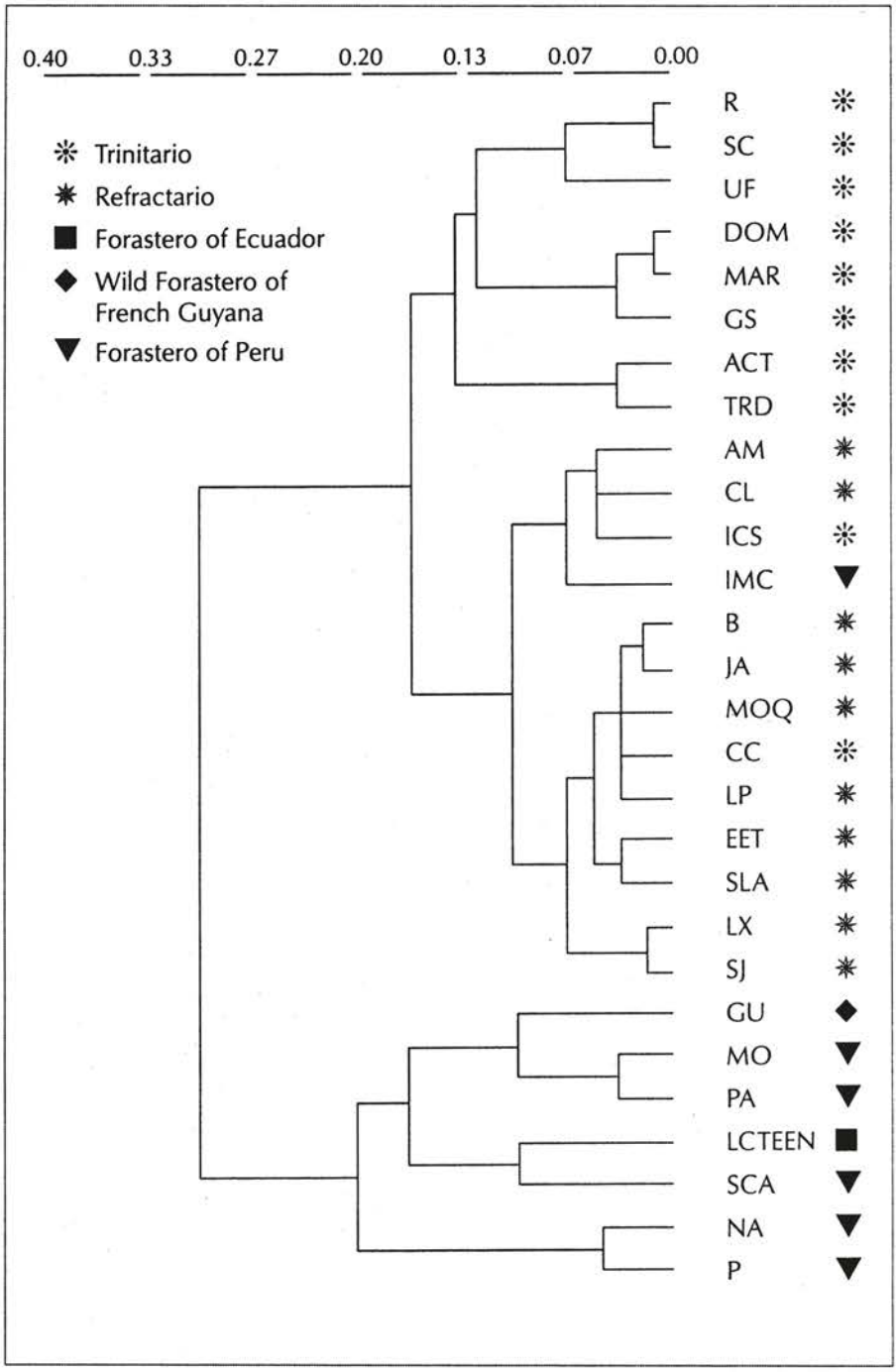

Fig. 2. Tree constructed by the neighbour joining method from non-biased Nei distances of various populations analysed with isozymes (Sounigo et al., 1996)

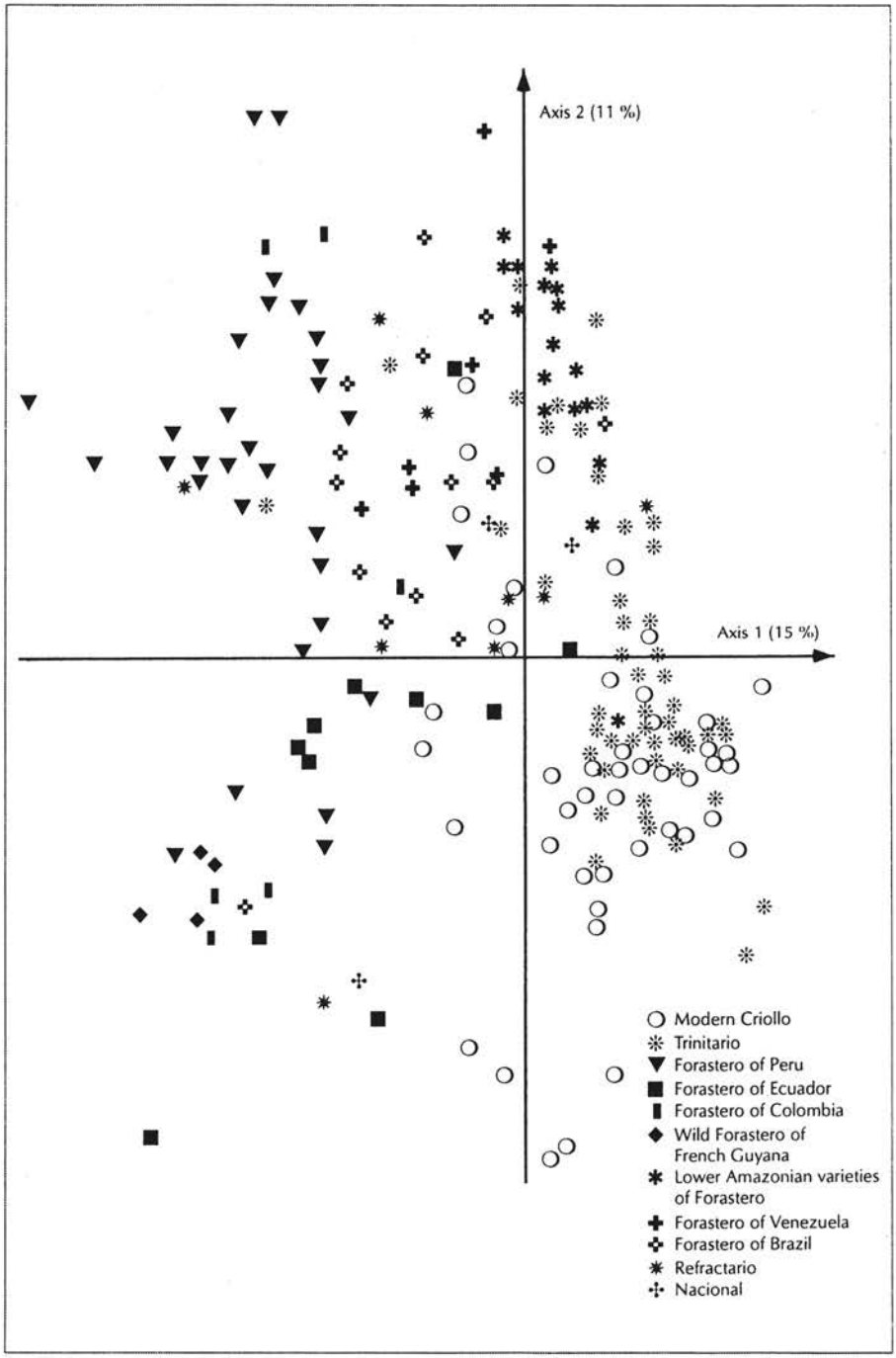

Fig. 3. Primary CA plane of 201 individuals studied for their nuclear diversity using 31 cDNA probes (Laurent et al., 1994).

Criollo and Trinitario on the other. The upper-Amazonian populations of Forastero and the modern varieties of Criollo have high variability. The cultivated varieties of Criollo are superposed on the pool of Trinitario hybrids.

In this analysis, a marked differentiation appears between the wild Forastero of Guyana on the one hand and the lower-Amazonian or Venezuelan Forastero on the other. The Refractario are found on the same area of distribution as the Forastero. The two clones of Nacional are close to certain Refractario.

A second RFLP study was focused on the diversity of the forms of Criollo cultivated in Central America and Venezuela (Motamayor et al., 2001). It covered 208 genotypes, which came from collections from the oldest plantations of Venezuela, without taking into account agronomic criteria of production, from the Lacandona forest of Mexico, close to Mayan archaeological sites where subspontaneous cacao are found that are probably the descendants of cacao cultivated in ancient times by the Mayas, and from the Yucatan. These collections comprise representations of pure varieties of Criollo cultivated in ancient times, which present varied forms of pod: oval and smooth, like those of Porcelana, or highly verrucose, like those of Pentagona. The analysis of this material was complemented by that of Trinitario, Forastero, and modern varieties of Criollo in the collections in Venezuela, Mexico, and Costa Rica.

About 50 clones and 30 probes common to the study of Laurent et al. (1994) were then used to compare the results of the two types of study. The allelic frequencies of 26 mapped loci and the mean heterozygosity per locus for each group show a high level of polymorphism among the Forastero of the upper Amazon and the Trinitario. The largest number of alleles per locus (2.2) is observed in the groups of Trinitario and upper-Amazonian Forastero, and the highest Nei genetic diversity value (0.41) is found in the Trinitario group. Similar values of mean genetic diversity (0.37) were found among the upper-Amazonian Forastero studied and the group of clones of modern Criollo. This diversity of the modern Criollo has also been indicated by other authors (Engels, 1986; Figueira et al., 1994; Laurent et al., 1994; Lerceteau et al., 1997).

On the other hand, the Criollo corresponding to the ancient cultivated types are poorly polymorphic and show the lowest values of genetic diversity (0.02). Similarly, they have a low level of mean heterozygosity per locus (0.02), contrary to the modern varieties of Criollo (0.43), which, with the Trinitario (0.40), present the highest values. Moreover, within these ancient varieties, almost no genotypic difference was found between morphological types that are quite different, such as Porcelana, Pentagona, and Guasare of Venezuela or the Criollo of the Lacandona forest in Mexico.

A CA from 64 polymorphic RFLP bands revealed a clear differentiation between the Forastero and the old varieties of Criollo (Fig. 4). The Trinitario are superposed on the modern varieties of Criollo as in the preceding analyses.

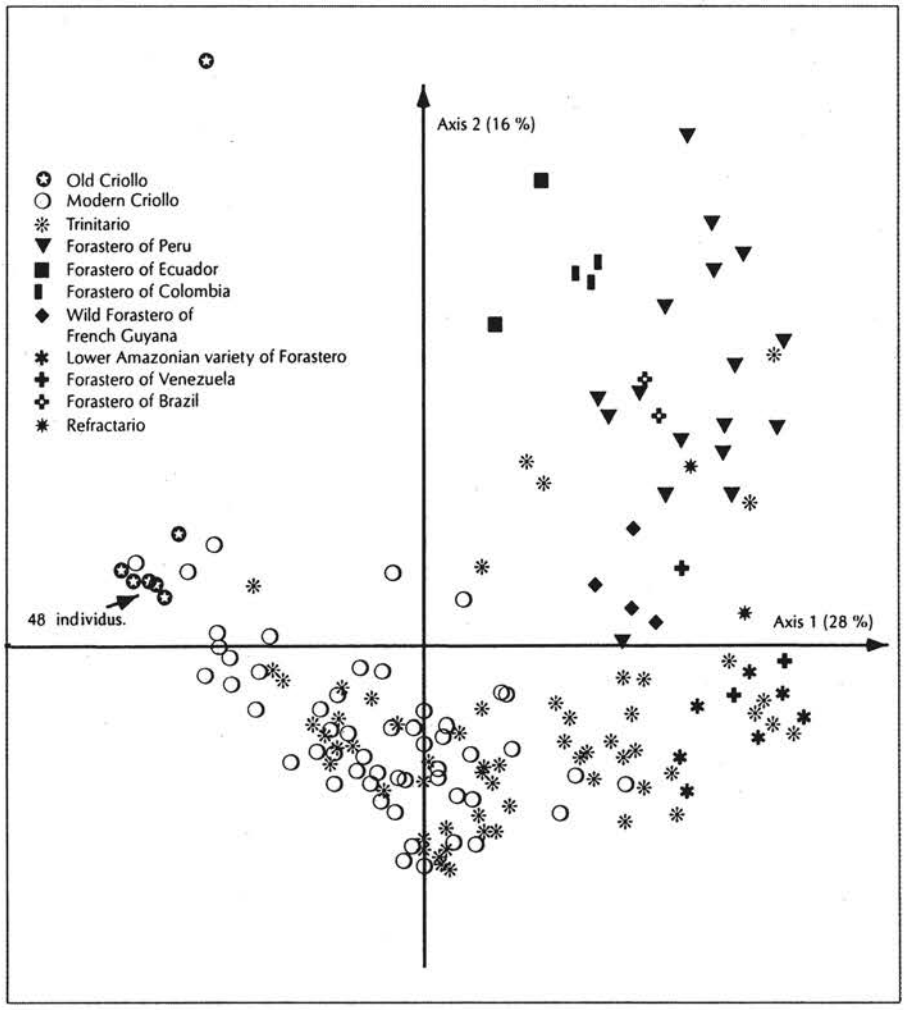

Fig. 4. Primary CA plane of 208 individuals comprising the old and new varieties of Criollo, studied for their nuclear diversity using 26 nuclear RFLP probes (Motamayor et al., 1997).

The modern Criollo clones in the collections appear generally more heterozygous than the ancient varieties, which can be explained by a selection of clones based not only on characters of bean quality, but also on certain agronomic characters of vigour, production, or disease resistance. The more vigorous types could correspond to the more or less hybridized forms resulting from introgressions of Forastero genes in the pure Criollo types cultivated in ancient times. The diversity observed among the modern varieties of Criollo could be explained by introgressions of Forastero genes—more or less significant and involving different portions of the genome—and by the diversity of Forastero at the origin of these introgressions.

In a study by Laurent et al. (1994), the diversity revealed on the first CA plane showed a differentiation between Forastero, on the one hand, and Criollo and Trinitario on the other hand, along axis 1. However, the variability appeared identical for these two groups along axis 2, which could suggest that the variability of Forastero for the characters of axis 2 may be the source of the diversity found within the modern Criollo and Trinitario. It appears that this second hypothesis can be ruled out. Indeed, in a study by Motamayor et al. (1997), for which supplementary probes were used, the diversity revealed by the first CA plane was structured differently. The diversity of Forastero could not explain that of the modern Criollo and the Trinitario, and the majority of clones of these two latter groups range from the ancient Criollo to the lower-Amazonian Forastero varieties. Moreover, the highly homozygous nature of the ancient varieties of Criollo, indicated by RFLP and microsatellites over the entire genome (Lanaud et al., 1995; Risterucci et al., 2000), allows us to make use of a reference genotype for all the chromosomes characterizing the pure Criollo. The observation of about 137 Trinitario, 10 modern varieties of Criollo, and 8 Forastero from Lower Amazonia using 16 highly polymorphic microsatellite markers seems to indicate that the same genotype or the same homogeneous population of lower-Amazonian Forastero is the source of the majority of Trinitario. For all the loci, it is always the same Forastero allele, different from the allele present in the ancient Criollo, that is found in most of the Trinitario, while the microsatellites reveal great allelic diversity in the Forastero group (Motamayor, 2001). Thus, according to these results, the different combinations of parts of the Forastero genome introgressed in the pure Criollo seem to be the source of the variability observed within Trinitario and the modern varieties of Criollo. The results of Laurent et al. (1994), which show an identical variability between the groups revealed by axis 2 of the CA, are undoubtedly linked to the low polymorphism revealed by the RFLP markers, which for the most part have only two possible alleles, common to all the groups. The use of a larger number of RFLP markers and more polymorphic markers, such as microsatellites that have revealed numerous alleles, enables us to identify the probable genetic origins of the pool of modern Criollo and Trinitario. These results thus indicate a very similar genetic structure between what is called the Criollo, which includes a majority of modern hybrid Criollo, and the Trinitario.

Lerceteau et al. (1997) analysed the genetic diversity of 59 Nacional clones from Ecuador, as well as 29 Forastero, 29 Trinitario, and 9 modern varieties of Criollo. Forty-three genomic probes, coded in terms of locus, were used. The intragroup genetic diversity was identical for Forastero (0.33), Trinitario (0.31), and the modern Criollo (0.31). It was lowest in Nacional (0.19). The heterozygosity was highest in Trinitario and modern Criollo, and lowest for Nacional. Among the latter, certain genotypes sampled in the very old plantations of Ecuador seem almost totally homozygous.

Hierarchical clustering was done using modified Rogers distances, calculated either between morphogeographic groups or between the geographic origins of clones alone. The discrimination between groups was best when the clones were classified as a function of their geographic origin (G_{ST} = 0.23 against 0.16). Considering the morphogeographic groups, the Criollo and Trinitario are the closest groups, while the Nacional are the most distant from all the other groups. These results confirm the genetic specificity of Nacional and their marked differentiation in relation to the Criollo.

The percentage of heterozygous loci of clones was estimated on the basis of 31 RFLP probes. Among the Nacional clones, two groups of individuals can be identified, BCH and SA, which have a very high rate of homozygosity varying from 90% to 100% and contain the same alleles. These trees come from two plantations that date from about 100 years ago, located about 450 km apart in the northern and southern parts of Ecuador. The present Nacional variety thus constitutes a highly hybrid population and, just like the genesis of modern Criollo and Trinitario, the BCH and SA clones, which are highly homozygous, could represent the ancestors of this population.

The diversity of cytoplasmic DNA

Yeoh et al. (1990), in a study on the structure of the chloroplast genome of cacao, pointed out the small size of this genome, which is of the order of 100×10^3 bases. The cytoplasmic diversity of 177 genotypes was analysed at the mitochondrial and chloroplast level by Laurent et al. (1993b) using mitochondrial heterologous probes (ATP-synthetase of sunflower, cytochrome oxidase of wheat) and a chloroplast heterologous probe (Rubisco of spinach).

Two chloroplast profiles were observed, A and B. Seventy per cent of clones have the profile A and belong to all the morphogeographic groups. Clones having a profile B are mostly modern Criollo, Trinitario, and some Forastero of the lower Amazon, Colombia, and Peru (Scavina).

The mitochondrial probes enabled the detection of 44 mitochondrial profiles, of which 35 contain 5 clones or fewer each. These highly variable and minor types are essentially made up of Criollo and Trinitario. Among the 9 remaining types, type 1 includes two thirds of Forastero types, including genotypes of Guyana, Venezuela, Brazil, Peru, Colombia, and Ecuador. The other major type, type 2, combines a majority of Criollo and Trinitario (26 clones) and some lower-Amazonian genotypes. By means of a CA, the diversity of mitochondrial DNA of clones studied can be visualized (Fig. 5). It is interesting to observe that, unlike with the nuclear DNA, the diversity of mitochondrial DNA is much greater in modern varieties of Criollo than in Forastero, which seem poorly variable. Recently, this mitochondrial polymorphism was indicated also among the pool of old Criollo of Mexico and Venezuela, which are also nearly totally homozygous at the nuclear level.

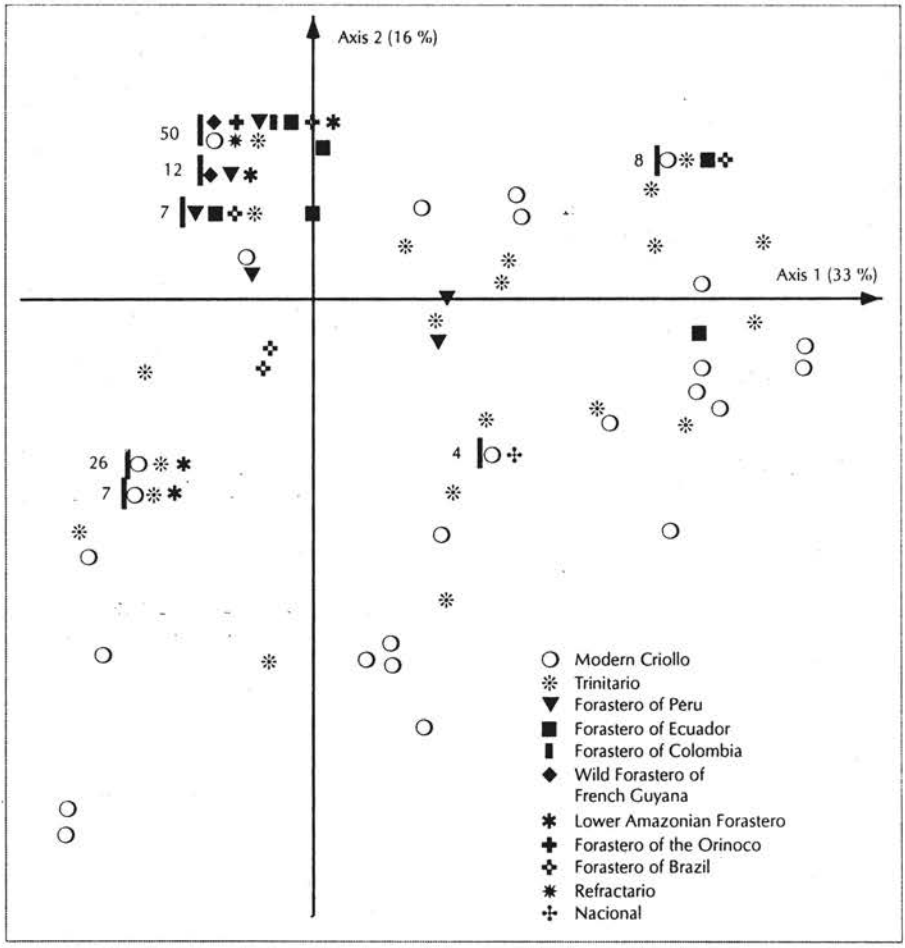

Fig. 5. Diversity of mitochondrial DNA of 177 individuals belonging to different morphogeographic groups and represented on the first plane of a CA (Laurent et al., 1993b).

Molecular diversity revealed by RAPD

Easy access to RAPD markers has allowed several researchers to use them to identify or study cacao diversity (Wilde et al., 1992; Russel et al., 1993; Figueira et al., 1994; N'goran et al., 1994; Ronning and Schnell, 1994; de la Cruz et al., 1995; Ronning et al., 1995; Sounigo et al., 1996; Lerceteau et al., 1997; Whitkus et al., 1998).

Figueira et al. (1994) analysed the diversity of genotypes belonging to three groups—Criollo, Trinitario, and Forastero—using 128 RAPD markers corresponding to amplified fragments taken from 23 primers. A continuous variation was observed between these groups, and they could not be differentiated from each other. On the other hand, a fairly clear structuration

was observed between wild and cultivated clones. That structuration is confirmed by analysis of ribosomal DNA. Considering these results, the authors proposed a new classification of cacao based not only on the three traditional groups, Criollo, Trinitario, and Forastero, but on the groups of wild cacao and cultivated cacao.

At the same time, Russel et al. (1993) studied the diversity of 25 clones resulting from three Forastero populations of Peru and Ecuador using 9 RAPD primers generating 75 bands. Despite the reduced sample, the Shannon indexes of diversity reveal greater diversity within populations than between populations. Multivariant analyses, CA and HC, reveal a genetic differentiation among the three populations of upper-Amazonian Forastero in relation with their geographic origin.

N'goran et al. (1994) analysed the genetic diversity of 106 genotypes that belong to different morphogeographic groups including the modern varieties of Criollo, using 19 primers. After thoroughly screening the bands, 49 polymorphic and reproducible bands were retained for the analyses. Out of 36 hybridized bands of DNA, 12 corresponded to highly repeated and dispersed sequences, 12 to unique sequences, and 12 to sequences that were repeated a few times. The HC, established from factorial coordinates of a CA (Fig. 6), shows a clear structuration on the first plane between Forastero and modern Criollo, as well as a clear differentiation between upper-Amazonian Forastero and lower-Amazonian Forastero (N'goran et al., 1994). If only those bands corresponding to unique sequences are considered, no clear diversity structuration appears. On the other hand, a structuration appears between modern Criollo and Forastero when only those bands corresponding to highly repeated sequences are considered, while they do not differentiate the lower-Amazonian Forastero from the upper-Amazonian Forastero.

Lerceteau et al. (1997) studied the diversity of 155 clones belonging to the groups Nacional, Forastero, Trinitario, and modern Criollo, using 40 RAPD bands resulting from 18 primers. The diversity was structured using a principal components analysis (PCA). Axis 2 indicates essentially the specificity of wild Forastero of Guyana, while the 1-3 plane reveals the structuration of other populations. Despite the continuous variations from one group to another due to intense genetic exchange, a clear differentiation is observed between Nacional and Criollo. The indication of this structuration as well as the information supplied by the RFLP on these same clones suggest that the highly homozygous Nacional clones such as BCH and SA, from very old plantations, could represent part of the original pool of Nacional. These clones were subsequently widely hybridized with other introduced clones. The high homozygosity could be a common character of pure varieties of Nacional.

Sounigo et al. (1996) used RAPD to analyse the diversity of 149 clones belonging to the groups Forastero, Trinitario, and Refractario. The calculation of the Shannon index of diversity shows that the most variable populations

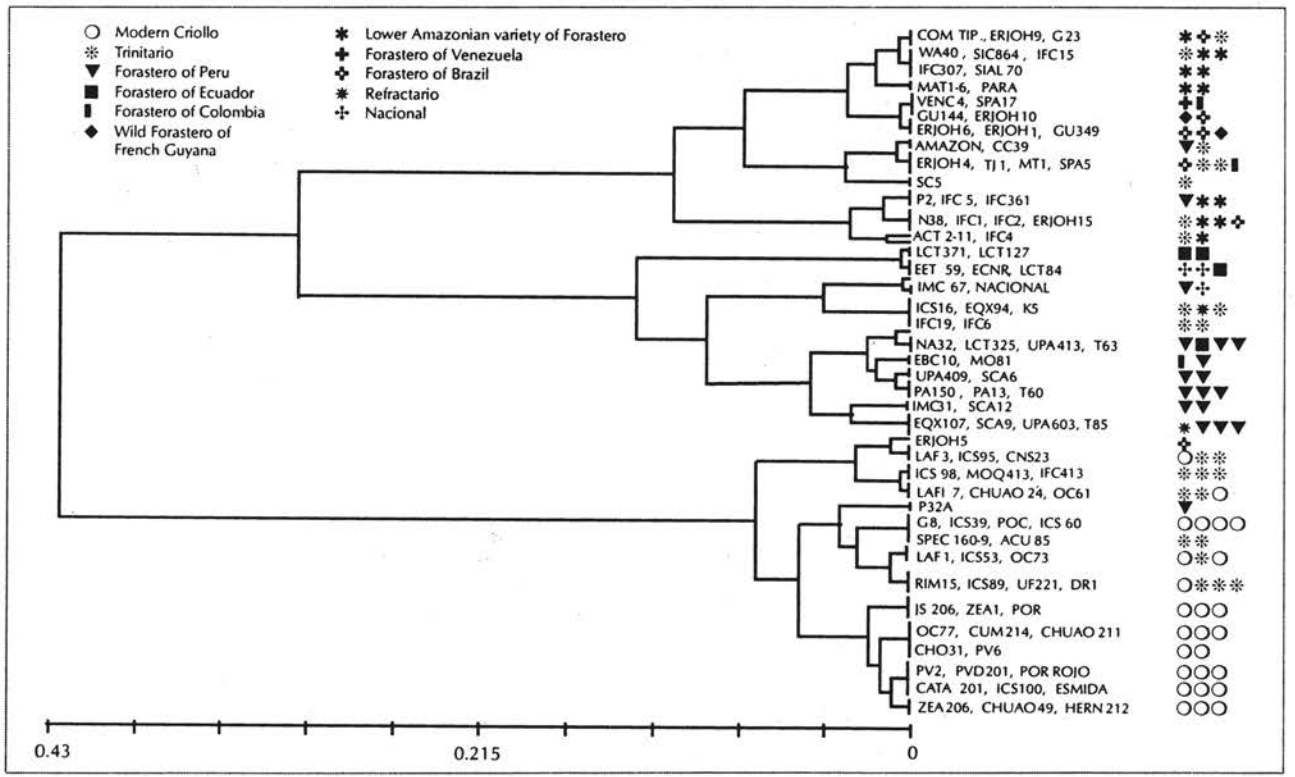

Fig. 6. Hierarchical clustering constructed by the neighbour joining method from coordinates of individuals of the seven first axes of the CA (N'goran et al., 1994). The diversity of individuals has been evaluated using 49 polymorphic RFLP bands.

are LCTEEN (Ecuador Forastero, 0.85), IMC (Peru Forastero, 0.72), and CL (Refractario, 0.71). The least variable are the populations B, JA (Refractario, 0.29 and 0.47), NA (Peru Forastero, 0.42), and GU (Guyana Forastero, 0.50). An additive tree constructed from Rogers distances shows a structuration of the populations into three groups. The first includes the three populations of Peru, NA, IMC, and SCA. The second combines the populations of Forastero, PA and MO of Peru, LCTEEN of Ecuador, and GU of Guyana. The third group contains the populations of Refractario and Trinitario.

In this RAPD analysis, the populations of Forastero are grouped differently from what is observed with the enzymatic markers (Sounigo et al., 1996). In particular, the IMC population is close to other populations of Forastero, while it is associated with Trinitario in enzymatic analysis. The SCA population appears closer to the IMC and the NA of Peru, while with isozymes it is closer to the population of the Ecuador LCTEEN. The population of wild Guyana cacao seems most distant from all the other populations (Fig. 7), while with the isozymes it is quite close to the MO and NA populations.

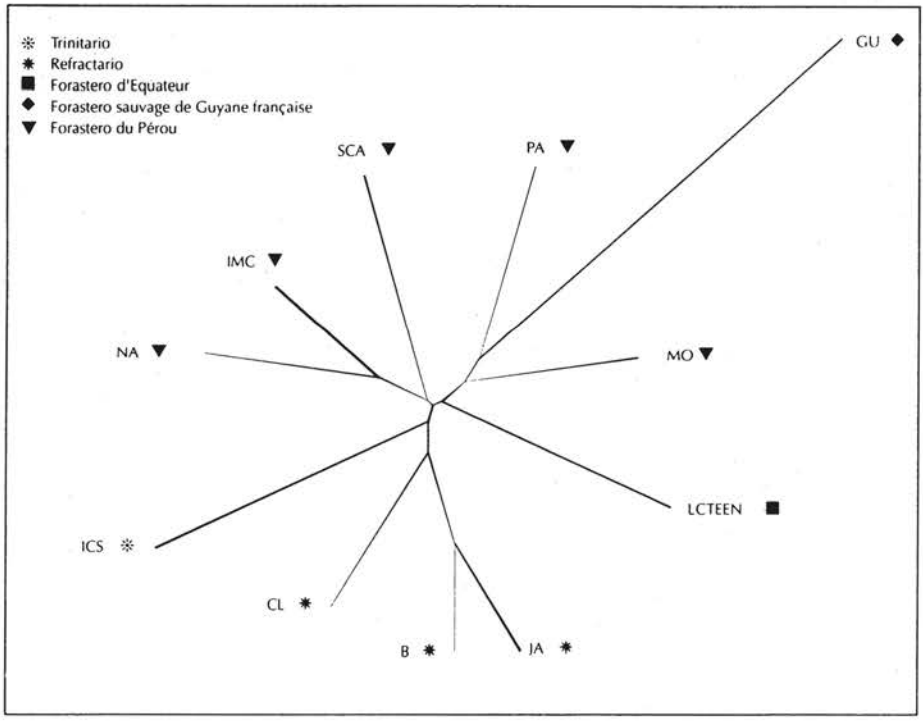

Fig. 7. Tree constructed by the neighbour joining method from Rogers distances obtained on RAPD data of Sounigo et al. (1996).

De la Cruz et al. (1995) used RAPD to analyse 42 genotypes corresponding to wild plants from either Mexico (forested islands of the Yucatan and Chiapas) or the upper Amazon, to cultivars of Criollo, Forastero, and Trinitario, and to a representative of a related genus, *Herrania*. From the calculations of dissimilarity indexes, a phylogenic tree was constructed according to the method of closest neighbours. The tree shows a greater similarity between the cultivated Criollo and the wild cacao of South America than between the cultivated Criollo and the wild plants of Mexico. The authors suggest that the wild trees found in the ancient sacred forests of the Mayans do not exist in the present collections and that they could be the closest representatives of the ancient Mayan cultivars.

In a later study, Whitkus et al. (1998) used RAPD to analyse a sample of 86 individuals including, in particular, 26 wild genotypes collected in the Lacandona forest in the state of Chiapas and 5 individuals from the forested islands of Yucatan, in Mexico, as well as the wild and cultivated clones of Central America. Unlike the populations and cultivars of Central America, where the intrapopulation diversity is always higher than the interpopulation diversity, the diversity found between the two Mexican populations is higher than that within each region. This situation reflects the low level of polymorphism found in the two populations. A structuration of the diversity similar to that of de la Cruz et al. (1995) has been observed. The two populations of Mexico seem well differentiated and original. These results conform to the hypothesis of Cuatrecasas (1964) about the natural distribution of cacao in Central America. The authors emphasize nevertheless the lack of affinity between the cultivated Criollo and the wild plants observed in Mexico. According to them, the cacao collected in the Lacandona forest could represent the wild cacao. The plants sampled in the island forests cultivated in ancient times by the Mayans could on the other hand have been introduced from wild populations. In this case, they represent a subsample of populations present in the Lacandona forest, which would explain the difference between the two Mexican populations. They also would correspond to the material closest to the cacao cultivated by the Mayans.

Our own analyses, done simultaneously on the population of the Lacandona forest and on that of the Yucatan, indicate a perfect genetic similarity between these two populations, unlike the analyses of Whitkus et al. (1998). This contradiction perhaps originates in the fact that the analyses of Whitkus et al. (1998) possibly included young wild plants, morphologically very similar to those of *T. cacao*, but not belonging to that species. These plants, observed in the Yucatan, could explain the genetic distances obtained by Whitkus et al. (1998) between cacao from the two places.

CONCLUSION

The intense genetic exchange that has occurred during the past three centuries between the wild and cultivated populations of cacao often makes them difficult to classify.

Morphological descriptors can be used to differentiate the populations of different geographic origin and to structure their diversity. However, they sometimes give a biased picture of the diversity of the genetic resources. For example, our analyses suggested that the ancient varieties of Criollo, highly varied as to their morphology, ranging from the Porcelana type with smooth pods to the Pentagona type with particularly verrucose pods, are genetically highly homogeneous and highly homozygous. Human selection could in this case have contributed to the fixation and conservation of very different morphological types, resulting in occasional mutations.

Genetic markers, which reflect the structure of the entire genome, are the tools of choice to reveal the diversity and relationships between wild and cultivated cacao. Different types of markers, nuclear (RAPD, RFLP from genomic or cDNA probes, microsatellites) or cytoplasmic, contribute varying and complementary types of information, indicating the variation of sequences that evolve at different rates. Codominant markers such as RFLP and isozymes provide more precise information on the genetic structure of populations and individuals.

From identification of alleles present in the genotypes as well as the gradients observed in the levels of heterozygosity of clones within hybrid cultivated populations such as Nacional of Ecuador or Criollo and Trinitario, we can understand the genesis of these populations and discover their most probable ancestors (Lerceteau et al., 1997; Motamayor et al., 1997). In these two cases, the ancestors probably constitute a very restricted genetic base.

A wide genetic diversity has been indicated between the wild populations of Forastero and within some of these populations originating in the upper Amazon, particularly Ecuador. A continuous variation can be observed among them. However, in several analyses, the Forastero of Guyana seem differentiated from other types of Forastero, including those of the lower Amazon or of Orinoco, which they are closest to geographically. A differentiation between Forastero and Criollo has been observed with all the markers, with a more marked differentiation between Forastero and ancient Criollo revealed by nuclear RFLP. Analysis of Nacional has also shown their originality with respect to Forastero and Criollo simultaneously. Nacional is clearly distinct from these two groups.

Thus, the diversity of the species could be structured diagrammatically around four unequal groups: ancient Criollo, pure varieties of Nacional, wild Forastero of Guyana, and a wider pool comprising the Amazonian Forastero (of the upper and lower Amazon) and those of the Orinoco, even though this pool comprises populations that are differentiated from one another according

to their geographic origin. This structuration differs from the usual classification of cacao into three groups: Forastero, Criollo, and Trinitario. On the one hand, the genetic structure of Trinitario is similar to that of the Criollo group, which in this classification includes a majority of modern Criollo. On the other hand, within the group called Forastero are included cacao that are well-differentiated, such as the wild populations of Guyana or those of the original variety Nacional.

The classification that we propose can be supported by the paleoclimatic history of the South American continent during the Quaternary era, during which considerable ecological upheavals occurred. The succession of periods of glaciation interrupted by phases of heating of the climate resulted in an alternation of dry and humid periods in the tropical regions. During the dry periods, the tropical forest shrank so as to be limited to the forest islands, favouring the differentiation of some populations and the disappearance of others. During the humid phases, the forests extended anew from these reserve zones. Such reserve zones were identified in Guyana, Brazil, Colombia, and Venezuela, as well as the upper Amazon, where a vast reserve zone is indicated. This phenomenon, recognized for numerous animal and plant species (Simpson and Haffer, 1978), could explain the diversity and differentiation of certain populations of cacao. Through molecular study on the sequence of certain genes, the analysis of relationships and differentiation between populations could be refined, and the time of their divergence could be established. It would thus be possible to situate the differentiation of populations in relation to major ecological and geographic events.

The importance of the sampling to the conclusions drawn from various analyses must be emphasized. Genetic resources have often been collected with criteria of agronomic interest, which could give a biased picture of the diversity actually existing in a given region: this is the case of Forastero collected in Peru with a specific objective, resistance to witches' broom disease. For the Criollo, the primary studies pertained only to the clones present in the collections and collected for their agronomic characters. The great diversity revealed within this group reflected in fact the diversity of introgressions from a unique ancestral type identified subsequently using markers and studies conducted directly on the oldest plantations or close to remains of Mayan civilizations, which were the first to have domesticated cacao. From a wider sample and better representation of certain populations the structuration between different origins of cacao that has appeared in these analyses could be corroborated.

These analyses confirmed that only a small part of the genetic diversity is exploited in selection, as several authors have already mentioned (Bartley, 1979; Lockwood and End, 1992). The hybrid forms between Criollo and Forastero, which correspond most often to traditional forms of Trinitario, implied a very narrow genetic base in each of the two groups of origin. Other types of hybrids between Criollo and Forastero, exploiting all the genetic

richness of Forastero, could thus be used in selection. Moreover, only a few meioses separate the ancient Criollo from the formation of hybrid types, and the linkage disequilibrium between characters and markers could have been maintained. This situation would facilitate the exploitation of genetic resources of the Criollo and Trinitario type, using markers that frame the useful genes as markers of early selection. Similarly, for the Forastero, only a very small part of the genetic diversity has actually been included in the selection programmes, and a small number of Peru Forastero, often having strong parental links, have been widely diffused and integrated in selection programmes of all the producer countries.

Even though the existing diversity in the collections is far from having been fully exploited, the extent of diversity revealed on a sometimes small sample of certain populations testifies in favour of the pursuit of collection expeditions in regions rich in diversity to preserve the genetic resources of cacao.

APPENDIX

Plant Material

The enzymatic analyses of Lanaud (1987) were done on 12 modern Criollo, 17 Trinitario selected from America, 22 Trinitario selected from Africa, 64 lower-Amazonian Forastero selected from Africa, 19 Forastero collected by Lanaud (1986) along the Orinoco in Venezuela (VENC), 92 wild Forastero collected by Sallee (1987) in French Guyana, 19 cultivated cacao collected by Clément (1986) in French Guyana, 74 Forastero collected by Pound (1938, 1943) in Peru (39 individuals of GO, 8 IMC, 11 P, 7 NA, 8 PA, and 1 SCA), 10 Refractario collected in Ecuador by Chalmers (1968) (6 individuals of EQX) and Pound (1938) (4 individuals of MOQ), 7 Forastero collected during the Colombian expedition of Ocampo (1985) in Colombia (EBC), and 18 Forastero collected by J.B. Allen between 1980 and 1985 in Ecuador (LCTEEN).

The enzymatic analyses of Sounigo et al. (1996) were done on 482 clones of cacao belonging to 28 populations or groups of accessions present in the Trinidad collection: 6 populations of Forastero collected in Peru (IMC, MO, PA, NA, P, SCA: 112 individuals), 1 population of Forastero from Ecuador (LCTEEN: 16 individuals), 1 population of Forastero collected in French Guyana (16 individuals), 9 populations corresponding to Refractario (AM, B, CL, JA, LP, LX, MOQ, SJ, SLA: 245 individuals), and 10 populations of Trinitario (ACT, CC, DOM, ICS, GS, MAR, R, SC, TRD, UF: 114 individuals).

The RFLP analyses (Laurent et al., 1993b, 1994) were done on 201 genotypes belonging to different morphogeographic groups: 45 modern Criollo (ICS, CATA, BOC, CHUAO, OC, POR, JS, MT, PV, ZEA, LAF), 20 Forastero of Peru (P, PA, NA, MO, IMC, SCA), 12 Forastero selected from Africa from clones of Peru (T, UPA), 11 Forastero of Ecuador (LCTEEN), 6 Forastero of Colombia (EBC, SPA), 7 Forastero of the Orinoco valley in Venezuela (VENC), 4 wild Forastero of French Guyana (GU), 22 Forastero of Brazil (ERJOH, Comun, Para, SIAL, SIC, MAT, Catongo), 9 lower-Amazonian Forastero selected from Africa (Amelonado) (IFC, SF), 37 Trinitario selected from Central and South America and from the Caribbean (ACT, ICS, GS, CHUAO, CNS, WA, RIM, MT, TJ, SC, SGU, CC, UF), 4 Trinitario selected from Asia (WA, LAFI, DR, G), 12 Trinitario selected from Africa (SNK, IFC, N, W, ACU, K), 10 Refractario (EQX, MOQ), and 2 Nacional.

Enzymatic Analyses

The methods used as well as the genetic determination and protocols of extraction and indication are given in Lanaud (1986a). In the analyses of Lanaud (1987), 6 polymorphic enzymatic systems (PGI, PGM, ADH, MDH, PAC, ICD), corresponding to 9 loci, were used. The analyses of Sounigo et al. (1996) used 5 polymorphic enzymatic systems (PGI, ADH, MDH, PAC, ICD), corresponding to 6 loci.

RFLP Analyses

A genomic library bank and a cDNA library were constructed. Between 31 and 50 of these probes, for the most part mapped (Lanaud et al., 1995), were used to study the nuclear diversity. The heterozygosity level of clones was found using mapped markers with known genetic determinism. Moreover, heterologous probes were employed to study the cytoplasmic diversity (mitochondrial and chloroplast). The protocols are described in Laurent et al. (1993a, b, 1994).

RAPD Analyses

After a thorough selection of primers from Operon kits that generate reproducible bands, 19 primers producing 49 polymorphic bands were used to analyse the diversity (N'goran et al., 1994). The nature of the RAPD bands observed—unique or repeated sequence—was tested by Southern hybridization on the total DNA of cacao restricted by restriction enzymes.

Data Analysis

Several genetic parameters of diversity were calculated on the enzymatic and molecular data: percentage of polymorphic loci, percentage of heterozygosity, mean number of alleles per locus, Shannon index of diversity based on genotype frequencies, and Nei diversity index (Nei, 1978) based on allelic frequencies.

From the Nei and Shannon indexes, the total diversity was broken down into intrapopulation and interpopulation diversity according to the following formulas:

$$H_{intrapop} = H_{pop}/H_{total}$$

$$H_{interpop} = (H_{total} + H_{pop})/H_{total}$$

For all the data, multivariate analyses were generally used—correspondence analysis (CA) (Benzecri, 1973) or principal components analyses (PCA) for quantitative characters.

For the enzymatic data, the genetic distances of Nei on the allelic frequencies were used to measure distances between populations. A hierarchical clustering (HC) was constructed according to the UPGMA method.

For the RAPD data (N'goran et al., 1994), a HC was constructed from the coordinates of clones on the seven primary axes of the CA.

For the RAPD data of Sounigo et al. (1996), the Rogers and Tanimoto distances were calculated, and a tree was constructed according to the neighbour joining method.

REFERENCES

Allen, J.B. and Lass, R.A. 1983. London cocoa trade Amazon project: final report, phase 1. Cocoa Grower's Bulletin, 34: 1-71.

Amefia, Y.K. 1986. Sur le polymorphisme enzymatique des cacaoyers introduits au Togo. Doct. thesis, INA, Paris- Grignon, 146 p.

Atkinson, M.D., Withers, L.A., and Simpson, M.J.A. 1986. Characterisation of cacao germplasm using isoenzyme markers. 1. A preliminary survey of diversity using starch gel electrophoresis and standardisation of the procedure. *Euphytica*, 35: 741-750.

Barriga, J.P., Machado, P.R.F., de Almeida, C.M.V.C. and de Almeida, C.F.G. 1984. A preservação e utilização dos recursos genéticos de cacau na Amazônia brasileira. In: *IX^e Conférence Internationale sur la Recherche Cacaoyère*. Lagos, Nigeria, Cocoa Producers' Alliance, pp. 73-79.

Bartley, B.G. 1979. Global concepts for genetic resources and breeding in cacao. In: *VII^e Conférence Internationale sur la Recherche Cacaoyère*. Lagos, Nigeria, Cocoa Producers' Alliance, pp. 519-525.

Bekele, F.L. 1992. Use of botanical descriptors for cocoa characterization: CRU experiences. In: *International Workshop on Conservation, Characterisation and Utilisation of Cocoa Genetic Resources in the 21st Century*. Port of Spain, Trinidad and Tobago, CRU, pp. 77-91.

Bekele, F. and Bekele I. 1998. A sampling of the phenetic diversity of cacao in the international cocoa gene bank of Trinidad. *Crop Science*, 36: 57-64.

Bekele, F.L., Kennedy, A.J., McDavid, C., Lauckner, B., and Bekele, I. 1994. Numerical taxonomic studies on cacao (*Theobroma cacao* L.) in Trinidad. *Euphytica*, 75: 231-240.

Benzecri, J.P. 1973. L´Analyse des Données. 2. L´Analyse des Correspondances. Paris, Dunod, 616 p.

Bergman, J.F. 1969. The distribution of cacao cultivation in pre-Columbian America. *Annals of the Association of American Geographers*, 59: 85-96.

Bernoulli, G. 1869. Ubersicht der bis jetzt bekannten Arten von *Theobroma*. *Neue Denkschriften der allgemeinen schweizerischen Gesellschaft für die gesammten Naturwissenschaften*, 24: 1-15.

Bouharmont, J. 1960. Recherches cytologiques sur la frutification et l´incompatibilitié chez *Theobroma cacao* L. INEAC, Série scientifique no. 89, 117 p.

Braudeau, J. 1969. *Le Cacaoyer*. Paris, Maisonneuve et Larose, 304 p.

Burle, L. 1952. La production du cacao en Afrique occidentale française. Centre de recherches agronomiques de Bingerville, Bulletin no. 5, pp. 3-21.

Chalmers, W.S. 1968. Cacao collecting trip to the Oriente of Ecuador (November 1968). Port of Spain, Trinidad and Tobago, CRU, 4 p.

Cheesman, E.E. 1944. Notes on the nomenclature, classification and possible relationships of cacao populations. *Tropical Agriculture*, 21: 144-159.

Chevalier, A. 1946. Révision du genre *Theobroma* d´après l´herbier du Muséum national d´histoire naturelle de Paris. *Revue de Botanique Appliquée*, 26: 265-285.

Clement, D. 1986. Cacaoyers de Guyane: prospections. Café, Cacao, Thé, 30: 11-35.

Coe, S.D. and Coe., M.D. 1996. The True History of Chocolate. London, Thames and Hudson, 280 p.

Cope, F.W. 1962. The mechanism of pollen incompatibility in *Theobroma cacao*. *Heredity*, 17: 157-182.

Cuatrecasas, J. 1964. Cacao and its allies: a taxonomic revision of the genus *Theobroma*. *Contributions from the United States Herbarium*, 35: 379-614.

de la Cruz, M., Whitkus, R., Gomez-Pompa, A., and Mota-Bravo, L. 1995. Origins of cacao cultivation. *Nature*, 375: 542-543.

Engels, J.M.M. 1983a. A systematic description of cacao clones. 1. The discriminative value of quantitative characteristics. *Euphytica*, 32: 377-385.

Engels, J.M.M. 1983b. A systematic description of cacao clones. 2. The discriminative value of qualitative characteristics and the practical compatibility of the discriminative value of the quantitative and qualitative descriptors. *Euphytica*, 32: 719-733.

Engels, J.M.M. 1986. The systematic description of cacao clones and its significance for taxonomy and plant breeding. PhD thesis, Agricultural University, Wageningen, Netherlands, 125 p.

Engels, J.M.M., Bartley, B.G.D., and Enriquez, G.A. 1980. Cacao descriptors, their states and modus operandi. *Turrialba*, 30: 209-218.

Enriquez, G.A. 1985. Cursos sobre el cultivo del cacao. Turrialba, Costa Rica, CATIE, Serie Materiales de Enseñanza, no. 22, 239 p.

Enriquez, G.A. 1992. Characterisitcs of cacao Nacional of Ecuador. In: International workshop on conservation, characterisation and utilisation of cocoa genetic resources in the 21st century. Port of Spain, Trinidad and Tobago, CRU, pp. 269-278.

Enriquez, G.A. and Soria, J. 1968. The variability of certain bean characteristics of cacao (*Theobroma cacao* L.). *Euphytica*, 17: 114-120.

Figueira, A., Janick, J., and Goldsbrough, P. 1992. Genome size and DNA polymorphism in *Theobroma cacao*. *Journal of the American Society for Horticultural Science*, 117: 673-677.

Figueira, A., Janick, J., Levy, M., and Goldsbrough, P. 1994. Re-examining the classification of Theobroma cacao L. using molecular markers. *Journal of the American Society for Horticultural Science*, 119: 1073-1082.

Glendinning, D.R. 1966. The incompatibility alleles of cocoa: report for the period 1st October 1963–31st March 1965. Ghana, Cocoa Research Institute.

Guisan, 1825. De la culture du cacaoyer. In: *Traité sur les Terres Noyées de Guyane*, 2nd ed. Imprimerie du Roy, pp. 170-189.

Hamrick, J.L. and Godt, M.J.W. 1990. Allozyme variation in plant species. In: *Plant Population Genetics, Breeding, and Genetic Resources*. A.H.D. Brown et al. eds., Sunderland, USA, Sinauer, pp. 43-63.

Hamrick, J.L., Godt, M.J.W., and Sherman-Broyles, S.L. 1992. Factors influencing levels of genetic diversity in woody plant species. *New Forests*, 6: 95-124.

IBPGR. 1981. Report of the IBPGR working group on genetic resources of cocoa. Rome, IBPGR, ACP-IBPGR/80/56, 28 p.

Knight, R. and Rogers, H. 1955. Incompatibility in *Theobroma cacao*. *Heredity*, 9: 67-69.

Lachenaud, P. and Sallee, S. 1993. Les cacaoyers spontanés de Guyane: localisation, écologie et morphologie. *Café, Cacao, Thé*, 37: 101-114.

Lanaud, C. 1986a. Utilisation des marqueurs enzymatiques pour l'étude génétique du cacaoyer *Theobroma cacao* L. 1. Contrôle génétique et linkage de neuf marqueurs enzymatiques. *Café, Cacao, Thé*, 30: 259-270.

Lanaud, C. 1986b. Utilisation des marqueurs enzymatiques pour l'étude génétique du cacaoyer *Theobroma cacao* L. 2. 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, systèmes d'incompatibilité, haploïdes spontanés; leurs conséquences pour l'amélioration génétique de cette espèce. Doct. thesis, Université Paris XI, Orsay, 106 p.

Lanaud, C. and Bertaud, J. 1984. Mise en évidence de nouveaux marqueurs génétiques chez *Theobroma cacao* L. par les techniques d'électrophorèse. In: IX^{e} conférence internationale sur la recherche cacaoyère. Lagos, Nigeria, Cocoa Producers' Alliance, pp. 249-253.

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

Lanaud, C., Risterucci, A.M., N'goran, A.K.J., Clement, D., Flament, M.H., Laurent, V., and Falque, M. 1995. A genetic linkage map of *Theobroma cacao* L. *Theoretical and Applied Genetics*, 91: 987-993.

Laurent, V., Risterucci, A.M., and Lanaud, C. 1993a. Variability for nuclear ribosomal genes within *Theobroma cacao*. *Heredity*, 71: 96-103.

Laurent, V., Risterucci, A.M., and Lanaud, C. 1993b. Chloroplast and mitochondrial DNA diversity in *Theobroma cacao*. *Theoretical and Applied Genetics*, 87: 81-88.

Laurent, V., Risterucci, A.M., and Lanaud, C. 1994. Genetic diversity in cacao revealed by cDNA probes. *Theoretical and Applied Genetics*, 88: 193-198.

Lecomte, H. and Challot, C. 1897. Le Cacaoyer et Sa Culture. Paris, Carré et Naud, 121 p.

Lerceteau, E., Robert, T., Petiard, V., and Crouzillat, D. 1997. Evaluation of the extent of genetic variability among *Theobroma cacao* accessions using RAPD and RFLP markers. *Theoretical and Applied Genetics*, 95: 10-19.

Lockwood, G. and End, M. 1992. History, technique and future needs for cacao collection. In: International workshop on conservation, characterisation and utilisation of cocoa genetic resources in the 21st century. Port of Spain, Trinidad and Tobago, CRU, pp. 1-14.

Morris, D. 1882. *Cacao: How to Grow and How to Cure It*. Jamaica, pp. 1-45.

Motamayor, J.C. 2001. Etude de la diversite génétique et de la domestication des cacaoyers du groupe Criollo (*T. cacao* L.) a l'aide de marqueurs moléculaires. Doct. thesis, University of Paris XI, Orsay, 18 July 2001.

Motamayor, J.C., Risterucci, A.M., Laurent, V., Moreno, A., and Lanaud, C. 1997. The genetic diversity of Criollo cacao and the consequences for quality breeding. In: *I° Congreso Venezolano del Cacao y Su Industria*. Aragua, Venezuela.

Motamayor, J.C., Risterucci, A.M., and Lanaud, C. 2001. Cacao domestication I: genetic evidence of a South American origin of cacao cultivated by the Mayans. *Heredity*.

Myers, J.G. 1930. Notes on wild cacao in Surinam and in British Guiana. *Kew Bulletin*, 1: 1-10.

Nei, M. 1978. Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics*, 89: 583-690.

N´goran, J.A.K. 1994. Contribution à l´étude génétique du cacaoyer par les marqueurs moléculaires: diversité génétique et recherche de QTLs. Doct. thesis, Université Montpellier II, Montpellier, France, 131 p.

N´goran, J.A.K., Laurent, V., Risterucci, A.M., and Lanaud, C. 1994. Comparative genetic diversity of *Theobroma cacao* L. using RFLP and RAPD markers. *Heredity*, 73: 589-597.

Ocampo, R.F. 1985. Informe sobre recolección de germoplasma de *Theobroma cacao* L. *El Cacaotero Colombiano*, 31: 24-29.

Paradis, L. 1979. Le cacao précolombien: monnaie d´échange et breuvage des dieux. *Journal d´agriculture traditionnelle et de botanique appliquée*, 26: 3-4.

Pittier, H. 1930. A propos des cacaoyers spontanés. *Revue de botanique appliquée*, 10: 777.

Pittier, H. 1933. Degeneration of cacao through natural hybridization. *Journal of Heredity*, 36: 385-390.

Pound, F.J. 1938. Cocoa and witches´ broom disease (*Marasmius perniciosus*) of South America with notes on other species of *Theobroma*. Port of Spain, Trinidad and Tobago, Yuille´s Printery.

Pound, F.J. 1943. Cocoa and witches´ broom disease: report on a recent visit to the Amazon territory of Peru. In: *The Archives of Cocoa Research*, vol. 1. H. Toxopeus, eds., London, ACRI-IOCC.

Rabouin, L.M. and Paulin, D. 1993. Etude de l´héritabilité de descripteurs floraux. In: XI[e] Conférence Internationale sur la Recherche Cacaoyère. Lagos, Nigeria, Cocoa Producers´ Alliance, pp. 449-450.

Reyes, H. 1992. Criollo cacao germplasm in Venezuela. In: International workshop on conservation, characterisation and utilisation of cocoa genetic resources in the 21st century. Port of Spain, Trinidad and Tobago, CRU.

Risterucci, A.M., Lanaud, C., N´goran, J.A.K., and Pieretti, I. 2000. A saturated linkage map of *Theobroma cacao*. In: XII[e] Conférence Internationale sur la Recherche Cacaoyère. Lagos, Nigeria, Cocoa Producers´ Alliance.

Ronning, C.M. and Schnell, R.J. 1994. Allozyme diversity in a germplasm collection of *Theobroma cacao* L. *Journal of Heredity*, 85: 291-295.

Ronning, C.M., Schnell, R.J., and Kuhn, D.N. 1995. Inheritance of random amplified polymorphic DNA (RAPD) markers in *Theobroma cacao* L. *Journal of the American Society for Horticultural Science*, 120: 681-686.

Russel, J.R., Hosein, F., Waugh, R., and Powell, W. 1993. Genetic differentiation of cocoa (*Theobroma cacao* L.) populations revealed by RAPD analysis. *Molecular Ecology*, 2: 89-97.

Sallee, B. 1987. Compte rendu de prospection dans les populations naturelles de *Theobroma cacao* L. du Haut Camopi (Guyane française). Kourou, Guyana, CIRAD, 45 p.

Schumann, K. 1886. Fam. Sterculiaceae g. *Theobroma*. In: *Flora Brasiliensis*, vol. XII, part III, pp. 70-77.

Simpson, B.B. and Haffer, J. 1978. Speciation patterns in the Amazonian forest biota. *Annual Review of Ecology and Systematics*, 9: 497-518.

Soria, J. 1970a. Principal varieties of cocoa cultivated in tropical America. *Cocoa Growers´ Bulletin*, 19: 12-21.

Soria, J. 1970b. The present status and perspectives for cocoa cultivars in Latin America. Proceedings of the Tropical Region American Society for Horticultural Science, 14.

Soria, J. 1973. Survey of crop genetics resources in their centers of diversity: first report. Rome, FAO, pp. 119-125.

Soria, J. and Enriquez, G.A. 1981. International cacao cultivar catalogue. Turrialba, Costa Rica, CATIE, Technical Series, Technical Bulletin no. 6, 156 p.

Sounigo, O., Christopher, Y., and Umaharan, R. 1996. Genetic diversity assessment of *Theobroma cacao* L. using isoenzyme and RAPD analyses. In: Annual Report 1996, CRU. Port of Spain, Trinidad and Tobago, CRU, pp. 35-51.

Toxopeus, H. 1985. Botany, types and populations in cocoa. In: *Cocoa*, 4th ed. G.A.R. Wood and R.A. Lass, eds., London, Longman, pp. 11-37.

Wadsworth, R.M. and Harwood, T. 2000. International Cocoa Germplasm Database ICDG version 4.1. London International Financial Futures and Options Exchange and the University of Reading, UK.

Warren, J.M. 1994. Isozyme variation in a number of populations of *Theobroma cacao* L. obtained through various sampling regimes. *Euphytica*, 72: 121-126.

Whitkus, R., de la Cruz, M., and Mota-Bravo, L. 1998. Genetic diversity and relationships of cacao (*Theobroma cacao* L.) in southern Mexico. *Theoretical and Applied Genetics*, 96: 621-627.

Wilde, J., Waugh, R., and Powell, W. 1992. Genetic fingerprinting of *Theobroma* clones using randomly amplified polymorphic DNA markers. *Theoretical and Applied Genetics*, 83: 871-877.

Wood, G.A.R. 1991. A history of early cocoa introductions. *Cocoa Growers' Bulletin*, 44: 7-12.

Yeoh, H.H., Chung, D.K., and Fritz, P.J. 1990. *Theobroma cacao* chloroplast DNA: isolation, molecular cloning and characterization. *Café, Cacao, Thé*, 34: 173-178.

Yidana, J.A., Kennedy, A.J., and Withers, L.A. 1987. Variation in peroxidase isozymes of cocoa (*Theobroma cacao* L.). In: X^{e} Conférence Internationale sur la Recherche Cacaoyère. Lagos, Nigeria, Cocoa Producers' Alliance, pp. 719-723.

Young, A.M. 1994. *The Chocolate Tree: A Natural History of Cacao*. Washington, D.C., Smithsonian Institution Press, 200 p.