

Recurrent selection of cocoa populations in Côte d'Ivoire: comparative genetic diversity between the first and second cycles

N. D. POKOU¹, J. A. K. N'GORAN¹, PH. LACHENAUD², A. B. ESKES³, J. C. MONTAMAYOR⁴, R. SCHNELL⁴, M. KOLESNIKOVA-ALLEN⁵, D. CLÉMENT⁶ and A. SANGARÉ¹

¹Centre National de Recherche Agronomique (CNRA), 01 B.P. 1740 Abidjan 01, Côte d'Ivoire, E-mail: pokoudesire@yahoo.fr; ²Centre de Coopération Internationale en Recherche Agronomique pour le Développement, CIRAD-Bios, UPR "Bioagresseurs de pérennes", TA-A31/02, 34398 Montpellier cedex 5, France; ³Bioversity International/CIRAD, c/o INIBAP, Parc Scientifique Agropolis II, 34397 Montpellier-Cedex 5, France; ⁴United State Department of Agriculture/Agricultural Research Service (USDA-ARS), 13601 Old Cutler Road, Miami, Florida; ⁵International Institute of Tropical Agriculture (IITA) PMB 5320, Ibadan, Nigeria; ⁶CEPLAC/CEPEC/SEGEN, Rod. Ilhéus-Itabuna, km 22. Itabuna, BA, 45.600-000, Brazil

With 5 figures and 5 tables

Received May 14, 2008/Accepted August 23, 2008

Communicated by S. Korban

Abstract

In Côte d'Ivoire, the cocoa breeding programme has been based on the creation of hybrids between different genetic groups. From 1990 onward, a reciprocal recurrent selection programme has been set up with the purpose of improving simultaneously the characteristics of the two main genetic groups: Upper Amazon Forastero (UA) and a mixture of Lower Amazon Forastero (LA) and Trinitario (T). Based on data obtained from 12 microsatellite primers, the genetic diversity and genetic distances of the parental populations used in the first and second selection cycles are presented. The results revealed that the diversity of populations UA0 and UA1 on the one hand and (LA+T)0 and (LA+T)1 on the other is similar. The genetic distances were small between the parental populations used for the first and second cycles. Genetic diversity was greater in the UA group than in the LA + T group. The number of rare and of private alleles was reduced for both genetic groups, as well as the number of the frequent alleles in the LA + T group.

Key words: cocoa — genetic diversity — Lower Amazon — Upper Amazon — selection — Trinitario

The cocoa tree (*Theobroma cacao* L., Malvaceae) is native to Amazonian forests (Bartley 2005). The trees are sub-divided into three large morphological groups: Criollo, Forastero and Trinitario (Cheesman 1944). The Criollo group, the first to be domesticated, produces fine-flavour chocolate but is low yielding and very susceptible to pest and diseases. Within the Forastero group, a distinction is made between Lower Amazon Forastero (LA), which is the first to be cultivated in that group, and Upper Amazon Forastero (UA). The LA group includes the Amelonado type, which is the most widely cultivated type in this group and which shows often low genetic variability. Amelonado is fairly vigorous, more resistant to diseases than Criollo, and quite productive, although late bearing and less well adapted to growing areas without shade than the UA group. The UA group is very diverse and has been collected from the Upper Amazon region from the end of the 1930s (Pound 1938, 1943). Many UA accessions are vigorous, widely adapted including to full sunlight, early bearing, more resistant to diseases and pests than other groups but have often small beans. They have been widely used in breeding programmes worldwide. The Trinitario (T) group is derived from crosses between Criollo and Amelonado. Trinitarios show wide morphological variation, relatively large beans, high vigour and good productivity but are quite

susceptible to pest and diseases (Eskes and Lanaud 2001, Paulin and Eskes 1995).

Cocoa genetic improvement programmes have been based on the creation and selection of bi-parental crosses (hybrid varieties) with yield as the main criterion. In Côte d'Ivoire, since the introduction of UA materials in 1954, hundreds of 'single pair' crosses between different groups have been made and the best crosses, mainly of the UA × LA type (Amelonado), have been selected and distributed to farmers. Varieties of this type are high and early yielding in Côte d'Ivoire (Besse 1977). Due to the spreading of *Phytophthora* Pod rot (Ppr), emphasis was made on the resistance to diseases in addition to yield. However, the breeding strategy uses only part of the genetic diversity available in collections and does not lead to a continuous improvement of selection traits, as is possible with reciprocal recurrent selection (RRS) (Gallais 1978).

From 1990 onward, a RRS scheme was adopted in Côte d'Ivoire based on two complementary genetic groups: selected UA accessions (providing good combining ability for early yield, adaptation and for resistance to Ppr) and a mixture of LA and T accessions (providing good yield potential, self-compatibility, large beans and good flavour). LA and T were mixed because these groups are genetically close; the inclusion of T types aimed at increased adaptation to the full sun, high yield, some level of resistance to Ppr and large beans. The main objective of the RRS scheme is to improve resistance traits and increase simultaneously the yield. The initial proposal includes firstly two cycles of recurrent selection (RS), to improve traits with good heritability (e.g. bean size and resistance to Ppr), followed by RRS to exploit additive as well as dominant gene effects (Clément et al. 1994).

The first RS cycle, containing the two complementary genetic group and more than 250 intra-group progeny, was completed in 1999 (Lachenaud et al. 2001). Most of the second cycle parental clones were selected not only as single trees in the first cycle intra-group progeny but also some additional parental clones were added (Lachenaud et al. 2001). Some parents of the base populations were not anymore represented in the second cycle (no descendants were selected). This selection procedure might have affected the genetic diversity within and between the groups. Furthermore, the genetic identity of several parental

clones used in the programme was not always sufficiently known and mis-identifications in cocoa germplasm collections are quite common (Lachenaud et al. 2001). The objective of this study was to analyse the coherence of the genetic identity and the evolution of genetic diversity in the parental populations of the UA and LA + T groups between the first and second selection cycles by using neutral polymorphic molecular markers [simple sequence repeat (SSR) markers].

Material and Methods

Plant material: First-cycle parental populations: The parental populations used for the first cycle within-group crosses of the RS programme are Upper Amazon population identified as (UA)0 and the mixture of Lower Amazon and Trinitario identified as (LA + T)0. These clones were selected based on phenotype values, combining abilities and genetic diversity provided by isozyme markers (Lanaud, 1987; Clément et al. 1994).

UA0 contains 20 parental clones that belong mostly to the collections made by Pound (1938, 1943) in the Upper Amazon region (Table 1). Some clones are known to possess good resistance to Ppr (SCA6, P7, PA150, P19A, IMC57, T79/501, T60/667 and MO98), good combining ability for yield (SCA6, PA150, P7, T60/887, T79/501, IMC67, UPA409 and NA32) or relative low vigour (AMAZ15-15, PA150, SCA6, T60/887) (Clément et al. 1994, Lachenaud et al. 2001).

The 28 parental clones used for the (LA + T)0 population (Table 1) contain Amelonado types selected in Africa (IFC1, IFC5, N38 and IFC303) or in America (MAT1-6, MAT1-9 and IFC371 -a Catongo type-), Trinitarios selected in Africa (ACU85, IFC6, IFC8, IFC11, IFC14, IFC29, IFC304, R15POS, SNK12, W41) or in America (CC10, ICS6, ICS46, ICS84, ICS89, ICS95, GS29, UF221, UF676).

Second-cycle parental populations: The second-cycle parental populations, named UA1 and (LA + T)1, respectively, consisted mainly of trees selected in the first cycle intra-population crosses (individual-family selection).

The UA1 parental population is composed of 40 genotypes, 28 of which were selected from the first cycle crosses and 12 were introduced as additional genotypes. The twelve additional genotypes included three clones (UPA402, IFC705 and H1258) that had shown good combining ability for yield, six trees selected in UA crosses (involving PA150, PA121, PA13, IMC67 and P19A parents) and three trees selected in selfed progeny obtained from UPA409, UPA419 and UPA608 that were selected for good yield and resistance to Ppr in the leaf disc test (Lachenaud et al. 2001).

The (LA + T)1 parental population was also composed of 40 genotypes, 31 of which were trees selected in the first cycle within-group crosses. The nine additional genotypes included three African Trinitarios (IFC7, IFC10 and IFC18; Besse 1976) and six Amelonado trees selected for yield and low Ppr incidence. Two of these latter trees

were selected in the progeny IFC5 × IFC15 and four in Amelonado varieties previously grown in French Guiana (IRA1, DCG1-1, DCG1-2 and DCG2-1) (Lachenaud et al. 2001).

Methods: DNA extraction: DNA was isolated from 0.5 g of leaves collected from trees growing at the experimental station of CNRA in Divo, Côte d'Ivoire. Leaves were cleaned, frozen in liquid nitrogen, ground. DNA isolation was performed as described by Risterucci et al. (2000) with a buffer containing 100 mM Tris-HCl pH 8.0, 2% MATAB, 20mM EDTA, 1% PEG6000 and 0.5% sodium sulphite. The DNA solution was then purified using the phenol-chloroform method (Karakousis and Langridge 2003).

Molecular analysis: *In vitro* amplification has been performed by Polymerase chain reaction (PCR) with 13 microsatellite primers (Lanaud et al. 1999). Each primer was used to amplify 2 ng of DNA in 5 µl of reaction mixture using the PTC 200 instrument (manufactured by MJ Research, Watertown, MA, USA). DNA migration and staining were performed with the ABI 3100 (Applied Biosystem, Foster City, CA, USA) machine according to the method described by Bhattacharjee et al. (2004).

Genetic diversity parameters: The alleles of the microsatellite loci were scored according to their size. The data obtained were used to estimate the following genetic diversity parameters: average number of alleles per locus, percentage of polymorphic loci, observed heterozygosity (Ho) and gene diversity or expected heterozygosity (He) (Nei 1987). The estimated value of total gene diversity (Ht) was subdivided into within-population (Hs) and between-population (Dst) diversity, where Ht = Hs + Dst. The coefficient of gene differentiation (Gst) is calculated as Gst = Dst/Ht and is an estimate of the proportion of the diversity present between the populations in relation to the total diversity. The genetic distance between the populations was calculated by Nei's (1978) unbiased method GENETIX (Genetix 4.05.2, CNRS, Montpellier, France), TFPGA (TFPGA 1.3, North Arizona University, CA, USA) and F-STAT (F-STAT 2.9.3.2, Institute of Ecology, Lausanne, Switzerland) software packages were used to calculate the genetic diversity parameters.

Factorial correspondence analysis: The factorial correspondence analysis (FCA) was used to study the degree of genetic similarity between individuals (Escoffier and Pagès 1998). In this analysis, codes 1 and 0 were used to indicate respectively the presence and absence of an allele at a given locus. The analysis was performed by XL-STAT software (XL-Stat 7.5.2, Microsoft Company, Redmond, WA, USA).

Results
Allele richness

Among the 13 microsatellite primers tested, 12 produced scorable alleles. Primer mTcCIR1 had to be removed from the

Table 1: The first recurrent selection cycle parental clones of the two complementary genetic groups UA (Upper Amazon) and LA + T (Lower Amazon and Trinitario), and clones incorporated in the second cycle parental populations

Genetic groups	
UA	LA + T
Cycle 1 parental population AMAZ15-15, IMC6, IMC57, IMC67 ¹ , IMC78, MO81, MO98, NA32, NA58 ¹ , P7, P19A, PA150, SCA6, PA4 ¹ , T60/887, T79/501, UPA401 ¹ , UPA409, UPA413, <u>UPA134¹</u>	Cycle 1 parental population IFC1, IFC5, IFC6, IFC8, IFC11, <u>IFC14</u> , IFC304, IFC29 ¹ , IFC303, SNK12, W41, ICS6 ¹ , ICS95, ICS89, UF221 ¹ , UF676, GS29, ICS84, WA40, ICS46, IFC371, ACU85, N38, CC10, <u>MAT1-6¹</u> , MAT1-9, VENC4-11, <u>R15POS¹</u>
Clones added in cycle 2 IFC705, H1258, UPA402	Clones added in cycle 2 DCG1-1, IRA1, DCG2-1, DCG1-2, IFC7, IFC10, IFC18.

Underlined clones were not included in the molecular analyses.
¹Clones not represented in the second cycle.

Table 2: Allelic richness in the four populations studied and modifications between the two selection cycles for each genetic group

Type of alleles	UA sub-group			(LA+T) sub-group		
	UA0	UA1	Δ^1	(LA+T)0	(LA+T)1	Δ
Total number of which:	83	78	-5	70	57	-13
frequent alleles	59	58	-1	42	33	-9
rare alleles	24	20	-4	28	24	-4
Private alleles ²	5	4		5	3	
Lost alleles	-7			-16		
New alleles		+2			+3	

UA, Upper Amazon Forastero; LA, Lower Amazon Forastero; T, Trinitario.

¹Delta (Δ) is the change in the number of alleles between cycle 1 and cycle 2.

²Private alleles are alleles that are specific to each of the four sub-groups; all private alleles belonged to the rare alleles.

analyses due to the poor quality of its profiles. In total, 98 alleles were identified on the 12 loci. No population contained the totality of the alleles (Table 2). UA0 contained 83 alleles and UA1 had 78 alleles. In the second genetic group, (LA+T)0 contained 70 alleles and (LA+T)1, 57 alleles. The frequently occurring alleles among which 30 are common for the four populations were similar for the UA0 population (59) and UA1 population (58). In LA+T population, the number of frequently alleles was 42 in (LA+T)0 and 33 in (LA+T)1. There were relatively more rare alleles (with a frequency under 5%) in UA0 (24 alleles) than in UA1 (20 alleles). There were also relatively more rare alleles in (LA+T)0 (28 alleles) than in (LA+T)1 (24 alleles). A total of five private alleles in UA0 and (LA+T)0, four in UA1 and three in (LA+T)1 was found (Table 2). All private alleles were part of the rare alleles of these populations. Seven alleles have been lost in UA1 and sixteen in (BA+T)1. The number of new alleles introduced is two for UA1 and three for (BA+T)1.

Ninety-three alleles were identified in the two parental populations of the first cycle among which 57 was common. In the second cycle population, the number of alleles identified in the parental population is 84 among which 50 are common.

Genetic diversity

The 12 loci analysed were highly polymorphic in the parental populations both of cycle 1 and of cycle 2 of the RS programme. The average H_o for all the four populations and 12 loci was high (0.55). The H_o values were higher than average for eight loci and lower for four loci (Table 3). H_t was high for UA populations (0.72) and (LA+T) populations (0.59). Average H_s diversity was also high for UA ($H_s = 0.71$) and for (LA+T) ($H_s = 0.59$) while D_{st} diversity was low: $D_{st} = 0.002$ for UA and $D_{st} = 0.004$ for (LA+T). Hence, the G_{st} between populations was only 0.07. The H_o value for UA0 ($H_o = 0.53$) increased for UA1 ($H_o = 0.64$) whereas from (LA+T)0 to (LA+T)1 H_o decreased from 0.57 to 0.48. H_e was similar for UA0 (0.72) and UA1 (0.71) but higher for (LA+T)0 (0.62) than for (LA+T)1 (0.56). The average number of alleles was slightly higher for UA0 (6.91) than for UA1 (6.5). Also in the second genetic group, the average number of alleles is higher for (LA+T)0 (5.83) than for (LA+T)1 (4.75) (Table 4). H_o and H_e values were much higher for UA1 than for (LA+T)1 as well as the average number of alleles per locus: 6.50 for UA1 and 4.75 for (LA+T)1.

Table 3: Genetic diversity parameters for each simple sequence repeat locus

Locus	N	H_o	H_t	H_s	D_{st}	G_{st}
mTcCir 3	11	0.77	0.86	0.81	0.05	0.07
mTcCir 6	10	0.54	0.75	0.69	0.07	0.09
mTcCir 12	7	0.57	0.83	0.74	0.09	0.11
mTcCir 15	9	0.71	0.83	0.79	0.04	0.05
mTcCir 19	9	0.63	0.75	0.72	0.03	0.04
mTcCir 21	7	0.66	0.65	0.63	0.02	0.04
mTcCir 24	7	0.39	0.45	0.45	0	0
mTcCir 25	10	0.54	0.75	0.69	0.06	0.08
mTcCir 26	6	0.47	0.71	0.67	0.04	0.13
mTcCir 9	9	0.46	0.63	0.59	0.04	0.06
mTcCir 17	4	0.31	0.37	0.36	0.01	0.03
mTcCir 18	9	0.58	0.76	0.68	0.09	0.11
Total	98	0.55	0.70	0.65	0.05	0.07

N, number of alleles; H_o , observed heterozygosity; H_t , total gene diversity; H_s , within-population diversity; D_{st} , between-population diversity; G_{st} , coefficient of gene differentiation.

Table 4: Average number of alleles per locus (N), observed heterozygosity (H_o) and expected heterozygosity (H_e) per population (standard deviation in brackets)

	N	H_o	H_e
UA0	6.91	0.53 (0.23)	0.72 (0.14)
UA1	6.50	0.64 (0.18)	0.71 (0.17)
(LA+T)0	5.83	0.57 (0.13)	0.62 (0.14)
(LA+T)1	4.75	0.48 (0.17)	0.56 (0.17)

UA, Upper Amazon Forastero; LA, Lower Amazon Forastero; T, Trinitario.

Table 5: Nei's (1978) genetic distances between the parental populations of the first and second recurrent selection cycles

	UA0	UA1	(LA+T)0
UA1	0.007		
(LA+T)0	0.29	-	
(LA+T)1	-	0.31	0.013

UA, Upper Amazon Forastero; LA, Lower Amazon Forastero; T, Trinitario.

Population structuring

The two UA parental populations used for the first and second selection cycles had a very low genetic distance of 0.007 (Table 5). For LA+T, the genetic distance between the two populations was also low (0.013). However, between the two genetic groups the distances were quite high: 0.29 for the UA0 and (LA+T)0 populations and 0.31 for the UA1 and (LA+T)1 populations (Table 5).

Factorial analyses of correspondence

Factorial analyses of correspondence were performed to visualize population diversity and genetic similarity; firstly for each base population and additional individuals introduced in the second cycle (Figs 1 and 2), then with the two base populations together (Figs 3 and 4), and finally with the second-cycle populations together (Fig. 5).

Figure 1 shows that the UA0 parental population might consist of three sub-groups. The four IMC clones were close to

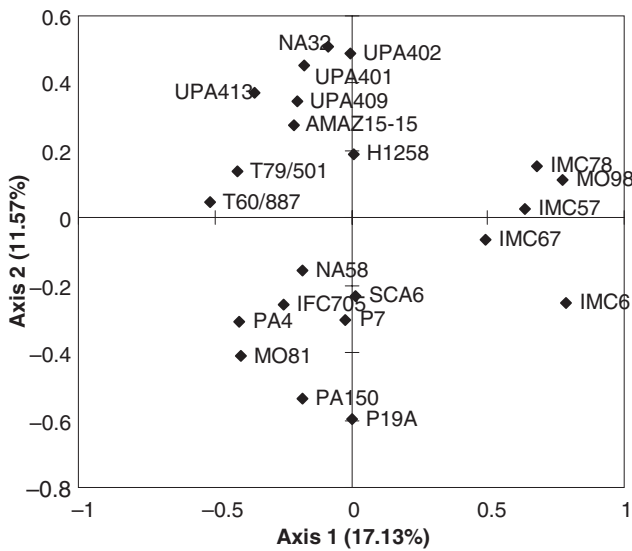


Fig. 1: Plane of axes 1 and 2 of the FCA for population UA0 plus three additional clones used in the UA1 parental population (UPA402, IFC705 and H1258)

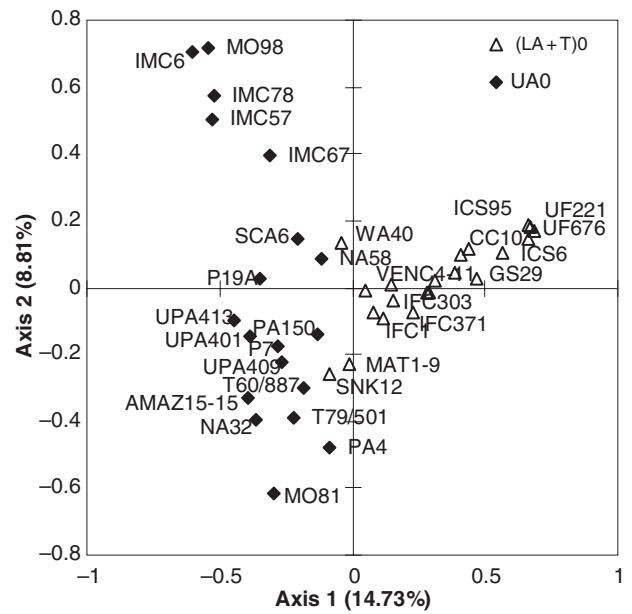


Fig. 3: Plane 1 and 2 of the FCA for the two parental populations used to create the first cycle inter-group crosses: UA0 and (LA + T)0

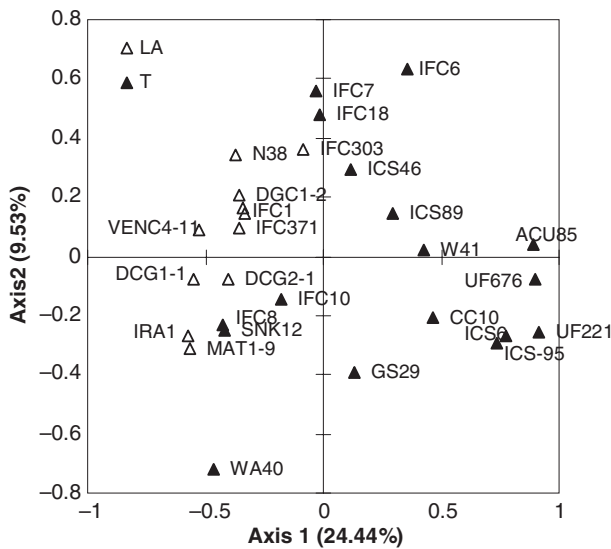


Fig. 2: Plane of axes 1 and 2 of the FCA for population (LA + T)0 plus seven additional clones used in the (LA + T)1 population (DCG1-1, IRA1, DCG2-1, DCG1-2, IFC7, IFC10, IFC18). IFC304 is superposed on IFC10, IFC11 on IFC7 and IFC5 on IFC1

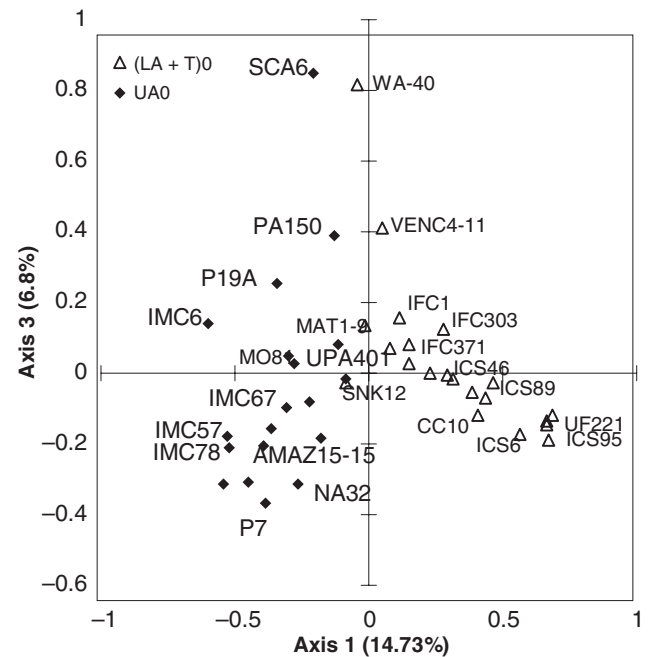


Fig. 4: Plane 1 and 3 of the FCA for the two parental populations used to create the first cycle inter-group crosses: UA0 and (LA + T)0

each other and also to MO98, while MO81 is located in a second group, close to PA4 (as in Fig. 3). A third group was composed of clones T79/501, T60/887 and H1258, NA32 (which is a parent of the previous three clones), four UPA clones (all progeny of IMC60 × NA34) and AMAZ15-15. The location of IFC 705, supposed to be an open-pollinated progeny of NA32 (Besse 1976), was a bit unexpected. The FCA plane of axes 1 and 3 (not shown) revealed that clone SCA6 was detached from the second group (as expected from its unique geographical origin among UA clones).

The FCA performed with (LA + T)0 showed a clear separation between the Amelonado clones (such as IFC1, IFC5, IFC371 and N38, grouped closely together), all located at the left side of Fig. 2, and the American Trinitario clones (UF,

ICS, GS and CC clones) located at the right side. Some supposedly African Trinitario clones (IFC8 and SNK12) grouped with Amelonado, whereas other African Trinitarios of the (LA + T)0 population were located at the upper centre-right side of Fig. 3, at the extreme end of the distributions of the American Trinitario and Amelonado clones. Two African Trinitario clones (ACU85 and W41) grouped well together with the American Trinitario clones. VENC4-11, from Venezuela, was located with the Amelonado clones. WA40, introduced from Wageningen in 1965 as a supposedly Trinitario clone, showed to be isolated in Fig. 2. This was

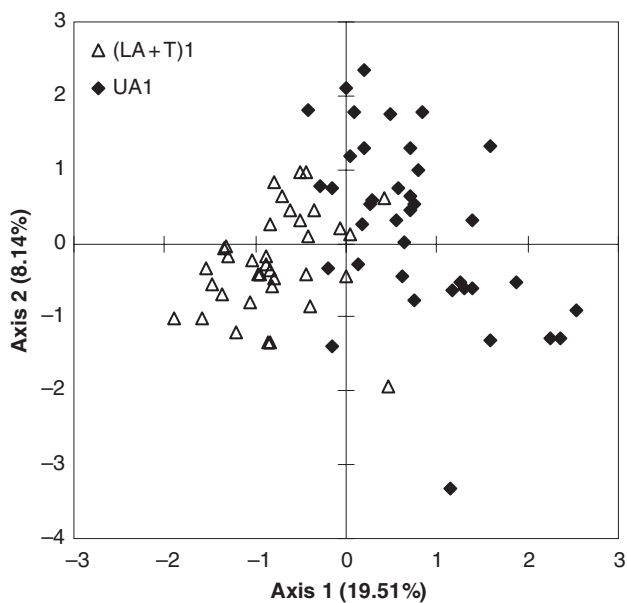


Fig. 5: Plane 1 and 2 of the FCA for the two parental populations used to create the second cycle inter-group crosses: UA1 and (LA+T)1

confirmed in Fig. 3, where this clone is very close to SCA6. The clones added to the LA+T group in the (LA+T)1 population (DCG1-1, DCG1-2, DCG2-1 and IRA1 from French Guiana, and IFC7, IFC10 and IFC18) were close to the well-known Amelonado clones.

When the two parental populations used for the first-cycle crosses were analysed together the first axis of the FCA separated the UA0 population from the (LA+T)0 population, while the second axis explained H_s variation (Fig. 3). The American Trinitario clones (ICS89, UF221, UF 676 and ICS95) were farther away from the UA materials than the Amelonado and African Trinitario individuals. Clones SCA6 and WA 40 were close to each other and clearly isolated from their respective groups by axis 3 (Fig. 4). The FCA of the second-cycle parental populations (Fig. 5) revealed higher variability for population UA1 than for (LA+T)1.

Discussion

Conformity of clones and homogeneity of populations

The results presented in this study led to answer some earlier doubts with regard the coherence of the LA+T population, which contains clones which genetic origin was less-well known, such as VENC4-11, clones from French Guiana and WA40 (Lachenaud et al. 2001). Our results show that VENC4-11 was correctly introduced into the LA+T population; in fact it appears to be very near to other pure Amelonado. The same appeared to be the case for the four selections made in cultivated cocoa from French Guiana which were introduced in the second cycle. However, in the case of WA40, Fig. 4 shows that this clone is very close to clone SCA6, from the UA group. That position seemed to be determined by alleles 358 of locus mTcCIR19, 209 of locus mTcCIR3 and 342 of locus mTcCIR 18, which are rare alleles found in SCA6 and WA40, and therefore suggest that the WA40 clone might be a progeny of SCA6. The analysis of the two genotypes revealed that WA40 might be a progeny of clone SCA6, and its location on the FCA graph suggested that it could be a hybrid between

SCA6 and a Trinitario clone. Original WA40 is considered as derived from two unselected plant; one of G8 \times DR38 and the other DR1 \times DR38 (Lockwood and Gyamfi 1979). G8 and DR1 are Criollo clones while DR38 is Trinitario. WA40 clone analysed here might be an off type.

Within the LA+T group, IFC303 showed a very good general combining ability during the first cycle intra-population crosses (Lachenaud et al. 2001), especially with Amelonado parents. This is surprising, as the other crosses among Amelonado genotypes did not provide good results. IFC303 is considered to be an Amelonado introduced from Ghana under the name P4/9: J1/14/5. However, the SSR loci obtained for ICF303 confirmed that this clone is really an Amelonado or closely related to the Amelonado group.

The close genetic similarity observed in this study between the Amelonado and African Trinitario genotypes had already been found using isozyme and RFLP markers (Lanaud 1986, Laurent et al. 1994). The strong presence of Trinitario alleles in population (LA+T)0 explains the higher H_o value for that population than in the UA0 population.

In population UA0, the two MO clones had ambiguous positions in the 1×2 plane of the FCA (Fig. 1). According to Boccara (Boccara M., personal communication), MO81 seems to be a mislabelled clone. Due to its position in the International Cocoa Genebank in Trinidad, it could be in fact a neighbouring PA accession. The proximity of PA4, PA150 and MO81 in Fig. 4 appears to confirm that hypothesis. Concerning MO98, this clone is not registered in the Trinidad collection, where it is supposed to come from (M. Boccara, personal communication). Our results showed clearly that MO98 is different from MO81 and genetically close to the IMC group, something which has been found earlier with RFLP markers (Motamayor et al. 2002). MO98 might therefore be a misidentification of one of the IMC clones.

An important finding was that the clones added to the second cycle parental populations UA1 and (LA+T)1 showed all good conformity with the groups to which they were added.

Comparison of population diversity

Gene diversity was higher in population UA0 than in population (LA+T)0, even with the number of individuals in UA0 being lower ($n = 20$) than that in (LA+T)0 ($n = 28$). The high genetic diversity of the UA group has already been shown before (Lanaud 1986; Laurent et al., 1993; N'Goran et al., 1994).

In all parental populations, H_o was lower than the H_e . The occurrence of a deficit in heterozygotes has been a common observation in cocoa populations (N'Goran et al. 2000). The biggest difference between H_o and H_e occurred in population UA0. As UA0 contains many genotypes that have been collected from the centre of diversity, this result would confirm a degree of allele fixation in the natural populations. The smaller difference between H_o and H_e in UA1 compared to UA0 is likely caused by some degree of random recombination of alleles in the cycle 1 crossing scheme.

The total number of alleles in the second-cycle populations decreased moderately (6%) for the UA group and more severely for the LA+T group (19%). The reduction was highest for the frequent alleles of the (LA+T) population, decreasing from 42 to 33. Therefore, this reduction in alleles might be due to a reduction of Criollo alleles that are specific to the LA+T population. However, the number of frequent

alleles in common between the UA and (LA + T) populations was not affected.

The small number of private alleles (in line with a low G_{st}) decreased, from five to four in the UA genetic group and from five to three in the (LA + T) genetic group. Therefore, selection slightly reduced H_e in the (LA + T) group, with a decrease of 0.62 to 0.56, while no significant change was observed in the UA group with values of 0.72 and 0.71, respectively.

Selection at the end of the first cycle led to elimination of five genotypes in each of the UA and LA + T groups from participation in the second cycle crosses, i.e. 25% of the initial parental genotypes for UA and 18% for (LA + T). At the same time, twelve and nine new genotypes were added to the UA1 and (LA + T)1 populations, respectively. The elimination and addition of genotypes between the first and second cycle selection cycles must be related to the disappearance of seven and sixteen alleles from the UA0 and (LA + T)0 populations and the appearance of two and three new alleles in the UA and (LA + T) groups, respectively (Table 2). The reduction of rare common alleles from 10 to 5 between the first and second-cycle parental populations did not affect the genetic distance, which remained stable (0.31 as opposed to 0.30). This must be ascribed to the fact that the genotypes that have disappeared in the cycle 1 selection process and that were newly introduced in the cycle 2 parental populations were representative for those populations.

Conclusion

Firstly, this study has made it possible to evaluate the correctness and conformity of the genotypes used in the parental populations of the two genetic groups that are at the basis of the cocoa RS programme in Côte d'Ivoire. The ambiguity of certain parental clones was elucidated. The genetic diversity of the two genetic groups, UA and (LA + T) was not greatly reduced after the first RS cycle. The addition of new parents at the beginning of the second selection cycle may have been partly responsible for that situation. The genetic diversity parameters of the UA group were unchanged, whereas the observed (and expected) heterozygosity of the (LA + T) group decreased slightly. Such homogenization may be favourable, as the composition of the original parental population (LA + T)0 was rather artificial and heterogeneous. The information provided by the molecular markers has proved very useful to guide the continuation of the RS programme in Côte d'Ivoire.

Perspectives for varietal output and for the RS programme

Inter-group crosses between some selected trees of the two populations UA1 and (LA + T)1 have been made and were planted in 2001 at Divo in varietal trials. The control varieties, which are recommended inter-group hybrid varieties, are the same as used in the intra-group crossing plots. This has been the first approach to select new inter-group hybrid varieties in the RS programme. However, the high genetic diversity within each of the genetic groups, confirmed in this study, may explain why several of the first and second -cycle intra-group crosses yielded as well or even better than the inter-group control varieties (Lachenaud et al. 2001, N. D. Pokou, J. A. K. N'Goran, Ph. Lachenaud, A. B. Eskes, M. Kolesnikova-Allen, D. Clément, and A. Sangaré, personal communication). This was especially the case for the UA intra-group crosses. These findings justify a second approach, i.e. selection of new

commercial varieties by further evaluation of the best within-group crosses together with the best intra-group crosses. These findings also have a bearing on the future of the RS programme. Instead of continuing the programme as a RRS programme, as was originally planned (Clément et al. 1994), it appears feasible to exploit both the best inter-group as well as the best intra-group crosses. The full exploitation of the second cycle for new variety output would best involve a comparison of crosses between the best parents selected in the second cycle, both in inter- and in intra-group crosses. Such a comparison would throw more light on how best the RS programme is to be continued. In the case that the inter-group crosses outperform the intra-group crosses, the programme should best be continued as a RRS programme. In case the performance of the inter- and intra-group crosses is similar, the improvement of the UA and (LA + T) groups between cycles should be based both on intra and inter-group cross performance.

Acknowledgements

Our thanks go to Peter Biggins for his contribution to English translation of the manuscript, to Dr Bhattacharjee for her help in the use of the software and to Mrs Emile Liadé and Sunday Taiwo, technicians at the CNRA/LCB and IITA/CBL laboratories, for their help in the laboratory experiments.

References

- Bartley, B. G. D., 2005: The Genetic Diversity of Cacao and its Utilization. CABI Publishing, London, 341.
- Besse, J., 1976: Note sur le matériel végétal représenté dans la collection de l' IFCC en Côte d'Ivoire. Institut française du café et du cacao, génétique cacaoyer Bingerville, Technical report of IFCC.
- Besse, J., 1977: La sélection générative du cacaoyer en Côte-d'Ivoire : bilan et orientation des recherches. In: Proc. 5th International Cocoa Research Conference, Ibadan, Nigeria, 95—103.
- Bhattacharjee, R., P. Aikpokpodion, M. Kolesnikova-Allen, K. Badaru, and R. Schnell, 2004: West African cocoa: a pilot study on DNA fingerprinting of the germplasm from the Cross River State of Nigeria. *Ingenic Newsl.* **9**, 15—20.
- Cheesman, E. E., 1944: Note on the nomenclature, classification and possible relationships of cacao populations. *Trop. Agric.* **21**, 100—116.
- Clément, D., A. Eskes, O. Sounigo, and J. N'Goran, 1994: Amélioration génétique du cacaoyer en Côte-d'Ivoire, présentation d'un nouveau schéma de sélection. In: Proc. 11th Internationale Cocoa Research Conference, Yamoussoukro, 451—457.
- Escoffier, B., and J. Pagès, 1998: Analyse factorielle simple et multiples : objectif, méthodes et interprétation. Dunod, Paris, 277 p.
- Eskes, A. B., and C. Lanaud, 2001: Cocoa. In: A. Charrier, M. Jacquot, S. Hamon, D. Nicolas, and M. K. Razdan (eds), *Tropical Plant Breeding*, 78—105. Repères CIRAD, Montpellier, France.
- Gallais, A., 1978: Amélioration des populations, méthodes de sélection et création des variétés. III Bases théorique pour l'étude de la sélection réciproque. *Ann. Amélioration plantes*, **28**, 637—660.
- Karakousis, A., and P. Langridge, 2003: A high-throughput plant DNA extraction method for marker analysis. *Plant Mol. Biol.* **21**, 95a—95f.
- Lachenaud, P., A. Eskes, J. A. K. N'Goran, D. Clément, I. Kébé, M. Tahí, and C. Cilas, 2001: Premier cycle de sélection récurrente en Côte d'Ivoire et choix des géniteurs du second cycle. In: 13 ème Conférence International sur la recherche cacaoyère, Kota Kinabalu, Malaisie, 9—14.
- Lanaud, C., 1986: : Utilisation des marqueurs enzymatiques pour l'étude génétique du cacaoyer : *Theobroma cacao* L. *Café Cacao Thé* **25**, 259—279.

- Lanaud, C., 1987: Nouvelles Données sur la Biologie du Cacaoyer (*Theobroma cacao* L.) Thèse de Doctorat D'état. Thèse Université de Paris-Sud, France. 106.
- Lanaud, C., A. M. Risterucci, I. Pieretti, M. Falque, A. Bouet, and P. J. L. Lagoda, 1999: Isolation and characterisation of microsatellites in *Theobroma cacao* L. *Mol. Ecol.* **8**, 2141—2152.
- Laurent, V., A. M. Risterucci, and C. Lanaud, 1994: Genetic diversity in cocoa revealed by cDNA probes. *Theor. Appl. Genet.* **88**, 193—198.
- Laurent, V., A. M. Risterucci, and C. Lanaud, 1993: Chloroplast and mitochondrial DNA diversity in *Theobroma cacao*. *Theor. Appl. Genet.* **87**, 81—88.
- Lockwood, G., and M. M. O. Gyamfi, 1979: The CRIG cocoa germplasm collection with notes on codes used in the breeding programme at Tafo and elsewhere. *Tech. Bull.* **10**, 60.
- Motamayor, J.-C., A. M. Risterucci, P. A. Lopez, C. F. Ortiz, A. Moreno, and C. Lanaud, 2002: Cocoa domestication I: the origin of the cacao cultivated by the Mayas. *Heredity* **89**, 380—386.
- N'Goran, J. A. K., V. Laurent, A. M. Risterucci, and C. Lanaud, 1994: Comparative genetic diversity studies of *Theobroma cacao* L. using RFLP and RAPD markers. *Heredity* **73**, 589—597.
- N'Goran, J. A. K., V. Laurent, A. M. Risterucci, and C. Lanaud, 2000: The genetic structure of cocoa populations (*Theobroma cacao* L.) revealed by RFLP analysis. *Euphytica* **115**, 83—90.
- Nei, M., 1978: Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics* **89**, 583—590.
- Nei, M., 1987: *Molecular Evolutionary Genetics*. Columbia University press, New York.
- Paulin, D., and A. B. Eskes, 1995: Le cacaoyer : stratégies de sélection. *Plant. Recherche Dév.* **2**, 5—18.
- Pound, F. J., 1938: Cocoa and Witches' Broom Disease (*Marasmius perniciosus*) of South America, with Notes on Other Species of *Theobroma*. Yuile's printerie, Port-of-Spain, Trinidad, 52.
- Pound, F. J., 1943: Cocoa and Witches' Broom disease (*Marasmius perniciosus*) In: H. Toxopeus (ed.), *Reprint Archives of Cacao Research*, 73—92. American Cocoa Research Institute, Eyre and Spottiswoode Limited, England.
- Risterucci, A. M., L. Grivet, J. A. K. N'Goran, I. Pieretti, M. H. Flament, and C. Lanaud, 2000: A high-density linkage map of *Theobroma cacao* L. *Theor. Appl. Genet.* **101**, 948—955.