HAL
open science

Genetic diversity and population structure in wild stands of cacao trees (Theobroma cacao L.) in French Guiana

Philippe Lachenaud, Dapeng Zhang

To cite this version:

Philippe Lachenaud, Dapeng Zhang. Genetic diversity and population structure in wild stands of cacao trees (Theobroma cacao L.) in French Guiana. Annals of Forest Science, 2008, 65 (3), pp.1. hal-00883372

HAL Id: hal-00883372

https://hal.science/hal-00883372v1

Submitted on 11 May 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Ann. For. Sci. 65 (2008) 310
© INRA, EDP Sciences, 2008
DOI: 10.1051/forest:2008011

Available online at:
www.afs-journal.org

Original article

Genetic diversity and population structure in wild stands of cacao trees (*Theobroma cacao* L.) in French Guiana

Philippe LACHENAUD[1]*, Dapeng ZHANG[2]

[1] Centre de coopération internationale en recherche agronomique pour le développement, UPR Bioagresseurs de pérennes, TA A-31/02, 34398 Montpellier Cedex 5, France
[2] USDA/ARS, Beltsville Agricultural Research Centre, Sustainable Perennial Crops Laboratory, Plant Sciences Institute, 10300 Baltimore Avenue, Bldg. 1. BARC-W, Beltsville, MD 20705, USA

(Received 23 March 2007; accepted 8 November 2007)

Abstract – The native cacao trees *(Theobroma cacao* L.) in south-eastern French Guiana represent a noteworthy fraction of the genetic diversity of this species. Several scientific surveys have led to the collection of numerous accessions, which provides an opportunity for investigating the pattern of genetic diversity in natural populations. Using a capillary electrophoresis genotyping system, we fingerprinted 189 wild trees in 18 natural populations. Based on the 15 loci SSR profiles, we analyzed the intra- and inter-population variation and their relationship with control clones. The global allelic richness was 4.87 alleles per locus, but was 2.4 on average within individual populations; gene diversity was 0.368 and observed heterozygosity was low (0.160). Major structuring was discovered in the metapopulation of French Guianan wild cacao trees (Fst = 0.20), which could be explained by the biological characteristics of the cacao tree and the climatic events that affected the Guianan forests in the Quaternary period.

cocoa / genetic diversity / population structure / fingerprinting / French Guiana

Résumé – Diversité et structuration génétiques dans des populations naturelles de cacaoyers (*Theobroma cacao* L.) en Guyane française. Les cacaoyers spontanés (*Theobroma cacao* L.) du sud-est de la Guyane française représentent une part notable de la diversité génétique de l'espèce. Plusieurs prospections ont permis la collection de nombreuses accessions, permettant l'étude de la structuration de la diversité dans les populations naturelles. Cent quatre-vingt-neuf cacaoyers spontanés, issus de 18 populations, ont été génotypés grâce à un séquenceur capillaire. Les profils obtenus avec 15 marqueurs SSR ont permis d'analyser les variations intra et inter populations, et leurs relations avec des clones témoins. La richesse allélique globale est de 4,87 allèles par locus, mais de 2,4 en moyenne dans les populations ; la diversité génique est de 0,368 et l'hétérozygotie observée est faible (0,160). Une importante structuration existe dans la métapopulation de cacaoyers spontanés Guyanais, avec un Fst de 0,20, qui pourrait s'expliquer par les caractéristiques biologiques du cacaoyer et les événements climatiques ayant affecté les forêts guyanaises au Quaternaire.

cacao / diversité génétique / structuration génétique / génotypage / Guyane française

1. INTRODUCTION

The cacao tree (*Theobroma cacao* L., Malvaceae) originates from the neotropical rainforests, from Central America (where it could only be feral) to northern Bolivia. Two sub-species have been recognized, *T. cacao* subsp. *cacao* and *T. cacao* subsp. *sphaerocarpum* [10], which roughly correspond to the two cultivated groups called Criollo and Forastero. However, knowledge of the genetic structure of the species provided by molecular markers did not support the existence of those two sub-species. Instead, several morpho-geographical groups have been revealed by studies based on co-dominant molecular markers [30, 43]. The wild cacao populations in south-eastern French Guiana are one of those morpho-geographical groups; they have been known since 1729, then surveyed and collected between 1985 and 1995 in the basins of five rivers through several expeditions [8, 23, 25]. Today, these wild cacao trees are the main part of the ex situ collection at the CIRAD Paracou-Combi station (Sinnamary, French Guiana). A fraction of this germplasm has been distributed to other cacao genebanks, quarantine stations and various producing countries [28]. Characterization and evaluation have already been undertaken for some of these populations using biochemical, morphological and agronomic descriptors, while some clones have been used in breeding programs [26–28].

The current aggregated spatial distribution of cacao in neotropical rainforests (i.e., in small, low-density patches or demes) and its biological characteristics (e.g. rare flowers in the wild state, pollination by tiny insects with a limited sphere of action, cross-fertilizing with generalized self-incompatibility, rare and indehiscent fruits that do not fall to the ground when ripe, seed dispersal over short distances, recalcitrant seeds) predict that the cacao tree must display strong differentiation between populations or sub-populations, resulting from limited gene flow. In addition, the well-documented successive phases of Guianan forest contraction-expansion in the Pleistocene and Holocene epochs [5, 11, 32], combined with the likely isolation of the Guianan shield [3, 7, 44], must

* Corresponding author: philippe.lachenaud@cirad.fr

Article published by EDP Sciences and available at http://www.afs-journal.org or http://dx.doi.org/10.1051/forest:2008011

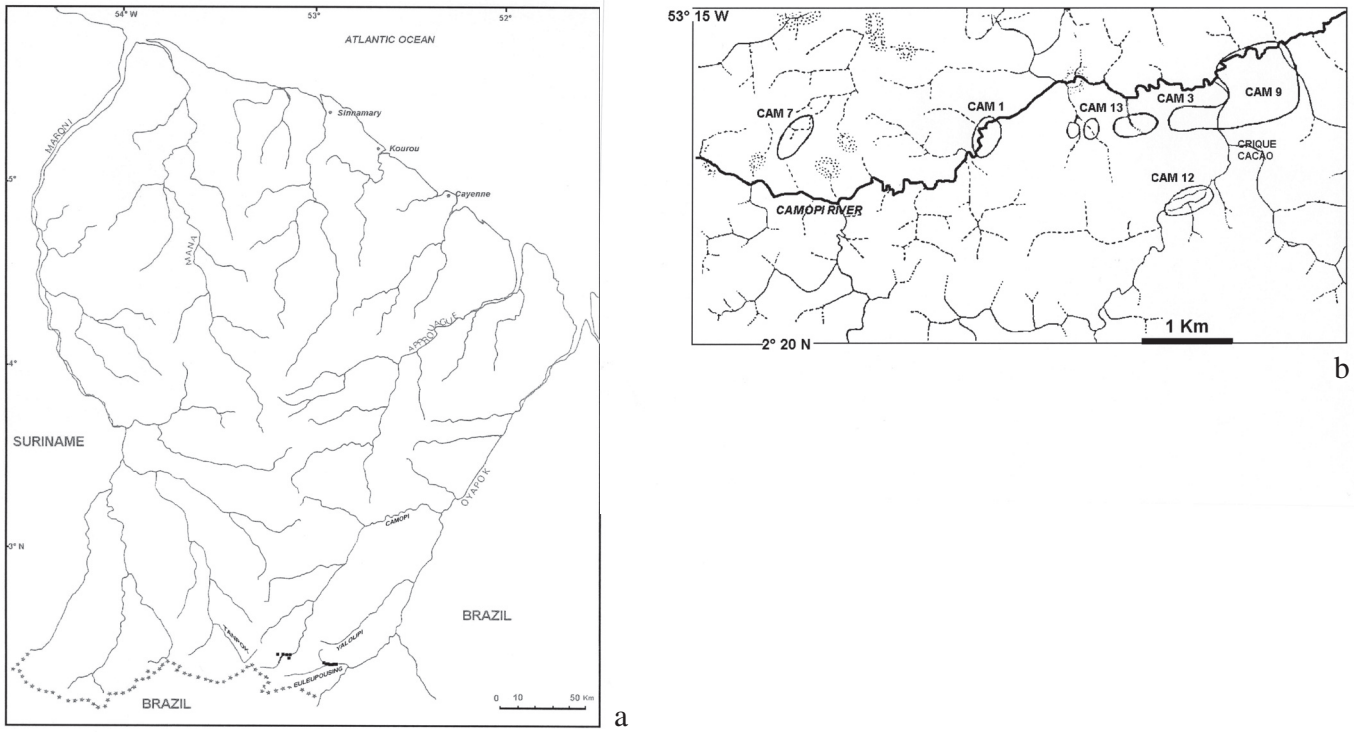

Figure 1. (a) Map of French Guiana and location of the seven populations (demes) studied (■, for more precise localization see [23, 25]). Other demes are found in Oyapok, Tanpok and Yaloupi river basins [23, 25]. (b) Detailed part of (a), showing the location of studied demes in the Upper-Camopi basin. Camopi 3 and Camopi 9 are separated by the Crique Cacao (Cacao creek).

have had major consequences for the isolation and scattering of cacao tree stands. These Cenozoic events may therefore have strengthened the species' natural tendency to differentiate, through founder and genetic drift effects. Nevertheless, to date, those hypotheses have yet to be tested on natural populations of cacao tree.

The study described here aimed at understanding the spatial genetic structure and assessing the genetic diversity in the wild cacao populations in French Guiana, using samples of 189 individual trees representing all the surveyed local wild cacao populations. These 189 trees accounted for 95% of the mother-trees maintained in the French Guiana cacao collection [28].

2. MATERIAL AND METHODS

2.1. Genetic material

The plant material comprised 189 wild trees and various controls. The wild cacao trees were selected in open progenies of wild mother-trees (or in some cases, were clones of mother-trees) planted in collection plots in Paracou-Combi. Pods (or in some cases, budwood) were originally collected from 189 mother-trees in 18 natural stands (demes) identified in the valleys of the Oyapok river, of its tributaries the Euleupousing, Yaloupi and Camopi, and of the Tanpok river (the latter belongs to the Maroni basin) (Fig. 1). The number of available trees varied between one and 48 individuals per deme; only the seven more numerous demes, usable in intra-population genetic diversity computation, are shown in Table I. Thirty-three trees, belonging to demes with lesser numbers of available genotypes (3 in Upper-Oyapok, 5 in Camopi, one in Tanpok and one in Yaloupi valleys [23, 25]) were only used in overall diversity computation.

Table I. Diversity parameters for the seven main populations (demes) and for the whole set of Guianan wild germplasm, compared with the "Primitive" material and the "Parinari" clone controls, with Cam = Camopi and Elp = Euleupousing.

Population	N	Na	He	Ho	F_{IS}
Cam 1	28	2.67	0.247	0.146	0.442
Cam 3	14	1.80	0.248	0.182	0.337
Cam 7	18	2.73	0.315	0.145	0.580
Cam 9	48	3.07	0.322	0.181	0.455
Cam 12	9	1.87	0.249	0.156	0.474
Cam 13	17	2.20	0.319	0.208	0.400
Elp	22	2.47	0.297	0.082	0.746
Mean of 7 demes	22.3	2.40	0.285	0.157	0.491
Total of 7 demes	156	4.53	0.356	0.158	0.624
Overall (18 demes)	189	4.87	0.368	0.160	0.626
"Primitive"	14	2.13	0.252	0.133	0.528
"Parinari"	15	3.00	0.443	0.480	-0.013

N = Sample size, N_a = average number of alleles per locus, H_e = expected heterozygosity (mean gene diversity), H_o = observed heterozygosity and F_{IS} = within-population inbreeding coefficient. The F_{IS} in all populations, except "Parinari", are significantly positive ($P < 0.05$).

The control samples were: (a) fourteen trees descended from germplasm formerly grown in plantations in northern French Guiana [8], called "Primitive" in this study. Twelve of these were of the Amelonado type, one had a Calabacillo-shaped pod and one had a Guianan-shaped pod [25]. These samples, with no Trinitario kinship, were included to test Lanaud's hypothesis [29] that cultivated material in French Guiana was completely different from the local wild cacao trees. (b) Fifteen "Parinari", Upper-Amazon Forasteros clones, originally collected from the bank of the Rio Marañon in Peru [2], from the collection of the Cacao Research Unit (CRU) at the University of the West Indies, Trinidad, and (c) seven international clones from collections at the CRU and CATIE (Centro Agronómico Tropical de Investigación y Enseñanza, Costa Rica): these were two Upper-Amazon Forasteros (PA 120 and IMC 67), three Lower-Amazon Forasteros (BE-3, Amelonado-15 and Amelonado-22), a verified Trinitario (ICS 95) and a Criollo (Criollo-13).

Four healthy young leaves were collected from each tree, and the samples sent by express mail to the USDA Beltsville Agricultural Research Center, Maryland. The fresh leaves were stored frozen at −20 °C until DNA extraction.

2.2. DNA isolation and SSR analysis

DNA was isolated from 50-mg samples of *T. cacao* leaf material using either the DNA XtractTM Plus kit (D2 BioTechnologies Inc., Atlanta, GA, USA) or the DNeasy® Plant System (Qiagen Inc., Valencia, CA, USA). The detailed protocol has been previously described [42, 49].

DNA amplification used 15 primer sets with sequences previously described [31, 42]. Primers were synthesized by Proligo (Boulder, CO, USA), and forward primers were 5'-labeled using WellRED fluorescent dyes (Beckman Coulter, Inc., Fullerton, CA, USA). PCR was performed as described in Saunders et al. [42]. The amplified PCR products were separated by capillary electrophoresis as previously described [42].

2.3. Data analysis

Intra-population genetic diversity was measured by allelic diversity, estimated gene diversity (H_e) [37], observed heterozygosity (H_o) and Fixation index within population (F_{IS}) [47] using PowerMarker v.3.0 [33]. The Exact HW test [20] was used to test deviation from Hardy-Weinberg equilibrium and was performed by the same program.

Inter-population variation was measured using the variation in allelic distributions among populations. The seven populations were defined a priori, following collectors' indications (Tab. I). Allelic composition was assessed using a contingency table test [45] as implemented in PowerMarker v.3.0 [33], with a null hypothesis of identical allelic distributions in all populations. The hierarchical structure was examined by the analysis of molecular variance (AMOVA, [15]), implemented in the Arlequin 3.0 software [16]. The total molecular variance was partitioned as components of among- and within-population. The significance of Φ statistics [15], which are molecular analogous to Wright's F-statistics [47], was tested by permutation for 1000 randomizations. The fixation index (F_{ST}) was calculated and tested for significance by FSTAT version 2.9.3 [17, 18]. Pairwise genetic distances [38] among populations were calculated using PowerMarker v. 3.0 (assuming that both mutations and genetic drift were implicated) and a UPGMA clustering was computed with 100 bootstrap replications. The dendrogram was then produced using the software TreeView version 1.6.6 [39]. To examine the relationship among all the individual accessions from the 18 populations, as well as their relationship with the cacao accessions from other geographical regions of South America, we used the procedure of Principal Coordinate Analysis (PCO) implemented in GenAlEx [40] for computation. The result was presented as a two-axis PCO plot.

3. RESULTS

3.1. Allelic richness and heterozygosities

All the microsatellite loci were polymorphic and validated the independence hypotheses (no pairs of loci were linked). A total of 73 alleles (4.87 alleles per locus) was identified in the metapopulation of Guianan wild cacao trees (Tab. I). In total, the seven main populations had 4.53 alleles per locus. However, within each population, the allelic richness was low, ranging from 1.80 in Camopi 3 to 3.07 in Camopi 9. In the controls, the fifteen "Parinari" samples had 3.0 alleles per locus and the "Primitive" samples had 2.13.

The expected heterozygosity (gene diversity) was 0.368 for all the wild trees and 0.356 for the seven main populations. Their observed heterozygosity was 0.160 and 0.158, respectively. Among the seven populations, the most diverse was Camopi 9 (He = 0.322) and the least diverse was Camopi 1 (He = 0.247; Tab. I). Within-population inbreeding coefficients (F_{IS}), ranging from 0.337 (Camopi 3) to 0.746 (Euleupousing), were all significant in the seven populations.

The Guianan "Primitive" material had a gene diversity of 0.252, which was below the average for the seven main wild populations. The "Parinari" group had the highest gene diversity (0.443).

3.2. Genetic proximities and population structure

The genetic proximities among the seven wild populations, the "Primitive" material and the "Parinari" group are illustrated by the dendrogram in Figure 2. The populations of the Camopi river were divided into two clusters. The first cluster includes populations Camopi 3, 9, 13 and 12 and the other includes populations Camopi 1 and 7, along with the "Euleupousing" population. The "Primitive" material formerly cultivated in French Guiana appeared to be more closely related to the "Parinari" UpperAmazon group than to the local wild populations.

The genetic proximities among all the wild trees, as well as the external controls, are illustrated by the plane of the first two PCO analysis axes (Fig. 3), which accounted for around 68% of total variability, and showed that all the Guianan germplasm was clearly distinct from the international clones used as controls, except a few exceptions belonging to the "Primitive" material. The "Primitive" germplasm appeared to fall between the wild accessions and the control clones.

Table II summarizes the pairwise differentiation between the populations using Nei's genetic distance and the Fixation

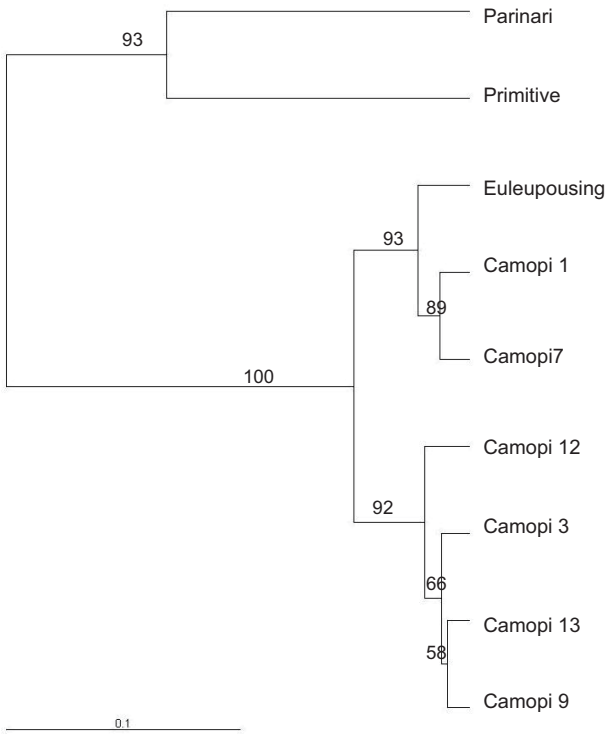

Figure 2. Dendrogram of seven wild cacao populations (demes) from south-eastern French Guiana and their relationship to the domesticated local clones ("Primitive" germplasm), and a population from Peruvian Amazon ("Parinari"), with bootstrap values.

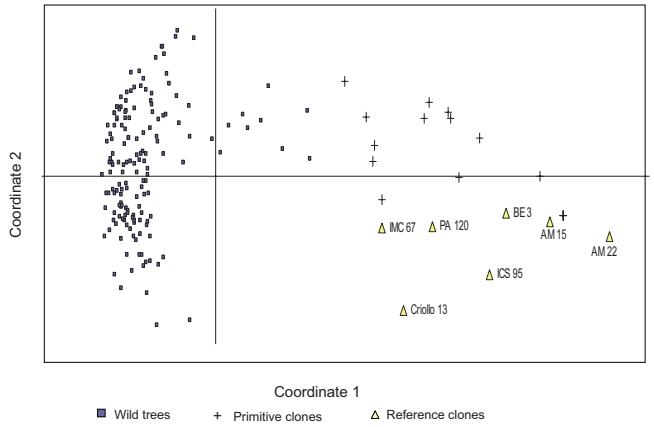

Figure 3. Principal Coordinate Analysis of the French Guianan wild and "Primitive" cacao trees and of the reference Forastero, Trinitario and Criollo germplasm accessions (First axis = 48.8% of total information and the second = 18.9%. AM 15 = Amelonado 15; AM 22 = Amelonado 22).

index (F_{ST}). For Nei's distances, we included the "Primitive" and "Parinari" accessions in the computation, but these two sets of controls were not included in the computation for F_{ST}, as they did not belong to the same natural metapopulation.

The largest F_{ST} was observed between the populations Camopi 7 and Camopi 12 (0.168), followed by Euleupousing – Camopi 12, with a F_{ST} value of 0.166. The smallest differentiation was found between the populations Camopi 13 and 9 (0.026). Nei's pairwise distances ranged from 0.032 (Camopi 9 – Camopi 3) to 0.237 (Camopi 12 – Camopi 7).

Nei's genetic distances confirmed that the "Primitive" materials and "Parinari" clones from the Peruvian Amazon were distant from the Guianan wild cacao trees; the mean genetic distance between the seven wild populations and the "Parinari" and the "Primitive" populations were 1.09 and 0.76, respectively. The "Primitive" material appeared to be closer to the "Parinari" than to the wild populations.

AMOVA analysis revealed highly significant intra- ($P < 0.01$) and inter- ($P < 0.001$) population variation. The majority of the variation (81%) was found within populations (Tab. III).

4. DISCUSSION

4.1. Genetic diversity

The overall allelic richness of the wild material was moderately high (4.87 alleles per locus), but quite low within each population (2.40 on average), which reflected the fact that the demes were well differentiated from each other. Using the same 15 SSR loci, Zhang et al. [48] found an average of 5.70 alleles per locus in a semi-natural population from Río Ucayali in Peru, and 3.68 in a managed population in the Río Huallagua valley. Using 11 microsatellite markers, Sereno et al. [43] reported a figure of 4.45 for 94 accessions representing 14 Amazonian basins. However, areas sampled in those studies were far larger than the area covered in French Guiana.

The total gene diversity of the Guianan wild cacao populations (He = 0.368) was much lower than those previously reported in other wild or cultivated cacao populations. Sereno et al. [43] reported a gene diversity of 0.589 in the natural populations in the Brazilian Amazon. Zhang et al. [48] reported a gene diversity of 0.61 and 0.74 in a semi-natural and a managed population from the Peruvian Amazon. The observed heterozygosity (Ho = 0.160) in the French Guiana wild cacao was also low compared with the populations from the Brazilian Amazon (Ho = 0.347; [43]). The low observed heterozygosity, which reflects the fixation of alleles, was also found within each individual deme, particularly the one from the Euleupousing river (Ho = 0.082).

These results confirmed those obtained by isozyme electrophoresis [27]. However, using individuals reared from seed in curated collections, under managed conditions, might tend to increase the survival rates of homozygous individuals, and this might lead to bias in H_o determination [21].

4.2. Population structure

The present study revealed a clear structure in the natural cacao populations. The high value of F_{ST} (0.199) reflected significant genetic differentiation among populations. That value was consistent with the large share of the "among-populations" variance component, i.e. 19% (Tab. III). Our results confirmed those obtained for F_{ST} with isozyme

Table II. Inter-population variation measured by Nei's distances (under the diagonal) and pairwise F_{ST} (above the diagonal), with Cam = Camopi and Elp = Euleupousing.

Population	Cam 1	Cam 3	Cam 7	Cam 9	Cam 12	Cam 13	Elp	"Primitive"
Cam 1	–	0.136	0.037	0.081	0.148	0.078	0.050	–
Cam 3	0.138	–	0.150	0.040	0.050	0.050	0.147	–
Cam 7	0.036	0.201	–	0.098	0.168	0.074	0.060	–
Cam 9	0.090	0.032	0.141	–	0.053	0.026	0.109	–
Cam 12	0.164	0.042	0.237	0.046	–	0.087	0.166	–
Cam 13	0.085	0.035	0.101	0.029	0.078	–	0.099	–
Elp	0.056	0.187	0.076	0.145	0.224	0.122	–	–
"Primitive"	0.811	0.867	0.718	0.776	0.805	0.772	0.545	–
"Parinari"	1.202	1.202	1.005	1.054	1.126	1.092	0.931	0.511

Population differentiations (F_{ST}) are all significant, except for Cam 1 – Cam 7 and Cam 9 – Cam 13.

Table III. Analysis of molecular variance (AMOVA) for SSR variation among and within seven natural cacao populations (demes) in south-eastern French Guiana, with Cam = Camopi and Elp = Euleupousing.

Source	d.f.	SSD[z]	MSD[y]	Variance component	% Total[x]	P value[w]
Among Pops	6	246.4	41.7	1.602	19%	0.01
Within Pops	149	1155.4	7.0	7.049	81%	0.001
Cam 1	28	152.2	5.4			
Cam 3	14	71.5	5.1			
Cam 7	18	139.4	7.7			
Cam 9	48	340.8	7.1			
Cam 12	9	52.4	5.8			
Cam 13	17	116.5	6.9			
Elp	22	177.5	8.1			
Total	155	1296.7	48.1	8.65		

[z] Sum of squared deviations.
[y] Mean squared deviations.
[x] Percent of total molecular variance.
[w] Probability of obtaining a larger component estimate. Number of permutations = 1000.

electrophoresis [27], where we found a F_{ST} of 0.197 (computed with expected heterozygosities) and of 0.227 (computed with Shannon's indices). These high values are typical of a cross-fertilizing perennial forest species having undergone some major genetic constraints, i.e. genetic drift.

The structure revealed by microsatellite markers also confirmed our earlier results obtained with morphological or biochemical markers [26, 27]: the populations of Camopi 9, 3, 12 and 13 were closely related and this group was separated from the populations of Camopi 1 and Camopi 7.

Pairwise F_{ST} values (Tab. II), which varied from 0.026 to 0.168, should be considered as low to high. Dutech et al. [12, 13] found in *Vacapoua americana* values from 0.008 to 0.195, and a high genetic differentiation among patches, with a global Fst of 0.11.

4.3. "Primitive" material

The origin of the "Primitive" germplasm was uncertain in this study. Apart from one or two exceptions, the "Primitive" clones were clearly separated from the controls of the Forastero (particularly the Amelonados), Trinitario and Criollo groups (Fig. 3) and occupied a position between the Guianan wild populations and the "Parinari" clones. However, they seemed closer to the latter (Tab. II, Fig. 2). It could not be ruled out that some samples might be hybrids between the indigenous wild cacao and some unknown imported foreign germplasm. Another possibility, such as the "historical" hypothesis proposed by Guisan [19], is that some of the "Primitive" individuals were indigenous to French Guiana, but from an as yet unknown population. The low gene diversity (0.252), along with the low observed heterozygosity (Ho) and highly positive inbreeding coefficient (F_{IS}) tend to be in favor of that hypothesis.

4.4. Diversity determinants, biological and historical factors

Our results on the considerable differentiation in populations with low allelic richness might be explained by the biological peculiarities of the cacao tree and by disturbances of climatic conditions that affected the Amazon and Guianan forests in the Quaternary period, as was the case for other forest species in French Guiana [14]. The cacao tree has a low capacity for dispersion. In the wild state, cacao trees usually multiply through vegetative propagation [1, 23, 41]. Flowers and fruits are rare, whereas there is often a profusion of trunks (up to 50; [1]) with secondary rooting and aerial roots [23, 41]. Natural stands of cacao trees form more or less extended aggregates, at low density. Pollinators, midges of the genus *Forcipomya*, only have a limited range that rarely exceeds twelve meters under growing conditions [24, 35, 36]. In cultivation conditions, only a small proportion (1–5%) of flowers set fruits [46]. The fruits are indehiscent with a thick, hard pericarp and do not abscise after senescence, remaining attached indefinitely to the tree. The recalcitrant seeds germinate within the fruit and can be easily lost if the seedlings are not provided with an avenue for escape by an external agent. Seed dispersal, which has rarely been observed under natural conditions, is ensured by animals, mainly monkeys, which open the pods and eat the seeds, dropping some of them to the

ground [1, 46], or seeds are swallowed whole without chewing and end up on the ground in faeces, where they germinate rapidly (endozoochory of the capuchin monkeys *Cebus nigrivittatus* and *C. apella* [41] or of the Red Howler monkey *Alouatta seniculus* [22]). As pods are rarely found close to the ground in the wild state, the role played by ground-dwelling caviomorph rodents such as the agouti (*Dasyprocta leporina*) is doubtless of little importance [1]. The role played by birds, some of which are able to open pods (*Ara* spp.) but unable to transport intact seeds, is unknown, but doubtless of little importance, and the role played by rivers has been convincingly contested [1]. Although the cacao tree is largely an outcrossing species [46], its low capacity for seed and pollen dispersal suggests that it would display substantial genetic structure among natural populations, with low to moderate variability in the demes [13, 34].

The Guyana shield, like the Andean Piedmont and the Atlantic forest of Brazil, is a biogeographical entity that displays a high level of biodiversity and endemism, which suggests that it may have been a refuge zone [44]. Even if the forest were able to survive on the whole, as claimed by Colinvaux et al. [9], the last period of glaciation, the Würm (22000-13000 BP), had the greatest impact on the vegetation of northern South America and the current period bears the imprint of the biogeographical reorganizations of that time [4]. In French Guiana, during the Holocene epoch, various clues (notably, traces of Palaeozoic wildfires) have revealed alternating dry periods (during which the forest could burn over great expanses) and wet periods where it could expand and recover lost ground [4, 5, 32]. The tree species of the current Guianan forest are particularly animal-dispersed, which generally enables rapid re-colonization and is typical of frequently disturbed forests. However, some species with a low colonizing capacity also exist and are often highly dispersed, in an inexplicable way by current forest dynamics. One common hypothesis is that those patches are remnants of ancient distributions remodeled by Palaeozoic wildfires, and subsequent expansion [5]. For instance, the palm *Astrocaryum sciophylum* displays patchy distribution. The seeds exhibit long dormancy (several years) and are dispersed by rodents, which bury them in nearby stores; apparently fewer than one seed per tree germinates each year. Consequently, it is a poor colonizer and spreads slowly (maximum of 2.3 m per year, [6]). The cacao tree could therefore be in the same situation (like the Wacapou *Vouacapoua americana* [14]) and may have survived dry episodes in French Guiana along rivers, in gallery-forests, or in marshy assai-palm (*Euterpe oleacera*) groves.

Acknowledgements: The authors would like to thank Stephen Pinney and Eric Tillson for their contributions to the genotyping. Sue Mischke and Lambert Motilal are thanked for reviewing and editing the manuscript, and Peter Biggins for translating sections of the text.

REFERENCES

[1] Allen J.B., Geographical variation and population biology in wild *Theobroma cacao*, Ph.D. thesis, University of Edinburgh, 1988.

[2] Bartley B.G.D., The genetic diversity of cacao and its utilization, CABI Publishing, Wallingford, 2005.

[3] Bush M., Amazonian speciation: a necessarily complex model, J. Biogeogr. 21 (1994) 5–17.

[4] Caron H., Organisation et dynamique de la diversité génétique de cinq espèces arborées de la forêt guyanaise, Ph.D. thesis, Université de Montpellier, 2000.

[5] Charles-Dominique P., Blanc P., Larpin P., Ledru M.-P., Riéra B., Sarthou C., Forest perturbations and biodiversity during the last ten thousand years in French Guiana, Acta Oecol. 19 (1998) 295–302.

[6] Charles-Dominique P., Chave J., Dubois M.-A., De Granville J.-J., Riéra B., Vezzoli C., Colonization front of the understorey palm *Astrocaryum sciophilum* in a pristine rain forest of French Guiana, Global Ecol. Biogeogr. 12 (2003) 237–248.

[7] Clapperton C.M., Nature of environmental changes in South America at the last Glacial Maximum, Palaeogeogr. Palaeoclimat. Palaeoecol. 101 (1993) 189–208.

[8] Clément D., Cacaoyers de Guyane. Prospections, Café Cacao Thé 30 (1986) 11–36.

[9] Colinvaux P.A., Irion G., Räsänen M.E., Bush M.B., Nunes de Mello J.A.S., A paradigm to be discarded: Geological and paleoecological data falsify the Haffer & Prance refuge hypothesis of Amazonian speciation, Amazoniana 16 (2001) 609–646.

[10] Cuatrecasas J., Cacao and its allies. A taxonomic revision of the genus *Theobroma*, Contributions from the United States National Herbarium 35, 1964.

[11] De Granville J.-J., Rain forest and xeric flora refuges in French Guiana, in: Prance G.T. (Ed.), Biological diversification in the tropics, Columbia University Press, N.Y., 1982, pp. 159–181.

[12] Dutech C., Seiter J., Petronelli P., Joly H.I., Jarne P., Evidence of low gene flow in a neotropical clustered tree species in two rainforest stands of French Guiana, Mol. Ecol. 11 (2002) 725–738.

[13] Dutech C., Joly H.I., Jarne P., Gene flow, historical population dynamics and genetic diversity within French Guianan populations of a rainforest tree species, *Vouacapoua americana*, Heredity 92 (2004) 69–77.

[14] Dutech C., Maggia L., Tardy C., Joly H.I., Jarne P., Tracking a genetic signal of extinction-recolonization events in a neotropical tree species, *Vouacapoua americana* Aublet in French Guiana, Evolution 57 (2004) 2753-2764.

[15] Excoffier L., Smouse P.E., Quattro J.M., Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data, Genetics 131 (1992) 479–491.

[16] Excoffier L., Laval G., Schneider S., Arlequin ver. 3.0: An integrated software package for population genetics data analysis, Evol. Bioinform. Online 1 (2005) 47–50.

[17] Goudet J., FSTAT (vers. 1.2): a computer program to calculate *F*-statistics, J. Hered. 86 (1995) 485–486.

[18] Goudet J., FSTAT, a Program to estimate and test gene diversities and fixation indices (V. 2.9.3), 2001. Available from http://www.unil.ch/izea/softwares/fstat.html.

[19] Guisan J., De la culture du cacao, Chap. IX, in: Traité sur les terres noyées de Guyane, Imprimerie du Roy 2e éd., 1825.

[20] Guo S.W., Thompson E.A., Performing the exact test of Hardy-Weinberg proportion for multiple alleles, Biometrics 48 (1992) 361–372.

[21] Hamrick J.L., Murawsky D.A., Nason J.D., The influence of seed dispersal mechanism on the genetic structure of a tropical tree population, Vegetatio 107/108 (1993) 281–297.

[22] Julliot C., Sabatier D., Diet of the red howler monkey (*Alouatta seniculus*) in French Guiana, Int. J. Primat. 14 (1993) 527–550.

[23] Lachenaud P., Sallée B., Les cacaoyers spontanés de Guyane. Localisation, écologie, morphologie, Café Cacao Thé 37 (1993) 101–114.

[24] Lachenaud P., Sounigo O., Production de semences de cacaoyer en Côte d'Ivoire : Influence des proportions relatives des parents d'un champ semencier biclonal sur la quantité, la qualité et l'origine génétique des semences, Café Cacao Thé 35 (1991) 105–112.

[25] Lachenaud P., Mooleedhar V., Couturier C., Wild cocoa trees in French Guiana. New surveys, Plant. Rech. Dev. 4 (1997) 25–32.

[26] Lachenaud P., Bonnot F., Oliver G., Use of floral descriptors to study variability in wild cocoa trees (*Theobroma cacao* L.) in French Guiana, Genet. Resour. Crop Evol. 46 (1999) 491–500.

[27] Lachenaud P., Sounigo O., Oliver G., Genetic structure of Guianan wild cocoa (*Theobroma cacao* L.) described using isozyme electrophoresis, Plant. Genet. Ressour. Newsletter 139 (2004) 24–30.

[28] Lachenaud P., Sounigo O., Sallée B., Les cacaoyers spontanés de Guyane française : état des recherches, Acta Bot. Gallica 152 (2005) 325–346.

[29] Lanaud C., Nouvelles données sur la biologie du cacaoyer (*Theobroma cacao* L.) : diversité des populations, système d'incompatibilité, haploïdes spontanés. Leurs conséquences pour l'amélioration génétique de cette espèce, Ph.D. thesis, Université de Paris-Sud 11 (Orsay), 1987.

[30] Lanaud C., Motamayor J.C., Sounigo O., Le cacaoyer. in: P. Hamon et al. Eds., Diversité génétique des plantes tropicales cultivées, Collection Repères, Cirad, 1999.

[31] Lanaud C., Risterucci A.M., Pieretti I., Falque M., Bouet A., Lagoda P.J.L., Isolation and characterization of microsatellites in *Theobroma cacao* L., Mol. Ecol. 8 (1999) 2141–2143.

[32] Ledru M.-P., Blanc P., Charles-Dominique P., Fournier M., Martin L., Riera B., Tardy C., Reconstitution palynologique de la forêt guyanaise au cours des 3000 dernières années, C. R. Acad. Sci. 324 (1997) 469–476.

[33] Liu K., Muse S.V., PowerMarker: an integrated analysis environment for genetic marker analysis, Bioinformatics 21 (2005) 2128–2129.

[34] Loveless M.D., Hamrick J.L., Ecological determinants of genetic structure in plant populations, Ann. Rev. Ecol. Syst. 15 (1984) 65–95.

[35] Lucas P., Étude de la pollinisation du cacaoyer par l'emploi d'éléments marqués, in: Actes 5e Conférence Internationale sur les Recherches Cacaoyères, Ibadan, Nigeria, 1977, pp. 134–144.

[36] Massaux F., Tchiendji C., Missé C., Decazy B., Étude du transport du pollen de cacaoyer par marquage au ^{32}P, Café Cacao Thé 20 (1976) 163–172.

[37] Nei M., Molecular Evolutionary Genetics, Columbia University Press, New York, 1987.

[38] Nei M., Genetic distance between populations, Am. Nat. 106 (1972) 283–292.

[39] Page R.D.M., TreeView version 1.6.6. Available at: http://taxonomy.zoology.gla.ac.uk/rod/treeview.html, distributed by the Division of Environmental and Evolutionary Biology, Institute of Biomedical and Life Sciences, University of Glasgow, Glasgow, 2001.

[40] Peakall R., Smouse P., GenAlEx 6: genetic analysis in Excel. Population genetic software for teaching and research, Mol. Ecol. Notes 6 (2006) 288–295.

[41] Sabatier D., Prévost M.-F., Une forêt à cacaoyers sauvages sur le haut-Camopi, en Guyane française, Laboratoire de Botanique, Centre Orstom de Cayenne, 1987.

[42] Saunders J.A., Mischke S., Leamy E.A., Hemeida A.A., Selection of international molecular standard for DNA fingerprinting of *Theobroma cacao*, Theor. Appl. Genet. 110 (2004) 41–47.

[43] Sereno M.L., Albuquerque P.S.B., Vencovsky R., Figueira A., Genetic diversity and natural population structure of cacao (*Theobroma cacao* L.) from the Brazilian Amazon evaluated by microsatellites markers, Conserv. Genet. 7 (2006) 13–24.

[44] Van der Hammen T., Ice age tropical South America: What was it really like?, Amazoniana 16 (2001) 647–652.

[45] Weir B.S., Cockerham C.C., Estimating F-statistics for the analysis of population structure, Evolution 38 (1984) 1358–1370.

[46] Wood G.A.R., Lass R.A., Cocoa, 4th ed., Tropical Agriculture Series, Longman, London, 1984.

[47] Wright S., The interpretation of population structure by F-statistics, with special regard to systems of mating, Evolution 19 (1965) 395–420.

[48] Zhang D., Arevalo-Gardini E., Mischke S., Zúñiga-Cernades L., Barreto-Chavez A., Adriazola del Aguila J., Genetic diversity and structure of managed and semi-natural populations of cocoa (*Theobroma cacao*) in the Huallaga and Ucayali Valleys of Perú, Ann. Bot. (Lond.) 98 (2006) 647–655.

[49] Zhang D., Mischke S., Goenaga R., Hemeida A.A., Saunders J.A., Accuracy and reliability of high-throughput microsatellite genotyping for cacao clone identification, Crop Sci. 46 (2006) 2084–2092.