African Journal of Biotechnology Vol. 5 (16), pp. 1441-1449, 17 August 2006
Available online at http://www.academicjournals.org/AJB
ISSN 1684–5315 © 2006 Academic Journals

Full Length Research Paper

Genetic diversity in cocoa germplasm of southern Cameroon revealed by simple sequences repeat (SSRS) markers

M.I.B. Efombagn[1*], O. Sounigo[2], S. Nyassé[1], M. Manzanares-Dauleux[3], C. Cilas[2], M. A.B. Eskes[2] and Kolesnikova-Allen[4]

[1]Institute of Agricultural Research for Development (IRAD), P.O. Box 2067 or 2123, Yaoundé, Cameroon.
[2]CIRAD-CP, TA 80/02, 34398 Montpellier Cedex 5, France.
[3]Agrocampus de Rennes, 65 Rue de St Brieuc, 35042, Rennes, France.
[4]Central Biotechnology Laboratory, International Institute of Tropical Agriculture (IITA), Ibadan, Nigeria.

Accepted 29 May, 2006

The range of polymorphism of about 194 cocoa accessions collected in farms in Southern Cameroon during field surveys and 71 Trinitario and Upper Amazon clones available in genebanks on-station was assessed using 13 SSR markers. The gene diversity, genetic differentiation and genetic similarities were analysed for the different populations. In total, 282 alleles were detected within all the populations studied. The farm accessions were strongly differentiated based on their geographical origin, with accessions coming from the East province clustering together with local Trinitario accessions from the genebank. Accessions from the Centre-South provinces clustered with Amazon and hybrid accessions, suggesting more uptake of seed garden materials in farms in these provinces. The genetic diversity parameters indicate that the farmers' planting material is not highly diverse, and is genetically close to parental genotypes available in genebanks. However, some promising Upper Amazon clones (T-clones) that have also been used as parents of released hybrid varieties were genetically distant from the accessions. This result suggests that the progenies of these parents have so far been poorly used in the cocoa farms surveyed. The consequences of these findings for cocoa breeding in Cameroon are discussed.

Key words*:* Cocoa, SSR markers, polymorphism, agronomic traits, breeding.

INTRODUCTION

Cocoa (*Theobroma cacao*) belongs to the family Malvaceae (Alverson et al., 1999). Cocoa originated from South and Central Americas and was introduced into Cameroon from Sao Tome and Principe in the nineteen century (Nya Ngatchou, 1981). In Cameroon, cocoa cultivation has been expanded since the 1960s, and is currently one of the major sources of revenues of rural households (1 to 2 millions of people) of the forest agro-ecological zones in the country (South and South-Western parts). Cocoa is grown in more than 200 000 farms and the total cocoa growing surface is estimated to be 400 000 hectares. The national dry bean cocoa production remains stable, around 120,000 tonnes per year, but has recently increased to 140, 000 tonnes (ICCO, 2004). However, cocoa cultivation is actually subjected to several constraints such as the ageing of the trees (mainly in Southern Cameroon where 40% of cocoa trees were planted before 1960), poor soils, susceptibility to black pod disease (causing 80-90% losses without chemical control according to Despréaux et al. (1989)) and susceptibility to cocoa mirids. Hence, the average production level per farm is currently very low, especially so in the Southern production zones of the country, where yields vary between 150 to 300 kg dry cocoa per ha.

*Corresponding author. E-mail: efombagn@yahoo.fr. Tel: 237 942 76 46 / 223 89 49.

The genetic diversity (origin and type of material grown) of cultivated cocoa remains largely unknown. However, it is estimated to be constituted mainly of traditional varieties, notably the so-called "German Cocoa" which is the results of natural hybridisation between Trinitario and Amelonado types (first populations introduced in the country); and populations introduced from Ghana (Upper Amazon origin) that appear to have been distributed also in Western Cameroon. It is expected that a relatively small percentage of hybrid varieties have been used. The seed garden parents were mainly locally selected Trinitario types and also introduced Upper Amazon clones (UPA and T-clones). The lack of manual pollination in the seed gardens has probably induced a high rate of self-pollination among the seed distributed to the farmers. The cocoa growers are presently not very satisfied with the hybrid varieties, because of tree mortality ("die-back" of which the cause is not well identified) and the reported higher susceptibility to disease and pests. Therefore, in general, the farmers harvest cocoa pods on their grown trees to obtain seed for new plantings (farm extension or replacement of dead trees).

The objective of the current study is to assess, with the means of molecular tools, the genetic diversity and the level of homozygosity of cocoa accessions in farmers fields. This variation is compared to the variation of accessions from seed gardens and from on-station cocoa collections available in Cameroon. The results may also have an impact on cocoa breeding strategies, with regard to the need to enlarge the genetic diversity for economically important traits, such as disease resistance, and with regard to the possible use of farm accessions in future selection programmes.

MATERIAL AND METHODS

Cocoa genotypes

About 194 cocoa accessions collected on-farm and 71 clones available in collections of the Institute for Agricultural Research for Development (IRAD) at Nkoemvone Station (South Cameroon) were selected to constitute the subset for the study. The accessions were collected in different farms distributed widely in three different administrative provinces of the Southern Cameroon's cocoa growing areas (Figure 1): South, East and Centre provinces. A description of the accessions is given below.

Breeders' accessions: They are composed of parental clones used in seed gardens for production of hybrid seed distributed to farmers from the 1960s onward, breeders' selections (cloned F1 progeny trees) and other accessions available in collections that have not been used in breeding. These include the following accessions:

- Local Trinitario clones: *SNK* ("Selection of Nkoemvone"): SNK7, SNK13, SNK16, SNK29, SNK30, SNK31, SNK32, SNK42, SNK48, SNK52, SNK67, SNK109, SNK111, SNK181, SNK213, SNK344, SNK348 , SNK376, SNK377, SNK392, SNK404, SNK413, SNK415, SNK416, SNK450, SNK460, SNK461, SNK467, SNK476, SNK479, SNK480, SNK504, SNK505 and SNK506.
- Cloned F1 hybrid genotypes: *SNK600* series (local Trinitario x Upper Amazonian genotypes): SNK600, SNK602, SNK604, SNK605, SNK607, SNK608, SNK609, SNK610, SNK611, SNK613, SNK614, SNK615, SNK616, SNK618, SNK619, SNK620, SNK622, SNK623, SNK624, SNK625, SNK627, SNK628, SNK630, SNK631, SNK632, SNK633, SNK634, SNK635 and SNK636.
- Upper Amazon Forastero clones (*UPA*; *T-Clones*): T60/1174, T60/887, T60/897, T79/467, T79/501, UPA134 and UPA143.

Farm' selections: Collection of farm accessions was carried out in farmers' fields in different geographical locations. The sampling methodology has been defined based on geographic representation or by density of cocoa farms in the regions of collection. Samples were collected in different administrative divisions distributed in three provinces constituting the Southern Cameroon cacao agro-forest ecology. All the accessions collected in each division were identified as follows (capital letters in italics represent the code of each division, and the number of accessions selected is in parenthesis):

- South Province: *SDL* : Dja-et-Lobo (9) ; *SVN*: Vallée-du-Ntem (18); *SMV*: Mvila (12).
- Centre Province: *CMA*: Mefou-et-Akono (10) ; *CMF* : Mefou-et-Afamba; *CLK*: Lekié (14); *CMI*: Mbam-et-Inoubou (9); *CMK*: Mbam-et-Kim (33); *CNM*: Nyong-et-Mfoumou (10); *CNS*: Centre Nyong-et-So'o (7).
- East Province: *EHN*: Haut-Nyong (19); *ELD*: Lom-et-Djerem (35).

Farm selections originated from the same division were identified in the study as a geographical population.

Discs of 5 mm diameter were collected from young, green and healthy leaves of selected trees and conserved in small tubes containing an NaCl-CTAB-Azide (70 g NaCl, 3 g CTAB, 0.04 g Na azide and 200 ml distilled water) solution.

DNA extraction

DNA was extracted following the basic DNA extraction developed at the Central Biotechnology Laboratory of IITA, Ibadan, Nigeria (Bhattacharjee et al., 2004). Four leaf discs per accession stored at 4°C in the NaCl-CTAB-Azide Buffer were removed with the help of forceps from eppendorf tubes, and placed inside a 1.2 ml propylene strip tubes with strip caps (Marsh Biomarket, USA, available as 12 x 8-well strips), and containing two pre-chilled 4 mm chrome-plated grinding balls bearing. The balls were dispensed using automatic balls dispenser. After adding the leaf discs, two more balls were then placed on the top of the leaf discs to ensure proper grinding of the leaf tissue.

For the extraction of 200 samples, enough CTAB buffer (7% w/v CTAB, 1.4 M NaCl, 20 mM Tris-HCl, ph 8.0, 0.17% β-mercaptoethanol), corresponding to 450 µl per sample, were dispensed into a disposable falcon tube and incubated in 65°C water bath. 450 µl of warm CTAB was added to each sample and capped tightly with polyethylene strip caps. The tubes across two racks (each with 96 samples) were balanced before loading into the GenoGrinder 2000 (Spex CertiPrep Inc, USA). Samples in GenoGrinder 2000 were grinded following the device's instructions, at 500 strokes/min for 10 -15 min. The grinded samples are incubated for 30 min in a 65°C water bath with occasional mixing. Tubes were removed from water bath and 400 µl of CIA (chloroform : isoamyl alcohol, 24:1) solvent are added to each sample. To mix well, the samples were tightly capped and inverted 2-3 times. The samples' plates were centrifuged at 3700 rpm in an Eppendorf

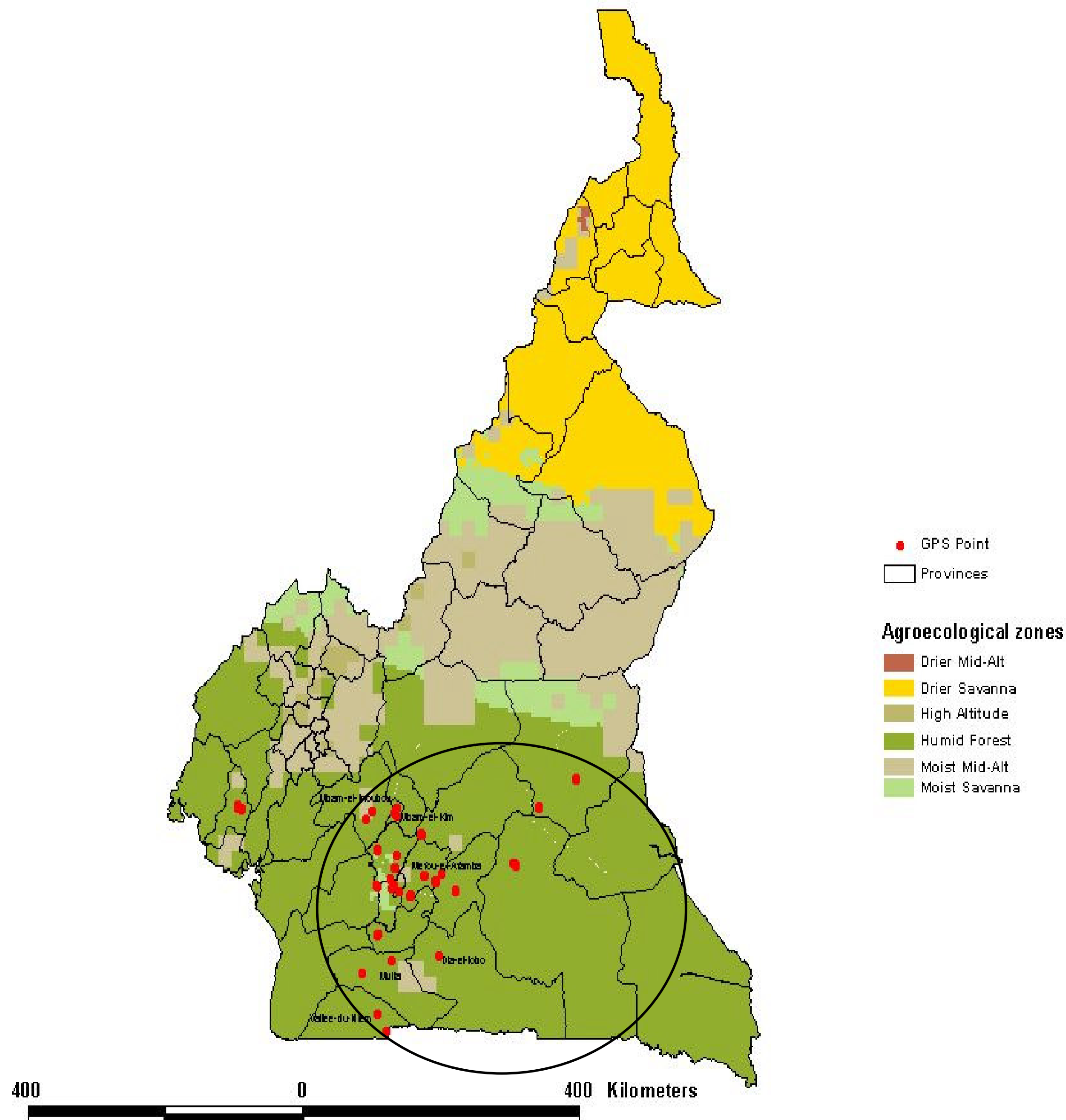

Figure 1. Country's map with the collection sites (in red) in the humid forest agro-ecological zones (South, Centre and East Provinces).

centrifuge model 5810 with (Swing-bucket rotor model A-2-DWP) for 30 min. About 300 µl of the supernatant formed were transferred into fresh strip tubes without disturbing the interface. Thereafter, the balls were removed and the used strip tubes were discarded.

To precipitate the DNA, 210 µl of 0.7 volume of cold Isopropanol were added to the supernatant and mixed by inverting the tubes. The plates were centrifuged at 3700 rpm at 30 min. The supernatant was carefully decanted without disturbing the pellet, and air dried for 20 min. To eliminate RNA, 203 µl of law-salt TE (200 µl $TE_{10}E_1$) and RNase (3 µl) are added in each sample, and placed at 37°C for 1 h. 200 µl of Phenol : Chloroform : Isoamylalcohol (PCI, 25:24:1) are added to each sample and mixed well by inverting the tubes, and centrifuged at 3700 rpm for 10 min. The aqueous layer formed is transferred to fresh strip tubes using multi-channel pipettes. 200 µl of CIA are added, mixed well by inverting the tubes 2-3 times and centrifuged at 500 rpm for 5 min. The newly formed aqueous layer is transferred into fresh tubes. To precipitate DNA, 15 µl ($1/10^{th}$ vol) 3 M sodium acetate and 300 µl of (2 vol) 100% ethanol are added to each sample (tube), and placed in a freezer at – 20°C for 5 min. After centrifugation at 3700 rpm for 15 min, the supernatant is carefully decanted without disturbing the pellet. The pellets are washed with 200 µl 70% ethanol and quickly centrifuged for 5 min. After pouring off the supernatant, the washed pellets are air-dried for 1 h. Each pellet is re-suspended in 100 µl low-salt TE and stored at 4°C. Aliquots of 3 µl of freshly extracted genomic DNA were electrophoresed on 0.8% agarose gel, stained with ethidium bromide, and visualised under UV trans-illuminator for quality DNA assessment. DNA quantity was estimated from the

Table 1. SSRs and originated tube names, dye (colour), forward and reverse primer sequences, annealing temperature, nucleotide repeats, and allele sizes.

Primer Name	Tube Name	Dye (Colour)	Forward and Reverse Sequence (5'-3')	Annealing Temperat. m(°C)	Nucleotide repeats structure	Allele size (bp)
mTcCIR 1	Y16883	HEX	gca ggg cag gct cag tga tgg gca acc aga aaa cga t	51	$(ct)_{14}$	198-260
mTcCIR 3	Y16977	6-FAM	cat ccc agt atc tca tcc att ca ctg ctc att tct ttc ata tca	46	$(ct)_{20}(ta)_{21}$	330-400
mTcCIR 6	Y16980	6-FAM	ttc cct cta aac tac cct aaa t tta agc aaa gca atc taa cat a	46	$(tg)_{7}(ga)_{13}$	107-191
mTcCIR 9	Y16983	NED	acc atg ctt cct cct tca aca ttt ata ccc caa cca	51	$(ct)_{8}n_{15}(ct)_{5}$	108-170
mTcCIR 12	Y16986	NED	tct gac ccc aaa cct gta att cca gtt aaa gca cat	46	$(cata)_{4}n_{18}(tg)_{16}$	224-325
mTcCIR 15	Y16988	HEX	cag ccg cct ctt gt tag tat ttg gga ttc ttg atg	46	$(tc)_{19}$	321-370
mTcCIR 17	Y16990	HEX	aag gat gaa gga tgt aag aga g ccc ata cga gct gtg agt	51	$(gt)_{7}n_{4}(ga)_{12}$	170-260
mTcCIR 18	Y16991	6-FAM	gat agc taa ggg gat tga gga ggt aat tca atc att tga gga ta	51	$(ga)_{12}$	218-260
mTcCIR 19	Y16992	HEX	cac aac ccg tgc tga tta gtt gtt gag gtt gtt agg ag	46	$(ct)_{28}$	130-200
mTcCIR 21	Y16994	6-FAM	gtc gtt gtt gat gtc ggt ggt gag tgt gtg tgt ttg tct	46	$(tc)_{11}n_{5}(ca)_{12}$	176-210
mTcCIR 24	Y16996	HEX	ttt ggg gtg att tct tct ga tct gtc tcg tct ttt ggt ga	46	$(ag)_{13}$	250-320
mTcCIR 25	Y16997	NED	ctt cgt agt gaa tgt agg ag tta ggt agg tag ggt tat ct	46	$(ct)_{21}$	185-290
mTcCIR 26	Y16998	NED	gca ttc atc aat aca ttc gca ctc aaa gtt cat act ac	46	$(tc)_{9}c(ct)_{4}tt$	250-300

Dye colour: FAM= Blue, HEX=Green, NED=Yellow.

migration of bands intensity in comparison to a marker ladder.

DNA amplification (polymerase chain reaction)

Isolated DNA of 270 samples in total was amplified using 13 cocoa SSR primers (Lanaud et al., 1999). Table 1 presents the list of primers with their Dye (colour), Forward and Reverse Primer sequences, annealing temperature, nucleotide repeats, allele size and number of alleles. The forward primer for each of the marker was labelled at the 5' end of the oligonucleotide with one of the following fluorescent dyes: (HEX) 4,7,2',4',5',7'-hexachloro-6-carboxyfluo-rescein, (6-FAM) 6-carboxyfluorescein and (NED) 7',8'-benzo 5'-fluoro-2',4,7 trichloro-3-carboxyflourescin. The total volume of PCR was 5 µl that contained 1 µl of freshly extracted DNA (2.5 ng concentration), 1 µl of 5x PCR buffer (10 mM Tris-HCl pH 8.3, 50 mM KCl), 1 µl of 25 mM $MgCl_2$, 0.5 µl of each forward and reverse primers, 0.2 µl of 10 mM dNTPs (dATP, dCTP, dGTP, dTTP) and 0.1 µl of 5 U Taq polymerase (Bioline, UK). Amplifications were carried out in a Gradient Cycler PTC 200 (MJ Research, USA). The PCR cycle consisting of initial denaturing at 94°C for 5 min followed by 35 cycles of 94°C for 30 s, 51°C annealing for 1 min and 72°C for 1 min. This was followed by further primer extension at 72°C for 7 min. PCR products were then stored at 4°C, prior to visualisation by an automated capillary electrophoresis.

Capillary electrophoresis and data collection

DNA fragment analysis (separation of amplification products) was achieved by capillary electrophoresis in sixteen capillary arrays (36 cm each) of the ABI PRISM 3100 Genetic Analyser machine (Applied Biosystems, USA), and containing POP 4 (Performance Optimised Polymer) matrix and urea. A master mix consisting of 0.5 µl of diluted amplification product (1:19), 9.5 µl formamide (Hi-Di) and 0.5 µl Rox 500 size standard (Applied Biosystems Inc., USA) was prepared for DNA fragments separation during electrophoresis. The master mix was denatured at 95°C for 5 min and stored at 4°C. A modified run module, which consisted of a 60ºC run temperature, 10 s injection time, 15 kV run voltage and 20 min run time, was used for all samples. Each PCR sample was separated with 0.1 µl of an internal-lane size standard (GeneScan - 500 Rox Standard, Applied Biosystems, USA).

The data obtained were analyzed using GENESCAN software v3.7 (Applied Biosystems Inc., USA). Fragment sizes were automatically calculated to two decimal places using the "local Southern" Method algorithm with reference to the internal-lane size

Table 2. Genetic parameters for SSRs markers: number of alleles per locus, observed (Hobs) and expected (H_E) heterozygosity, PIC (polymorphic information content), F-Statistics (FIT, FIS, FST), Hardy-Weinberg (HW) probabilities, et Shannon genotypic diversity index.

Primer	Number of alleles	Hobs	H_E	PIC	FIT	FIS	FST	HW Prob.	Shannon Index
mTcCIR1	21	0.715	0.786	0.757	0.095	0.151	0.147	0.7742	1.90
mTcCIR3	35	0.745	0.877	0.866	0.153	0.143	0.150	0.0000	2.66
mTcCIR6	13	0.535	0.577	0.527	0.079	0.150	0.144	0.8976	1.23
mTcCIR9	25	0.385	0.722	0.680	0.478	0.129	0.130	0.0000	1.82
mTcCIR12	31	0.728	0.890	0.879	0.187	0.142	0.147	0.2328	2.51
mTcCIR15	22	0.478	0.794	0.778	0.404	0.123	0.142	0.0000	2.23
mTcCIR17	19	0.602	0.773	0.747	0.230	0.142	0.142	0.7970	1.81
mTcCIR18	15	0.479	0.726	0.692	0.358	0.147	0.124	0.0000	1.66
mTcCIR19	21	0.471	0.724	0.700	0.358	0.131	0.140	0.0000	1.91
mTcCIR21	24	0.532	0.617	0.598	0.140	0.143	0.147	0.7011	1.63
mTcCIR24	13	0.174	0.246	0.237	0.294	0.136	0.144	0.0008	0.68
mTcCIR25	25	0.697	0.808	0.783	0.147	0.152	0.141	0.0611	2.03
mTcCIR26	18	0.414	0.768	0.734	0.471	0.127	0.132	0.0000	1.81

standard ranging from 35-500 bp (base-pairs). All the 282 SSRs allele sizes obtained were estimated by using 'GeneMapper' software v3.5 (Applied Biosystems Inc., USA) after defining the marker category. This produced a single or pair of peaks, as expected for the co-dominant markers such as microsatellites.

Statistical analysis

The analyses were based on general population genetics principles and approaches, including analyses of gene diversity, genetic differentiation and genetic similarities. With GENETIX 4.05 and F-STAT softwares, the following statistical parameters were used to assess the genetic diversity within and among loci and among the 17 populations (farm selections and clones) under study: allele frequencies; allelic richness; mean number of alleles per polymorphic locus; effective number of alleles per locus; mean observed and expected heterozygosity (H_o and H_e), percentage of polymorphic loci and average gene diversity (Nei, 1978). Bootstrap and Jacknife (Efron, 1982) were used to estimate standard deviations for these diversity parameters with confidence intervals at 95 and 99% levels.

Genetic differentiation was quantified by the F-statistics estimator (Fst), as described by Weir and Cockerham (1984), using Genetix 4.05 software. Fst values were estimated per allele, per locus and overall. GENETIX 4.05 performs bootstrapping (Efron, 1982) over loci and provides rigorous testing of hypotheses of genetic differentiation. In addition, the inbreeding coefficient (FIS) values referred as the proportion of the variance of a genotype within each population were also obtained with GENETIX 4.05.

The genetic similarity (gs) between individual accessions was calculated based on Nei's (1978) unbiased genetic distance. The genetic similarity and distance matrix were obtained using GENETIX 4.05 and TFPGA 1.3 software. Cluster Analysis was performed considering all 270 genotypes (clones and farmers' accessions) individually, and a dendrogram was built on TFPGA 1.3 using the UPGMA (Unweighted Pair Group Method with Arithmetic mean) algorithmic method of clustering (Sokal and Michener, 1954). The distribution of the SSRs data was analyzed using the Factorial Analysis of Correspondences (FAC, GENETIX 4.05) generating three dimensional-based graphs to project the diversity obtained with all variables.

RESULTS

Genetic diversity at loci (markers) level (for all accessions and clones together)

Table 2 presents the values of the genetic parameters used for polymorphism analysis of the 13 SSRs markers used on all the 263 DNA samples of farm selections and clones. The average number of alleles per locus for the 13 markers is 21.19. The highest number of alleles is detected in the mTcCIR3 locus (35 alleles) and the lowest number in the mTcCIR6 and mTcCIR24 loci (13 alleles). A heterozygote deficit (H_E>H_O) is registered for all the markers used.

FIT values of different loci vary between 0.09 and 0.48. The inbreeding effect (relative lack of heterozygosity) in the entire farm population has been estimated by the FIS values for each locus. It varies between 0.127 and 0.151, and indicates consequently a variable rate of self-mating within the entire population studied. FST values for the entire population indicate a low variation between loci (0.12-0.16).

Mean values of Hardy-Weinberg probabilities (Haldane, 1954) for all the entire population have been estimated too for each locus. A deviation from the Hardy-Weinberg equilibrium is observed for most of the loci, except for mTcCIR1, mTcCIR6, mTcCIR17, mTcCIR21 and mTcCIR25. Therefore, the probability of transmission of one allele from one generation to another in all the farmers and breeders populations of the study is under the influence of evolution forces. Hardy-Weinberg equilibrium is not respected for most of the loci; measurement of heterozygosity at each locus is therefore not a suitable approach. Thus, Shannon diversity index that measures the genotypic diversity (instead of allelic diversity) was used (Table 2). From one locus to another, Shannon diversity index values are variable (0.68-2.66), this varia-

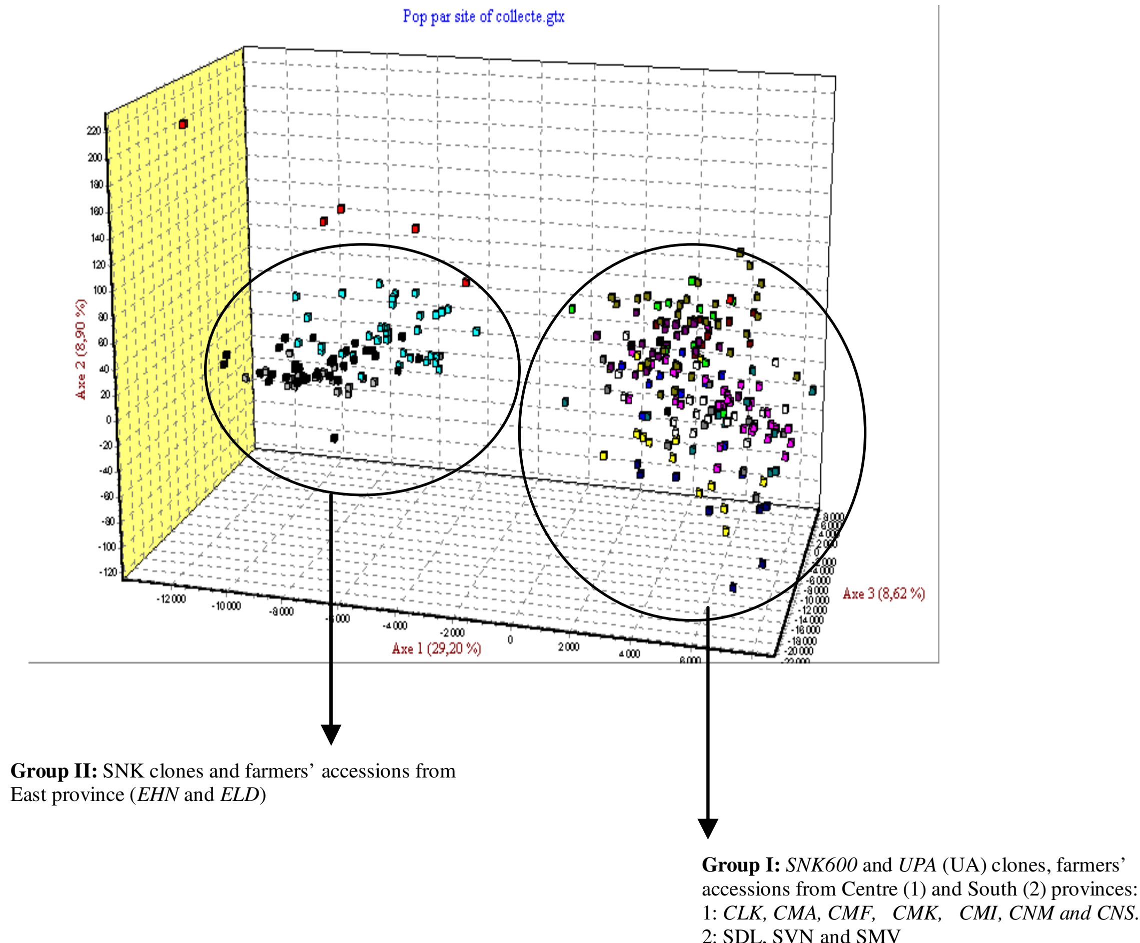

Figure 2. Factorial analysis of correspondences (FAC). Symbols with the same colour represent one geographic group of accessions, or individuals of the same type of breeders' clones (SNK, SNK600 or UPA). T-clones (UA) are represented separately by the symbols in red colour.

tion indicates a low level of diversity within the entire population studied.

Genetic variability among populations

A factorial analysis of correspondences (FAC) distributed all accessions into two main groups (Figure 2 and Table 3). Table 3 presents the values of calculated parameters from the FAC for each geographical population. The expected heterozygosity (H_E) obtained is variable for all populations. Population from Nyong-et-Mfoumou (CNM) division recorded the highest value. Among the breeders' accessions, the UA clones (UPA and T-clones) presented the highest H_E value (0.66). Among accessions, the observed heterozygosity (H_o) varies from 0.40 to 0.72. In terms of deficit of heterozygotes ($H_E < H_o$), only accessions collected in the Nyong-et-So'o Division registered a higher number of heterozygotes than expected.

However, the proportion of polymorphic loci obtained (92-100%) indicated that all the primers used were able to detect an allelic variation among accessions. The average number of alleles varied from one population to another. This difference was due to the variation of sample size of each population. The number of effective alleles per locus (alleles for which the frequency is greater than 0.05) varied from 2.84 to 5.30, and from 3.38 to 3.53 among farm accessions and breeders' accessions respectively.

Diversity structure: hierarchical classification of populations

The structure of the diversity obtained from the factorial analysis of correspondences is confirmed by the dendogram (Figure 3). The dendogram based on Nei's genetic distances (1978) clusters in 2 groups as follows:

- The first group is constituted by accessions from Centre and South provinces together with the SNK600 hybrid group.

Table 3. Diversity based on genetic parameters of FAC constituted groups. FAC group I: Farm accessions collected in the South and Centre provinces, and breeders accessions (SNK 600 and UA clones). FAC group II: Farm accessions collected in the East province and gene bank material (SNK and Upper Amazon Forastero: UA).

FAC group I							
Population	**Collection site**	**H_E**	**H n.b.**	**H_o**	**P(0.95)**	**Mean A/L**	**Effective A/L**
Farm accessions							
CLK	Lékie	0.65(0.20)[1]	0.67(0.21)	0.66(0.21)	1.00	6.53	3.92
CMA	Mefou-et-Akono	0.63(0.14)	0.66(0.15)	0.51(0.15)	1.00	5.53	3.53
CMF	Mefou-et-Afamba	0.68(0.16)	0.70(0.17)	0.65(0.20)	1.00	7.61	4.76
CMI	Mbam-et-Inoubou	0.62(0.22)	0.66(0.23)	0.53(0.32)	0.92	5.15	4.53
CMK	Mbam-et-Kim	0.60(0.21)	0.61(0.22)	0.51(0.25)	1.00	9.53	3.61
CNM	Nyong-et-Mfoumou	0.69(0.12)	0.72(0.12)	0.63(0.20)	1.00	6.00	4.46
CNS	Nyong-et-So'o	0.68(0.17)	0.74(0.18)	0.72(0.16)	1.00	5.30	5.30
SDL	Dja-et-Lobo	0.47(0.20)	0.50(0.21)	0.49(0.29)	1.00	4.23	4.07
SMV	Mvila	0.63(0.17)	0.66(0.18)	0.58(0.19)	1.00	6.76	4.30
SVN	Vallée du Ntem	0.57(0.21)	0.58(0.22)	0.54(0.27)	1.00	6.46	4.23
Genebank material							
SNK600	Nkoemvone	0.64(0.15)	0.65(0.15)	0.59(0.21)	1.00	9.00	3.53
UA (UPA)	Nkoemvone	0.66 (0.17)	0.70(0.18)	0.58(0.36)	1.00	5.23	3.43
FAC group II							
Population	**Collection site**	**H_E**	**H n.b.**	**H_o**	**P(0.95)**	**Mean A/L**	**Effective A/L**
Farmers accessions							
EHN	Haut-Nyong	0.34(0.24)	0.35(0.24)	0.16(0.16)	0.92	4.07	2.84
ELD	Lom-et-Djerem	0.61(0.21)	0.62(0.21)	0.55(0.18)	1.00	7.85	3.92
Gene bank material							
SNK	Nkoemvone	0.61(0.14)	0.61(0.14)	0.47(0.22)	1.00	7.38	3.38

H_E: Expected Heterozygosity H n.b.: non biased heterozygosity H_o:observed heterozygosity
P(0.95): proportion of polymorphic loci when the most frequent allele do not exceed 95%
Mean A/L: Mean number of alleles per locus effective A/L: Effective number of alleles per locus
[1]: standard deviation in parentheses

- The second group is constituted by accessions from East province together with local Trinitario clones (SNK).

DISCUSSION

Two main groups exist in farmers populations. One group (Group II) is made up of near Amelonado types (ELD, because most homozygous) and local Trinitario (because near to SNK types). It is interesting to note that the Amelonado types are still quite heterozygous with the SSR markers used. Such may be unexpected from their self-compatible mating type. It may mean that this population has still considerable variation that can be used in selection (such as for disease resistance or yield). Group II accessions are quite far apart from the T-clones, which could better be considered as a separate group. The fact that the T-clones are separate would indicate that they have not been used in the distribution of hybrid seed.

Group I is made up of a large hybrid swarm. Accessions from 3 divisions (SDL, SVN, SMV) are near to the SNK 600 group, which would indicate that farms in these divisions have used a lot of quite “pure” hybrid seed progenies from the seed gardens. The accessions in the other divisions belonging to Group I is probably made up of divers hybrid genotypes, that may have resulted from open pollination between traditional and hybrid genotypes. The UPA clones are grouped in Group I. This is likely due to the fact that the UPA seed garden parents have been widely used in the distribution of seed to farmers.

It is surprising that apparently no pure Trinitario or Amelonado types have been identified in the Centre and South provinces. This would indicate that intensive hybridisation has taken place between hybrid genotypes and local types, and that there has been a tendency for local types to disappear. This would warrant further studies and separate analyses of diversity within this group.

The T-clones are far away from the centre of Group I and Group II accessions. This means that there has been little uptake of the genes from these clones in farmers’ fields. This is rather unfortunate, as these clones appear

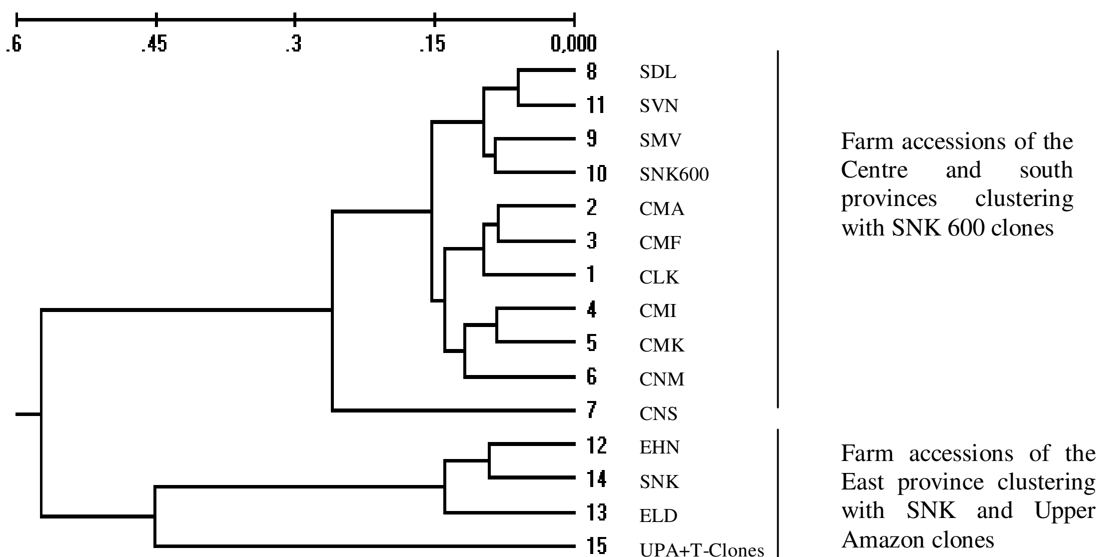

Figure 3. UPGMA Dendogram of farm population and Genebank material based on Nei distances (1978).

to harbour resistance to black pod disease.

The FAC clustered all the farm accessions into two main genetic groups based on their geographical origin. Ten out of 12 geographical populations of farm accessions have shown a level of genetic diversity similar to the ones observed in gene bank on-station. Only one less diverse and highly homozygote population (ELD) has been identified, which is likely to be constituted by Amelonado-type of accessions. The FAC analysis also shows a clear separation of Upper Amazon Forastero T-clones from all the other genotypes. However, these genotypes were used as parents of the released selected hybrid varieties. This suggests that the progenies of these parents were very poorly used in cocoa farms surveyed. This is rather unfortunate since these clones are known to be quite resistant to black pod disease (Nyassé et al., 2003a; Nyassé et al., 2003b). Anyhow, during the last five years, an effort has been made to distribute progenies issued from these T-clones.

All the SSRs markers used in this study, which had been developed at CIRAD by Lanaud et al. (1999), appeared to be highly polymorphic (282 alleles detected in total). For example, the expectedly more homozygous and self-compatible Amelonado population (ELD) appears still as quite heterozygous with these SSR markers. This may also indicate that this population is not as pure as one might think with Amelonado types.

The genetic parameters used show the existence of a narrow genetic relationship between certain populations of farm accessions and certain groups of breeders' clones. For example, the SNK clones (Trinitario types) are genetically close to EHN population of accessions collected in the East province. This would suggest that the farm accessions in the EHN Division are still largely made up of Trinitario types, and that no hybrids were distributed in this Division. Despite the variable sampling size of different populations, the average number of alleles per locus of all genotypes studied is similar to the results obtained in previous cocoa diversity studies in Central America (Motamayor, 2001; Motamayor et al., 2002).

In conclusion, this study has revealed the existence of a genetic diversity base on geographical origin of the accessions. The values of genetic parameters obtained in this study, mainly the values of heterozygosity indicate that the genetic basis of cultivated cocoa in Southern is genetically narrow. However, the allelic richness (282 alleles in total) of the total population suggested the presence of potential unique and rare alleles. This would suggest good possibilities for selection of improved material in farmers' fields. This is therefore in favour of evaluating selections made by farmers for low disease incidence or for high yield. Base on this study, it has been found that some cocoa clones known as resistant material (T-clones), and used in seed gardens in Cameroon were not found to be represented in the farm accessions. The use of this promising material in new breeding program is currently ongoing. In the near future, the molecular data generated will be correlated with information provided by farmers on productivity and black pod resistance, as well as with pod and bean traits collected during the farm survey. The studies will also be extended to the Western parts of Cameroon.

ACKNOWLEDGEMENTS

The authors thank the IITA Genetic Diversity Project for Africa and the CFC/ICCO/IPGRI project titled 'Cocoa

Productivity and Quality Improvement, a Global Approach' for their financial support.

REFERENCES

Alverson WS, Whitlock BA, Nyffeler R, Bayer C, Baum DA (1999). Phylogeny of the core Malvales: Evidence from ndhF sequence data. Am. J. Bot. 86: 1474-1486.

Bhattacharjee R, Kolesnikova-Allen M, Aikpokpodion P, Taiwo S, Ingelbrecht I (2004b). A semi-automated rapid method of extracting genomic DNA for molecular marker analysis in cocoa, Theobroma cacao L. Plant Mol. Biol. Report 22: 435a-435h.

Despréaux D, Clémént D, Partiot M (1989). La pourriture brune des cabosses du cacaoyer au Cameroun : mise en évidence d'un caractère de résistance au champ. Agronomie 9: 683-691.

Efron B (1982). The Jackknife, the bootstrap and other re-sampling plans. NSF-CBMS Regional Conference Series in Applied Mathematics, Monograph 38. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, Pennsylvania, USA.

Haldane JBS (1954). "The Origin of Life", New Biology 16, 12.

ICCO (2004). International Cocoa Organisation. Annual Report 2004. p. 32.

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

Motamayor JC (2001). Etude de la diversité génétique et de la domestication des cacaoyers du groupe Criollo (Theobroma cacao L.) à l'aide de marqueurs moléculaires. Ph D Thesis. Université Paris XI. p. 179.

Motamayor JC, Risterucci AM, Lopez PA, Lanaud C (2002). Cacao domestication I: the origin of the cacao cultivated by the Mayas. Heredity 89: 380-386.

Nei M (1978). Estimation of average heterozygosity and genetic distance from a small number of individuals. Genetics 89: 583–590.

Nya N'gatchou J (1981). Etat d'avancement des travaux de génétique et d'amélioration du cacaoyer au Cameroun. VIIth Intern. Cocoa Rech. Conf., Douala, Cameroon. pp. 507-511.

Nyassé S, Efombagn MIB, Bouambi E, Ndoumbè-Nkeng M, Eskes AB (2003a). Early selection for resistance to Phytophthora megakarya in local and introduced cocoa varieties in Cameroon. Trop. Sci. 43: 96-102.

Nyassé S, Efombagn MIB, Eskes AB (2003b). Selection for resistance to black pod and yield gains prediction by use of selected cocoa varieties in Cameroon. Plant Genetic Resour. 1: 157-160.

Sokal, R.R. and Michener, C.D. (1958) A statistical method for evaluating systematic relationships. University of Kansas Scientific Bull. 28: 1409-1438.

Weir BS, Cockerham CC (1984). Estimating F-statistics for the analysis of population structure. Evolution 38: 1358-137.