International Scholars Journals

International Journal of Plant Breeding and Genetics ISSN: 3167-5891 Vol. 3 (1), pp. 166-176, June, 2016. Available online at www.internationalscholarsjournals.org © International Scholars Journals

Full Length Research Paper

Combined use of parentage analysis and phenotypic assessment to evaluate the performances of cocoa (*Theobroma cacao* L.) varieties in farmers' field

Olivier Sounigo[1,2], Fidèle Nsouga[2], Olivier Fouet[3], Claire Lanaud[3], Ives Bruno Efombagn Mousseni[2]*

[1]CIRAD, UR Bioagresseurs, BP 2572 Elig-Essono, Yaoundé, Cameroun.
[2]IRAD P0 Box 2123, Yaoundé, Cameroun.
[3]CIRAD, UMR AGAP, TA A108/3, Avenue Agropolis, 34398 Montpellier, France.

Accepted 19 June, 2016

The genetic diversity of a cocoa population composed of 264 adult trees grown in anon farm plantation was assessed using 125 microsatellite markers to understand the relationship between the origin of the cultivated material and its agronomic features in the field. Analysis using a Bayesian model-based method allowed us to divide the cocoa trees into eight sub-populations (clusters) totaling 456 alleles with a coefficient of differentiation, F_{ST} = 0.1. Genealogical classes were generated within the entire population, and allowed to clearly identify 157 cocoa trees (59.46% of the total population) as derived from crosses between the two parental clones represented in bi-clonal seed-gardens, and were considered as true improved cocoa varieties while the genetic origin of 107 cocoa trees failed to be clearly identified. The parentage analysis carried out on the 157 trees identified as true F1-hybrids revealed that they belong to seven full-sib different progenies, derived from seven different bi-clonal seed-gardens. A comparison of the agronomical performances of three of the identified full-seed progenies showed differences for traits such as yield and susceptibility to pod rot caused by *Phytophthora megakarya*.

Keywords: Cacao, genetic structure, full-sib progenies, parentage analysis, morphological evaluation.

INTRODUCTION

Cacao (*Theobroma cacao* L.) is a diploid (Cope, 1984) preferentially allogamous tropical tree species, belonging to the family Malvaceae (Alverson et al., 1999). It is an important crop with industrial and nutritional values, and constitutes the principal source of raw material for making chocolate. It plays a major role in Africa, which hosts four of the five most cocoa producing countries worldwide, with Cameroon ranking 5th in production with an annual cocoa production neighboring 200,000 tons (ICCO, 2015).

Cocoa was first introduced in Cameroon between 1876 and 1886 (Ardener, 1996, Bartley, 2005). After the first breeding program implemented in the 1950s, based on clonal selection, a new program was initiated in the early 1960s, based on the selection of high yielding full-sib progenies, derived from crosses between the local clones selected in the 1950s and clones selected in Trinidad and Ghana, introduced in Cameroon in the late 1950s (Liabeuf, 1962).

This program consisted of the creation of around 300 full-sib progenies and the assessment of their yield in trial plots in the research stations in Nkoemvone (South Cameroon) and in Barombi-kang (South-West Cameroon). This program resulted in the selection of 24 full-sib progenies, released to farmers through the establishment of 36 bi-clonal seed gardens (plots estab-

*Corresponding author. E-mail: efombagn@yahoo.fr

lished with cuttings representing the 2 parents) in three sites in South, East, South West and Centre regions of Cameroon, between 1960 and 2002.

Unfortunately, no further assessment of the performances of the released improved cocoa varieties has been made after they were established in cocoa farms, except few farmers' interviews revealing that these improved varieties were generally appreciated for their level of precocity and yield, but were disliked for their high level of susceptibility to pod rot caused by *Phytophthora megakarya* (Efombagn et al., 2007). However, the main problem encountered with these cultivated varieties is the high variation in agronomic traits' of cocoa trees within a single plantation. This might be due to the mixture of seedlings from different progenies during the establishment of a plantation or replacement of dead cocoa trees, as well as the high heterozygosity that is commonly found in clonal hybrids in cocoa producing countries.

One of the major constraints faced by cacao breeding programs is the poor availability of uniform genetic background of the material to constitute the germplasm serving as working collections. In Cameroon for instance, most of the genotypes used as parental clones in varietal selection harbour a great level of heterozygosity in the genome. The progenies obtained are consequently considered as clonal hybrids rather than true F1-Hybrids with all the individuals being genotypically uniform. The direct consequence is the high level of outcrossing pattern usually exhibited on-farm as shown in Côte d'Ivoire by Pokou et al., 2014. Efombagn et al. (2008) have also shown that the cultivated cacao populations generally display a diversified structure, with a relatively high rate of genetic diversity. A great genetic variation within cacao progenies may impact significantly on its agronomic performance, specifically with yield components (Aikpokpodion et al., 2010; Cassia et al., 2013) and PPR resistance (Barreto et al., 2015) controlled by several genes.

In order to get information on the performance of the released cocoa varieties, a participatory breeding cocoa selection was initiated in Cameroon in 2005, consisting in releasing pods from the bi-clonal seed gardens to farmers, after identification of the genetic origin of these pods and following the performance of the derived progenies in on farm plots (Sounigo et al., 2012).

In addition, molecular and phenotypic data obtained from a linkage disequilibrium study (Sounigo, data not published) conducted on twenty year old cocoa trees in an on farm cocoa trial were used in an alternative approach for assessing on farm performances of the improved cocoa varieties. The molecular data of the cocoa trees in the field was compared to the one from the parental clones of the bi-clonal seed gardens used to produce at large scale the released progeny and parentage analyses were performed in order to identify the pedigree of the assessed cocoa trees. The performances of the identified full-sib progenies were then compared.

The present paper describes the genetic diversity of the whole set of cocoa trees studied and the agronomical performances observed within the true full-sib progenies identified in the on farm cocoa plot used for our study.

MATERIALS AND METHODS

Plant material and growing conditions

The study was conducted on a twenty-year old on farm plot located near the city of Yaoundé, in the central region of Cameroon.

This plot was set up using more than 3,000 seedlings derived from pods commercially released to farmers. These pods were harvested in bi-clonal seed gardens after natural pollination. No hand-pollination was performed in the bi-clonal seed-gardens of Cameroon in the early 1980s since the female parental trees had been selected for their self-incompatibility and the seeds derived from natural pollination were assumed to originate exclusively from crossing with the male parental trees present in the bi-clonal seed-gardens. However, a study made by Lanaud et al. (1987) showed that the self-incompatibility of the female parental trees failed to prevent selfing from occurring under natural pollination. The cocoa trees in our plot were assumed to be a mixture of trees derived from crosses and selfing.

The cocoa trees were cultivated under permanent shade provided by forest trees and no fungicide treatment was performed during the study, these two conditions being assumed to favor a high level of infection by *Phytophthora megakarya*, desirable for assessment of cocoa trees for tolerance to this disease.

Counting of pods was made on all the 3,000 cocoa trees at the beginning of the main harvest season (end of August 2005), and the 260 cacao trees with a high number of pods were selected for this study, in order to ensure a reliable assessment of certain traits, such as; tolerance to black pod disease, caused by *Phytophthora megakarya,* bean and pod traits, two traits that require a proper sample size.

METHODS

Genotyping

Since the assessed cocoa trees were also used for a linkage disequilibrium study aiming at identifying genome areas involved in several traits of interest (Sounigo, data not published), the set of markers selected for this study were those found to be close to PPR resistance QTLs identified by a previous meta QTL analysis and distributed along the ten chromosomes of the cacao genome in a consensus map (Lanaud et al., 2009). DNA was extracted from fresh leaves harvested from the 260 selected cacao

trees. 125 SSRs markers described by Lanaud et al. (1999) and Fouet et al. (2011) were used to genotype the corresponding samples. The position of these markers along the cacao linkage groups was presented in a reference map established recently by Allegre et al. (2012). After the extraction, DNA fragments were amplified and genotyped in the fluorescent PCR products and analyzed on Mega BACE™ 1000 Sequencer (Amersham Biosciences) as described by Fouet et al. (2011). Genotyping was carried out by the National Genotyping Center (CNG) in France following the protocol described by Fouet et al. (2011).

Analyses of markers

GenAIEx v 6.5 (Peakall and Smouse,2012) was used to determine the sample size, number of alleles, number of effective alleles, F_{ST} values, observed heterozygosity and expected heterozygosity for each of the 125 markers.

Population structure

The software STRUCTURE ver.2.3.3 (Falush et al., 2003; Pritchard et al., 2000) was ran to identify the genetic structure and the number of sub-populations among the 264 accessions. Analyses were ran using a burn-in period of 100,000 and MCMC (Monte Carlo Markov Chain) reactions of 200,000. The admixture and correlated option were used for the ancestry and allele frequency models. While running STRUCTURE, the hypothesis of two (k=2) to eight (k=8) subpopulations was considered. Individuals were assigned to a subpopulation only when a cluster membership probability was higher than 70%.

Identification of genealogical classes among accessions

The microsatellite data of the 264 accessions (F1 hybrids, admixed or selfed genotypes) were subjected along with bi-clonal seed gardens (BSG) progenitors to the analysis of hybrid ancestry using the method implemented in the program NEWHYBRIDS 1.0 (Anderson and Thomson, 2002). STRUCTURE can identify admixtures among any number K of parental populations, while NEWHYBRIDS assumes that hybrid classes originated after admixture of two parental species or parental populations (genetic groups of putative progenitors in our study) within species. NEWHYBRIDS uses an inheritance model defined in terms of genotype frequencies to compute for each individual, the posterior probability of inclusion in each of the following six classes: P1 (Female's BSG progenitors in our study), P2 (male's BSG progenitors), selfed P1 or selfed P2 genotypes (S1), F1, F2 and back cross (BC) genotypes between F1 and each of the two genetic groups of progenitors (P1 and P2). NEWHYBRIDS was run using the default parameters for the six genotype class frequencies, Jeffreys prior, a burn-in phase of 100 000 steps and 500 000 MCMC sweeps. The five Amelonado controls were subsequently used as a separate parental class to simulate the reality in farmers' field where genotypes may have originated from outcrossing between traditional and BSG-derived progenies.

Parentage analysis

Parentage analysis was performed only on the F1 hybrid accessions identified by NEWHYBRIDS. We used the software CERVUS 3.0 (Marshall et al., 1998) to detect the putative BSG progenitors of the F1 accessions, as well as of the selfed S1 accessions per candidate progenitor. CERVUS uses a simulation program to generate log-likelihood scores and provides a level of statistical confidence to assign paternity, maternity or both male and female parents. Parent pair analysis used in this study is the common parentage analysis when the candidate male and female parents are known (Marshall et al., 1998). The log-likelihood ratios were expressed as LOD scores, which are the natural logarithm of likelihood ratios products at each locus. The most likely progenitor was the one with the highest positive LOD Score. The CERVUS programme uses furthermore the number of mismatching loci between a candidate parent and a putative offspring for the parental assignment. Parent pair non-exclusion probability (Jamieson and Taylor, 1997; Marshall et al., 1998) for each candidate progenitor was calculated for each microsatellite used. The non-exclusion probability is the probability of not excluding a single unrelated candidate parent or parent pair from parentage of a given offspring at one locus.

Comparison of the genetic and the geographic matrices

The relationship between the genetic and the geographic distances was estimated with the mantel test procedures computed in GenAIEx v 6.5.The genetic distances as well as the geographic distance matrices were calculated and estimates between possible pairs of individual cocoa trees of the study were generated prior to the comparison test. In GenAIEx v 6.5, Mantel tests for Matrix Correspondence (Mantel, 1967) followed the method of Smouse and Long (1992), with the option for statistical testing by random permutation. The Mantel option allows tests for a statistical relationship between the elements of any two distance matrices with matching entries. Typical applications include testing for isolation-by-distance, for which a Nei genetic distance matrix (or log of the genetic distance) is compared with the geographic distance matrix.

Yield components

Tree yield related traits were measured namely; the number

of pods per tree (PodNum), the average weight of one fermented and dried bean (BeanWei) and the number of beans per pod (BeanNum).
Healthy (ripe) and rotten pods (unripe and ripe) were harvested every week during the period from April 2007 to March 2010 and counted, for each tree.
To estimate the average bean weight, the fresh beans from the pods of the same cacao tree were mixed to constitute a sample, fermented during five days and exposed to solar drying. The drying was completed when the amount of water within a bean was below 8%. 100 dried beans were randomly picked up in each sample and weighed.
The average bean number was obtained by averaging the number of beans counted fromfive separate pods harvested from the cocoa tree.

Disease Phenotyping

Incidence of PPR in the field (%PPR)
% of rotten pods was calculated according to the following formula;
%PPR = number of rotten pods / (number of rotten pods + number of healthy pods) * 100

Symptom scoring after artificial inoculation (PodTest)

Artificial inoculation was performed in the lab on mature but unripe harvested pods. The protocol was the one perfected by Iwaro et al. (2000), consisting in spraying a suspension of zoospores from a moderately aggressive isolate of *Phytophthoramegakarya* ($3x10^5$ zoospores/ml) on the pods placed in trays just after the harvest. The trays were then kept in obscurity at room temperature during 5 days.
Symptom assessments were made 5 days after inoculation, based on a scale describedby Iwaro et al. (2000), taking both the frequency and the surface of the lesions caused by *Phytophthoramegakarya* into account.
The number of pods used for artificial inoculation rangedbetween 4 and 20 per cacao tree.

Analyses of variance were carried out with S.A.S. (2000) software version 8 to compare accessions or groups of accessions. The Newman and Keuls (Cochran and Cox, 1957) test at 5% probability level was used to compare means of individual cacao trees and means of progenies.

RESULTS

Population structure

Bayesian clustering implemented within the STRUCTURE v 2.3.2.1 software identified K=7 as the most likely number of subpopulations. The criteria used to define the number of subgroups within the studied population, which are the position of break point in the L(*k*) curve and a peak in the Δk distribution, supported this value of K=7, and the corresponding sub-populations are depicted in Fig.1.

Genetic diversity within the sub-populations revealed by the population structure

Table 1 shows thevalues for diversity parameters for each of the seven sub-populations identified by STRUCTURE. The number of genotyped accessions ranged from 14 (Cluster 2) to 114 (Cluster 6) among the seven subpopulations. A total of 546 alleles were amplified across the accessions. The mean number of alleles across all accessions in each subpopulation ranged from 2.0 to 3.78.The number of alleles per locus across all accessions ranged from 2 to 12, and the number of effective alleles ranged from 1.59 to 2.36, with a mean of 2.06. Mean observed heterozygosity across the markers was 0.56 (ranged between 0.41 and 0.66).Population differentiation among all analyzed accessions was significant with a mean overall F_{ST} value of 0.105 ($p< 0.001$, AMOVA).

Identification of the genealogical structure within the farm population

The genealogical classes of the farm population in relation to the bi-clonal Seed Gardens progenitors used in Cameroon were determined by NEWHYBRIDS. The results suggested that 157out of 264cocoa trees (60% of the total population) issued from crosses between the two parents of the same BSG plots, corresponding to truly improved varieties, noted as F1 in the present paper. 37 cocoa trees (14%) were issued from crosses between two parents, but located in two different BSG plots. Two hypotheses can explain their existence: cross contamination among neighboring BSG plots and accidental introduction of wrong parent trees during the setting-up of the BSG plots, after mislabeling events. 21 trees (8%) were identified as derived from selfing of five clones (SNK16, SNK64, T79/501, ICS 40 and ICS 46) used as parents in the BSG. The genetic origin of the remaining 48 (18%) trees could be interpreted as belonging to either F2 or backcross (BC).
Such trees could originate from natural establishment of seedlings obtained from seeds from the trees originally set up in the plot, once these trees were able to produce fruits, or these seedlings could have been set up by the farmer in order to replace cocoa trees which died few years after plot establishment.
F2 and BC were considered as one group (identified as F2+BC in Table 2) because the probabilities of assignment were rather similar.F2+BC material was found as the likely product of natural pollination in farmers' field between F1 hybrids, or between F1 hybrids and traditional genotypes. The trees derived from the BSG crosses were distributed between four different

Table 1. Population statistics across and within each subpopulation revealed by STRUCTURE and averaged for all the 125 SSRs loci.

Subpop	N	NA	NE	Ho	He	GD	AR	F
1	21	2.776	2.360	0.665	0.513	0.523	2.089	-0.280 (0.017)
2	14	2.776	2.082	0.579	0.436	0.448	1.923	-0.298 (0.025)
3	37	2.888	2.066	0.518	0.421	0.425	1.877	-0.196 (0.019)
4	5	2.000	1.593	0.418	0.314	0.343	1.674	-0.294 (0.024)
5	19	3.128	2.118	0.565	0.441	0.450	1.948	-0.247 (0.019)
6	114	2.480	2.060	0.563	0.427	0.428	1.859	-0.306 (0.022)
7	23	2.624	1.989	0.621	0.463	0.470	1.899	-0.345 (0.018)
NC	31	3.784	2.227	0.553	0.485	0.493	2.033	-0.135 (0.013)
All	264	2.807	2.062	0.560	0.437	0.447	1.913	-0.261 (0.007)

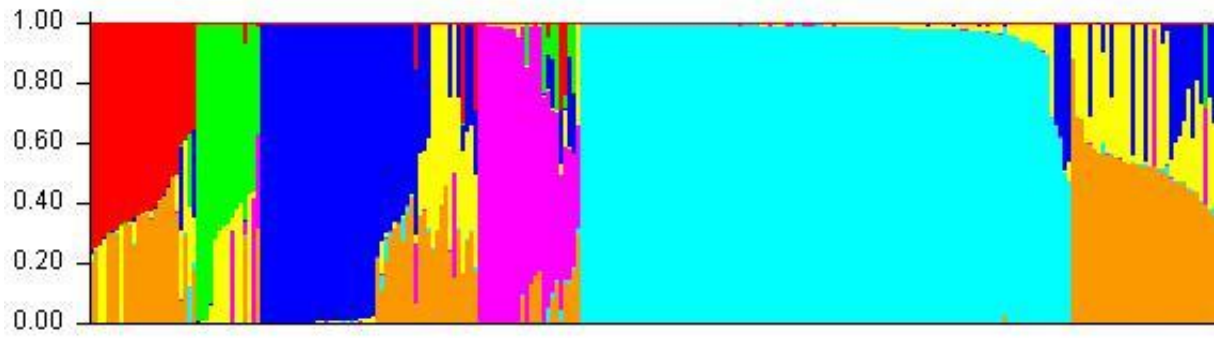

Figure 1. Diversity structure depictured for 264 accessions according to the STRUCTURE program (Pritchard et al. 2000).

clusters, whereas the two other genealogical classes (crosses without fidelity to BSG known combinations and F2/BC genotypes) were scattered in more clusters, including the group of accessions which did not belong to any of these subpopulations45 (NC= No Classified).

Parentage analysis

All the F1 genotypes were subjected to a parentage analysis to detect the true BSG progenitors in the studied population. The 37 F1 hybrid genotypes for which the two parents identified did not correspond to the known combinations of the BSG progenitors represented19.07% of the total F1 hybrid genotypes found in this study. These were therefore derived from unwanted hybridization in the BSG. More than 80% (157 trees) of that genealogical class was true F1 progenies corresponding to the known commercial varieties released from Seed gardens.The true 157 F1 genotypes were assigned to candidate BSG progenitors based on the highest positive LOD score and on the number of mismatched alleles registered (Table2). Seven full-sib progenies were identified, represented by the number of trees ranging between one (case of ICS95*SNK10) and 110 (case of SNK16*T60/1174). All the trees from the same full-sib progeny were assigned to the same cluster generated by STRUCTURE, showing a complete match between the genetic structure and the parentage analysis approaches.

Relationship between the genetic structure and the geographical distribution of the cacao trees

A graph of the relationship between elements from any two matrices was plotted (Fig. 2). A correlation coefficient for both genetic and geographic data matrices yielded a value of 0.011, with a test for a significant relationship by approximately 1,000 random

Table 2. Distribution of the cacao accessions in the different genealogical classes and identification of progenies issued from bi-clonal seed gardens (BSG).

Type of material	Progeny	Quantity of Trees	Corresponding Cluster (STRUCTURE)	Percentage
F1 Hybrids derived from BSG	ICS40*UPA134 SNK109*T79/501 SNK13*T79/501 SNK16*T60/1174 ICS95*SNK10 ICS95*SNK413 ICS95*UPA41 ***Sub total***	16 5 23 110 1 1 2 158	Cl1* Cl3 Cl3 Cl6 Cl5 Cl5 Cl5	60.0%
F1 Hybrids ofunknown origin		37	Cl1, 2, 3, 5, 6, NC	14.0%
Selfed BSG parents		21	Cl 3, 4, 7	8.0%
Other genotypes	F2 +BCs	48	Cl1, 2, 3, 4, 5, 7, NC	18.0%
Total		264		

Cl= Cluster revealed by the software STRUCTURE – **NC**: accessions no attributed to any of the seven clusters.

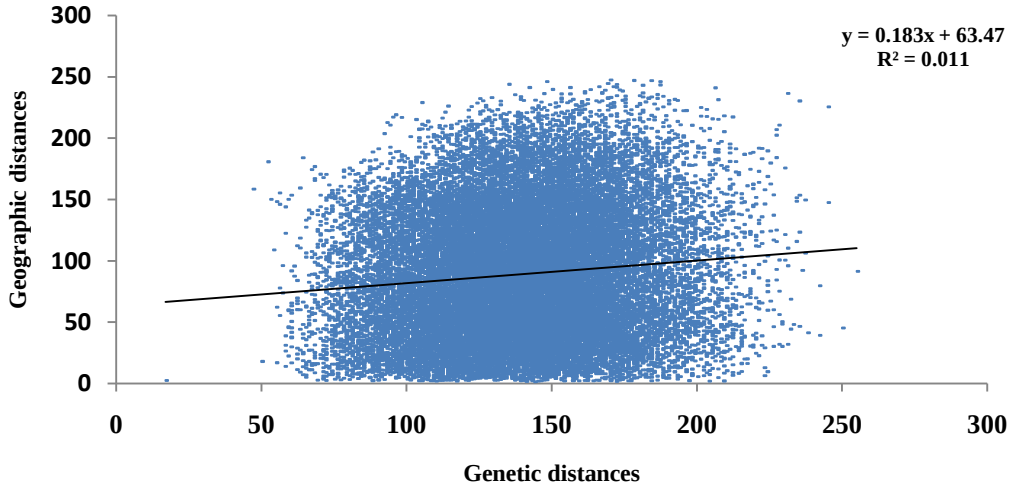

Figure 2. Relationship between the genetic and the geographic distances among the 264 accessions scattered in a single plot.

permutations. The null hypothesis was that there is no significant relationship between the two types of distances. However, a random correlation was not at the extreme (closer to +1 or – 1), indicating that the genetic structure of the cacao population was weakly related to the distribution of the trees within the plot.

Phenotypic evaluation of BSG progenies

The ranking of the five genetic origins represented by more than 10 trees for several traits is presented in Table 3.

The progeny derived from SNK 16 * T 60/1174 appeared as a higher yielder than the other genetic origins. Indeed, the mean cumulated yield of this progeny was estimated at9119 g per tree during the period from April 2007 to March 2010, corresponding to an estimated potential annual yield of 3.377 kg of dried cocoa bean at the recommended planting density of 1,111 cocoa trees/ha.

Table 3. PPR resistance and yield performance observed in the different subgroups of the studied population.

Type of comparison	Genetic origin	Number of Trees	%PPR	PodTest	PodNum	BeanWei	BeanNum	BeanWei/Pod	Optimal average yield	Minimal average yield
among full-progenies issued from BSG	ICS40*UPA134	15	44.8 (*6-76*) b	4 (*2.6-6.0*) a	170.3 (*46-356*)a	1.15 (*0.72-1.67*)ab	33.6(*12.5-45.4*)a	38.9(14.5-60.9)a	6625b	3869a
	SNK13*T79/501	23	62.8 (*33-82*) a	4.6 (*2.7-5.6*) b	184.6 (*27-434*)a	1.07 (*0.75-1.25*)a	35.5 (*29.0-68.1*)a	38.5(25.4-72.9)a	7107b	2500b
	SNK16*T60/1174	110	61.1 (*22-89*) a	4.4 (*2.0-7.0*)b	257.6 (*44-596*)b	1.06 (*0.78-1.66*)a	33.5 (*16.0-49.8*)a	35.4(17.5-71.2)a	9119a	3432a
	Full-sib progenies issued from BSG	158	59.7(6.1-89.4) a	4.5(2.0-7.6) b	233.3(27-596)a	1.1(0.7-1.7)c	33.7(12.5-68.1)a	36.4(14.5-72.9)b	8492a	3306a
among cacao trees from different genetic origins	progenies of Unknown origin	37	59 (11.2-88.7)a	5.3(2.5-6.9) a	203.7(63-506)a	1.02(0.69-1.39)c	34.5(19.5-47.5)a	35(13.5-57.4)b	8074a	3149a
	Progenies issued from selfing of BSG parents	21	60.5 (27.6-84.5) a	5 (2.5-6.9) ab	132.1(40-213)b	1.39(0.75-1.95)a	31.8(20.5-36.4)a	45.4(25.2-56.6.)a	5997c	2764b
	Other genotypes (F2 +BCs)	48	55.0(13.9-88.8) a	.95(2.0-7.6) ab	175.9(17-465)ab	1.20(0.76-1.83)b	35.3(24.0-49.5)a	42.3(22.9-73.1)a	7440b	3435a

%PPR: percentage of PPR-infected pods -**PodTest**: Symptoms scored on infected pods five days after inoculation – **PodNum**: Average number of pods per tree – **BeanWei**: Average bean Weight par tree – **BeanNum**: Average number of beans per pod– **BeanWei/Pod**: Average bean weight (gr.) per pod– optimal average yield: mean weight of cocoa per tree, calculated on the basis of all harvested pods, including the ripe and unripe rotten pods – minimal average yield: mean weight of cocoa per tree, calculated exclusively on the basis of the healthy harvested pods.

On the other hand, the two other progenies (SNK 13 * T 79/501 and ICS 40 * UPA 134) showed yields similar to trees which were not derived from the full-sib progenies officially selected for releasing to farmers. The potential yield values for these genetic origins ranged between 5997 and 8492 g of cocoa bean per tree, corresponding to a potential annual yield ranging between 2221 and 3145 kg/ha.

The lowest level of yield was observed for the trees derived from the selfing of parents represented in the BSG (5997 g/tree, corresponding to 2221 kg/ha).

Of course, one should not forget that the performances of the assessed trees probably did not reflect the average performances of the progenies since the assessed trees were not a randomly chosen sample but were selected on the basis of the number of pods they were bearing in August 2005. Consequently, the values obtained here are expected to be higher than the mean values of the progenies.

The progeny obtained from ICS 40 * UPA 134 showed a lower number of rotten pods (44.8%) than the four other genetic origins (55 to 62.8%). This lower level of susceptibility to black pod disease caused by *Phytophthora megakarya* was confirmed by the lower score observed for this progeny after lab inoculation of the pods (4.0) while this score ranged between 4.4 and 5 in the case of the other progenies.

Taking the % of rotten pods into account, the annual yield of the different

genetic origins ranged between 2500 and 3869 (g of cocoa /tree), corresponding to annual yields of between 926 and 1433 kg/ha.

The two highest yielding progenies were derived from ICS 40 * UPA 134 (1343 kg/ha) and SNK 16 * T 60/1174 (1,311 kg/ha), while the progeny derived from SNK 13 * T 79/501 showed a lower value (965 kg/ha) than the progenies which did not correspond to officially released genetic material (1,094 and 1,114 kg/ha).

Only the trees derived from selfing of parents of BSG were found to yield nearly as poorly as SNK 13 * T 79/501, with a yield of2764 g/tree, corresponding to 1024 kg/ha.

On the other hand, there was no statistical difference between the different genetic origins for the number of beans per pod and for the mean weight of one fermented and dried bean.

DISCUSSION

One of the objectives of our study was to investigate the structure of a cultivated cacao population composed of varieties released from commercial seed gardens. The genetic structure (Fig. 1)confirmed the presence of different sources of planting material within the plot. Efombagn et al. (2008) showed that a high level admixture in most cacao farms in Cameroon, due to hybridization and recombination of these genes from different genetic groups in seed gardens and in farmers' fields. In this study, seven different clusters were identified with 48 accessions (18% of the total population) being unclassified to any of these clusters. However, one cannot be sure that the figures observed for the 264 tree subsamples were also valid for the whole plot since this subsample was not identified randomly but based on the number of pods counted on the trees in September 2005.

It is important to note that the % of trees identified as derived from undesired selfing of parents in the BSG seemed very low (only 8% of the trees), compared to the levels of selfing observed in the BSG of Cameroon and other African countries, in a study undertaken in 1986 (Lanaud et al., 1987).

Among the diversity parameters estimated, the genetic differentiation among the clusters generated by STRUCTURE software appeared relatively low (F_{ST}= 0.1), when compared to wild population on which a F_{ST} of 0.2 was observed in another study (Lachenaud and Dapeng, 2008), this value obtained in our study indicated that the selection process leading to the development of released cacao varieties acted as an evolutionary force to narrow the diversity over time. All the basic values for the genetic parameters were similar to those found in similar cultivated populations studied in Côte d'Ivoire (Pokou et al., 2009;2014; Aikpokpodion et al.,2009; Opoku et al. 2007).

A knowledge of the genetic diversity and population structure of germplasm collections is an important foundation for crop improvement. The large variability in farm accessions can be explained by the large variation of cacao introduced in Cameroon at the beginning of the 20th century, as reported by Bartley (2005).The BSG progenies resulted from crosses between genetically distant parental clones, usually a Trinitario male crossed with an upper/lower Amazon female parent. These types of hybrid varieties have the advantage of revealing some superior tolerant or high yielding trees as a result of the high heterosis generated (Dias et al., 2003). Hybrids' breeding value gives a measure of probable advances for yield components like pods and bean traits as shown by Adewale et al. (2014) in a field experiment aiming at measuring the heterosis in hybrid progenies.

The parentage analysis revealed that at least seven full-sib progenies from bi-clonal seed gardens were released to the farmer for his plot. This number might be higher since we only genotyped a sub-sample of the trees (around 10%) and the sampling method was based on the number of pods produced in August 2005. It is thus possible that full-sib progenies received by the farmer from the extension unit and initially planted were not detected in our study, for several reasons; the farmer could have received only few pods from the parents of these progenies, these progenies were not able to establish properly in the plot or yielded poorly.

These reasons could also explain why the seven progenies from seed-gardens were not equally represented in the plot, the one from SNK 16 * T 60/1174 being largely represented (110 trees) while the six others were represented by numbers of trees ranging between 1 and 23.

No correlation was observed between genetic and geographical distances between the assessed trees, indicating that a mixture of seeds was probably made before planting. The absence of correlation indicates that phenotypic differences observed between the assessed progenies were due to genetic differences rather than to their small scale environment. Indeed, on farm cocoa plots are usually very heterogeneous in Cameroon, especially because of the presence of shade trees belonging to several species, having contrasted impacts on the neighbouring cocoa trees through their influence on the level of shade, on the soil fertility and on the competition for water (Isaac et al., 2007;Sonwa et al., 2007, Nygrenet Leblanc, 2009).

The assessment of the different types of progenies identified from the parentage analysis revealed a rather high level of potential yield (ranging between 2,400 and 3,400 kg of cocoa per year and per ha), the highest value being observed for the progeny derived from SNK 64 * T 60/1174.However, these values do not probably reflect the average yield of these progenies since they were calculated on the basis of the totality of the harvested pods, including unripe and ripe rotten pods and since the trees selected for our study were the ones showing high numbers of pods in August 2005.

The yield values must be considered with caution and it

is not clear if the comparison between the different progenies could be considered as reliable, since the samples size varied greatly among the assessed progenies. Indeed, since the 10% highest yielding trees in August 2005 were selected for the study, the progeny SNK 64 * T 60/1174 was expected to have been advantaged because of its large population size (110 trees). This large level of representation could have resulted from a predominance of this progeny among the seedlings before planting or from a higher level of survival of this progeny in the plot. On the other hand, one cannot rule out the possibility that this progeny is predominant in our sample because it actually yielded more than the others. It would be necessary to genotype another sample of randomly chosen cocoa trees to compare the representation of the different progenies in it with the one observed in our biased sample.

The progeny derived from ICS 40 * UPA 134 showed a lower level of susceptibility to pod rot caused by *Phytophthora megakarya*, when this trait was measured through the % of rotten pods (44.8% for this progeny versus 56.8 to 62.8% for the other progenies) and when it was measured through the scores obtained after artificial pod inoculation (4 versus 4.4 to 5.3 for the other progenies).

Considering the rate of rotten pods for the calculation of yields, the two progenies derived from ICS 40 * UPA 134 and SNK 64 * T 60/1174were better performing (3869 g/tree corresponding to 1343 kg/ha and 3432 g/tree corresponding to 1311 kg of dried cocoa /ha/year, respectively) than the other progenies (2500 g/tree, corresponding to 965 kg/ha to 3435 g/tree, corresponding to 1114 kg dried cocoa /ha).These differences in yields between released improved cocoa varieties were also observed in on farm progeny trials set up in Cameroon since 2006 (Sounigo et al., 2012) with the best performing variety yielding three times more than the poorest performing one. Unfortunately, no data was available yet from these trials for the varieties assessed in the present study, except for the one derived from SNK 13 * T 79/501, which appeared to yield moderately also in the on farm trial plot.

Our data indicated that two out of the three released improved varieties performed better than progenies not obtained from the seed-gardens. Similar results were observed in on-farm trial plots set up in 2007 (Sounigo et al., 2012) where only one out of the two varieties from bi clonal seed gardens under assessment performed better than unselected progenies obtained from pods harvested in farmers' plots.

Our data also indicated a low level of yield for the trees derived from selfing of the parents of the BSG. This confirmed the results from a study conducted in Côte d'Ivoire by N'goran et al. (2003) who found a difference of about30% between the yield of full-sib progenies and trees derived from selfing of the same female parent. This confirmed the need for performing hand-pollination in these plots for progeny release (Bastide and Sounigo, 1993) which dramatically reduced the rate of selfing.

On the other hand, the trees obtained from unknown origin and from open-pollination of the released varieties showed yields not significantly different from the full-sib progenies.

In this study, the trees inside this category were probably derived from pods harvested from the released of full-sib progenies inside the plot, or from spontaneous germination of seeds or voluntary planting by the farmer. The rather good performances of this type of trees was however good news since the use of pods harvested in commercially established cocoa plots has been a common practice in Cameroon during the last twenty years, whether to establish new cocoa plots or for replacing dead cocoa trees in old plots (Efombagn et al., 2008).

The progeny ICS40*UPA134 which is rather susceptible to PPR was more tolerant in the cacao plantation selected for this study. This result may suggest environmental effects as usually observed in farmers' field. The Mantel test revealed that, the correlation between the geographic and the genetic structures within the plot was very weak. Therefore, the cacao trees of the same progeny were scattered over several parts of the plot, and therefore subjected to micro environmental variations mainly caused by the lack of uniform distribution of the shade. With regards to PPR, recent studies carried out in Cameroon have shown the impact of environmental factors on cacao pod production dynamics and spread of black pod disease caused by *Phytophthora megakarya* in smallholder's plantations (Deberdt et al., 2008; Ndoumbè et al., 2009).A PPR-susceptible progeny may consequently appear as less infected by the disease if most of its trees are located in low-shaded areas of the farm.

The same progeny ICS40*UPA134 registered the lowest average number of pods per tree, but the greatest average bean weight and average number of bean per pod. The contradicting performances among yield components tallied with previous field observations in different cacao growing conditions. Engels (1985) and Tan (1990) found that Pod weight was positively correlated with number of beans per pod and bean weight, but showed no correlation with yield. Good agricultural practices could minimize the lack of convergence among yield component data.

In conclusion, our study showed that this plot set up with improved cocoa varieties released by extension contained cocoa trees which did not originate from the expected genetic origin, for several potential reasons involving the technique of seed production (selfing of the female parents or harvesting of mislabelled trees in the bi-clonal seed gardens) or events occurring after field establishment (natural or voluntary addition of unselected progenies).In this paper, we described a strategy based on phenotypic assessment combined with parentage analysis using molecular markers. This approach could be very useful when it comes to assess-

ing the performances of released varieties in on farm plots after a long time when the initial trial plots set up for selecting the best varieties do not exist anymore.

It is envisaged to extend this study to other cocoa plots in Cameroon and compare the results from this approach to results obtained from setting up new on farm plots (program ongoing since 2006). The approach described in the current paper is less time consuming (three years for yield assessment and few weeks for molecular analyses) compared to the establishment of new trial plots (five years latency between planting and reaching the plateau indicating stable yield and three years for yield assessment).Indeed, the constant reduction in cost related to molecular analyses makes the approach described in this paper more attractive.

REFERENCES

Adewale DB, Adeigbe OO, Sobowale OI, Dada OS (2014). Breeding Value of Cocoa (*Theobroma cacao* L.) for Pod and Bean Traits: A Consequential Advance in Nigerian Cocoa Breeding Program. Not Sci Biol. 6 (2):214-219.

Aikpokpodion PO (2010). Variation in agro-morphological characteristics of cacao, *Theobroma cacao* L., in farmers' fields in Nigeria, New Zealand Journal of Crop and Horticultural Science 38 (2): 157-170.

Aikpokpodion PO, Motamayor JC, Adetimirin VO, Adu-Ampomah Y, Ingelbrecht I, Eskes AB, Schnell, Kolesnikova-Allen M (2009). Genetic diversity assessment of sub-samples of cacao, Theobroma cacao L., collections in West Africa using simple sequence repeat markers. Tree Genetics and Genome 5(4): 699-711.

Allègre M, Argout X, Boccara M, Fouet O, Roguet Y, Bérard A, Thévenin JM, Chauveau A, Rivallan R, Clement D, Courtois B, Gramacho K, Boland Augé A, Tahi M, Umaharan P, Brunel D, Lanaud C (2012). Discovery and mapping of a new expressed sequence tag-single nucleotide polymorphism and simple sequence repeat panel for large-scale genetic studies and breeding of *Theobroma cacao* L. DNA Res.19(1): 23-35.

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

Anderson EC, Thomson EA (2002). A model-based method for identifying species hybrids using multilocus genetic data.Genetics 160: 1217-1229.

Ardener E (1996). Kingdom on Mount Cameroon: Studies in the History of the Cameroon Coast, 1500–1970. Berghahn Books, Oxford.

Barreto MA, Santos JCS, Correa RX, Luz EDMN, Marelli J, Souza AP (2015). Detection of genetic resistance to cocoa black pod disease caused by three *Phytophthora* species. Euphytica. DOI 10.1007/s10681-015-1490-4.

Bartley BGD (2005).The genetic diversity of cocoa and its utilization. CABI Publishing (ed), Wallingford, UK. 341 pp.

Bastide P, Sounigo O (1993). Production de semences hybrides de cacaoyer en Côte d'Ivoire. Essais de différentes pollinisations manuelles. *Café Cacao Thé* vol. 37 (4) : 295-302.

Cássia BR, Corrêa RX, Santos RC, Machado RCR, Luz EDN, Araújo IS, Ahnert D (2013). Inheritance of the Number of Ovules per Ovary and Selection of Cacao Genotypes Am. J. Plant Sci. 4: 1387-1392.

Cochran WG, Cox GM (1957). Experimental Designs, second ed. Wiley, New York, 76pp.

Cope FW (1984). Cacao Theobroma cacao (Sterculiaceae). In: Longman (ed.), Evolution of Crop Plants, London, pp. 285–289.

Deberdt P, Mfegue CV, Tondje PR, Bon MC, Ducamp M, Hurard C, Begoude BAD, Ndoumbe-Nkeng M, Hebbar PK, Cilas C (2008). Impact of environmental factors, chemical fungicide and biological control on cacao pod production dynamics and black pod disease (*Phytophthoramegakarya*) in Cameroon. Biol. Control. 44: 149–159.

Dias Luiz A, Marita J, Cosme DC, Everaldo G, Tânia MFS (2003). Genetic distance and its association with heterosis in cacao. Braz. Arch .Biol. Technol. 46: 339-347.

Efombagn MIB, Motamayor JC, Sounigo O, Eskes AB, Nyassé S, Cilas C, Schnell R, Manzanares-Dauleux MJ, Kolesnikova-Allen M (2008). Genetic diversity and structure of farm and genebank accessions of cacao (*Theobroma cacao* L.) in Cameroon revealed by microsatellite markers. Tree Genet.and Genome. 4: 821-831.

Efombagn MIB, Nyassé S, Sounigo O, Kolesnikova-Allen M, Eskes AB (2007). Participatory cocoa selection in Cameroon: *Phytophthora* pod rot resistant accessions identified in farmers' field. Crop Prot. 26(10):1467–1473.

Engels JMH (1985). A systematic description of cacao clones: V. hybrids in Papua New Guinea. Trop. Agr. Quantitative genetic aspects of several fruit characters. Cafe Cacao.

Falush D, Stephens M, Pritchard JK (2003). Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. Genetics 164:1567–1587.

Fouet O, Allègre M, Argout X, Jeanneau M, Lemainque A, Pavek S, Boland A, Risterucci AM, Loor G, Tahi GM, Sabau X, Courtois B, Lanaud C (2011). Structural characterization and mapping of functional EST-SSR markers in *Theobroma cacao.*Tree Genet.and Genomes7(4): 799-817.

Hladni N, Škoric D, Kraljevic-Balalic M (2008). Line x tester analysis of morphophysiological traits and their correlations with seed yield and oil content in sunflower (*Helianthusannuus L*.).Genetika 40:135-144.

ICCO (2015). Quarterly Bulletin of Cocoa Statistics, Vol. XLI, No. 3, Cocoa year 2014/15.

Isaac ME, Timmer V,Quashie-Sam SJ (2007). Shade tree effects in an 8-year-old cocoa agroforestry system: biomass and nutrient diagnosis of Theobroma Cacao by vector analysis. Nutr. Cycl. Agroecosyst. 78:155–165.

Iwaro AD, Sreenivasan TN, Butler DR, Umaharan P (2000). Rapid screening for *Phytophthora*pod rot resistance by means of detached pod inoculation. *In* Working procedures for cocoa germplasm evaluation and selection. IPGRI, Rome, pp. 109-113.

Jamieson A, Taylor St CS (1997). Comparisons of three probability formulae for parentage exclusion. Animal Genetics28: 397-400.

Lachenaud P, Zhang D (2008). Genetic diversity and population structure in wild stands of cacao trees (*Theobroma cacao* L.) in French Guiana. Ann. For. Sci. 65: 310.

Lanaud C, Fouet O, Clement D, Boccara M, Risterucci AM, Surejdeo-Maharaj S, Legavre T, Argout X (2009). A meta-QTL analysis of disease resistance traits of *Theobroma cacao* L. Mol. Breed.24: 361-374.

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

Lanaud C, Sounigo O, Amefia YK, Paulin D, Lachenaud P, Clément D (1987). Nouvelles données sur le fonctionnement du système d'incompatibilité du cacaoyer et ses conséquences pour la sélection. *Café*, *Cacao*, *Thé*31 : 267-277.

Liabeuf J, Lotode R, Maucourant P (1962). Rapport annuel. Station du cacaoyer de Nkoemvone. Direction de l'Agriculture. République Fédérale du Cameroun. 248 p.

Mantel N (1967). The detection of disease clustering and a generalized regression approach. Cancer Res. 27: 209-220.

Marshall TC, Slate J, Kruuk LEB, Pemberton JM (1998). Statistical confidence for likelihood-based paternity inference in natural populations.Mol. Ecol. 7: 639-655.

N'Goran JAK, Sounigo O, Clement D (2003). Comparaison des variétés utilisées par les cacaoculteurs ivoiriens. XIV *IntConf Cocoa Res,* Accra, Ghana, October 2003, 95-102.

Ndoumbe-Nkeng M, Efombagn MIB, Nyassé S, Nyemb E, Sache I, Cilas C (2009). Relationships between cocoa *Phytophthora*pod rot disease and climatic variables in Cameroon. Canadian J. Plant Path. 31: 309-320.

Nygren P, Leblanc HA (2009). Natural abundance of 15N in two cacao plantationswith legume and non-legume shade trees. Agroforest. Syst. 76:303–315

Opoku SY, Bhattacharjee R, Kolesnikova-Allen M, Motamayor JC, Schnell R, Ingelbrecht I, Enu-Kwesi L,Adu-Ampomah Y (2007). Genetic diversity in Cocoa (*Theobromacocoa*L.) germplasm collection from Ghana. J. Crop. Improv. 20: 73-87.

Peakall R, Smouse PE (2012). GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research-an update.Bioinformatics 28: 2537-2539.

Pokou ND, Motamayor JC, Tondo CL, Tahi M, Schnell R (2014). Molecular characterization of resistant accessions of cocoa (*Theobroma cocoa* L.) to *Phytophtora* pod rot selected on-farm in Côte-d'Ivoire. Inter. J. Agri. Biosci. 3(5): 236-240.

Pokou ND, N'Goran JAK, Lachenaud PH, Eskes AB, Motamayor JC, Schnell R, Kolesnikova-Allen M, Clément D, Sangaré A (2009). Recurrent selection of cocoa populations in Côte d'Ivoire: comparative Genetic diversity between the first and second cycles. Plant Breed. 128: 514-520.

Pritchard JK, Stephens M, Donnelly P (2000). Inference of population structure using multilocus genotype data. Genetics 155: 945 – 959.

SAS (2000).Statistical Analysis System.SAS/STAT Software Changes and Enhancements through Release 6.12. SAS Institute Inc., Cary. N.C., USA.

Smouse PE, Long JC (1992). Matrix correlation analysis in anthropology and genetics.Yearbook Phys. Anthropol. 35: 187-213.

Sonwa DJ, Nkongmeneck BA, Weise SF, Tchatat M, Adesina AA, Janssens MJJ (2007). Diversity of plants in cocoa agroforests in the humid forest zone of Southern Cameroon. Biodivers. Conserv. 16:2385–2400.

Sounigo O, Efombagn MIB, Feumba de Tchoua LF, Owono B, NsougaAmougou F, Bourgoing R, Mekondane R, Abolo D, Ten Hoopen GM, Nyassé S, Eskes A (2012). Cocoa variety assessment in on-farm progeny trials in Cameroon. In: Alliance of CocoaProducing Countries (COPAL). 17th ConférenceInternationalesur la RechercheCacaoyère. 17, 2012-10-15/2012-10-20, Yaoundé, Cameroun.

Tan GY (1990). Combining Ability Analysis of Yield and Its Components in Cacao. J. Am. Soc. Hort. Sci. 115(3):509-512.