

New insights on spatial genetic structure and diversity of *Coffea canephora* (Rubiaceae) in Upper Guinea based on old herbaria

Jean-Pierre Labouisse^{1,2,3,*}, Philippe Cubry⁴, Frédéric Austerlitz⁵,
Ronan Rivallan^{1,2,3} & Hong Anh Nguyen^{1,2,3}

¹CIRAD, UMR AGAP, FR-34398 Montpellier, France

²AGAP, Université de Montpellier, CIRAD, INRAE, Montpellier SupAgro, Montpellier, France

³Postal address: CIRAD, UMR AGAP, Avenue Agropolis, TA A-108/03, FR-34398 Montpellier Cedex 5, France

⁴DIADÉ, Université de Montpellier, IRD, FR-34394 Montpellier, France

⁵UMR 7206 Éco-Anthropologie, CNRS, MNHN, Université Paris Diderot, FR-75116 Paris, France

*Corresponding author: jean-pierre.labouisse@cirad.fr

CIRAD: Centre de coopération Internationale en Recherche Agronomique pour le Développement; AGAP: Amélioration Génétique et Adaptation des Plantes tropicales et méditerranéennes; CNRS: Centre National de la Recherche Scientifique; INRAE: Institut National de Recherche pour l'Agriculture, l'Alimentation et l'Environnement; DIADÉ: Diversité - Adaptation - Développement des plantes; IRD: Institut de Recherche pour le Développement; MNHN: Muséum National d'Histoire Naturelle; UMR: Unité Mixte de Recherche

Backgrounds and aims – Previous studies showed that robusta coffee (*Coffea canephora* Pierre ex A.Froehner), one of the two cultivated coffee species worldwide, can be classified in two genetic groups: the Guinean group originating in Upper Guinea and the Congolese group in Lower Guinea and Congolia. Although *C. canephora* of the Guinean group is an important resource for genetic improvement of robusta coffee, its germplasm is under-represented in *ex situ* gene banks and its genetic diversity and population structure have not yet been investigated.

Methods – To overcome the limitations of living collections, we explored old herbarium specimens collected in Guinea and Côte d'Ivoire and conserved at the Muséum National d'Histoire Naturelle, Paris. First, we reviewed the history of collection missions in both countries and how the *C. canephora* herbaria from the Muséum were assembled. Then, using 23 nuclear microsatellite markers, factorial and model-based Bayesian analyses, we investigated the genetic diversity of 126 specimens and 36 controls, analysed their distribution among the Congolese and Guinean groups, and estimated admixture proportions for each individual.

Key results – For the first time, we detected population genetic structure within the Guinean group of *C. canephora*. The Guinean genotypes can be assigned to five sub-groups with distinct geographic distribution, especially in Guinea where two sub-groups (Maclaudii and Gamé) are characterized by a low level of admixture due to geographical isolation.

Conclusions – We showed how combining a literature review and genetic data from old herbarium specimens can shed light on previous observations made by botanists and guide further actions to better preserve native coffee plants in forest remnants of West Africa.

Keywords – *Coffea canephora*; Côte d'Ivoire; genetic diversity; Guinea; herbarium; population genetic structure; robusta coffee; Rubiaceae; SSR markers; West African forest.

INTRODUCTION

Coffee is a popular beverage enjoyed by millions of consumers throughout the world and a major export commodity, with about nine million tons of green coffee produced yearly (ICO 2018: 7). Among the 124 species of the genus *Coffea* L. (Davis et al. 2006; Davis 2011), only two are widely cultivated: *C. arabica* L. (arabica coffee) and *C. canephora* Pierre ex A.Froehner (robusta coffee). A third species is much more rarely in cultivation, *Coffea liberica* W.Bull. ex Hiern (liberica coffee). While arabica remains the main traded species, robusta's share has gradually increased since the beginning of the twentieth century and now accounts for almost 40% of the world's coffee market (ICO 2018: 7). Robusta production has also been boosted by the demand for instant coffee in emerging countries and by the wide price gap between arabica and robusta (Wallengren 2017). While there are vast tropical lowland reserves suitable for the production of *C. canephora*, it is likely that, in a few decades, optimum cultivation requirements will become increasingly difficult to reach in many arabica coffee producing regions due to global warming (Davis et al. 2012). Arabica producers in South and Central America are already experiencing a devastating outbreak of coffee leaf rust (CLR), caused by *Hemileia vastatrix* Berk. & Broome (Avelino et al. 2015). *Coffea canephora* is more productive, less sensitive to high temperatures, and more resistant to CLR than *C. arabica*. Although robusta beverage is less prized than arabica, there is a high variability in quality traits within this species and the presence of consistent quantitative trait loci provides breeders with promising tools to improve the cup quality of robusta coffee (Leroy et al. 2011).

Coffea canephora is a perennial forest shrub with the largest natural distribution range for a coffee tree species, stretching from Guinea to Uganda with a discontinuity in the Togo-Benin region (Dahomey Gap) (Cubry et al. 2013b). The type specimen of the species *C. canephora* was collected in Gabon in 1895 by the Reverend Théophile Klaine and sent to the botanist Jean Baptiste Louis Pierre at the Muséum National d'Histoire Naturelle (MNHN) in Paris, France under the name "Café indigène des Ishira" [specimen Klaine 247 (P)]. Pierre named the species *Coffea canephora*; Pierre's name remained unpublished, however, and was validated by Froehner (1897).

Coffea canephora is a diploid, allogamous, and self-incompatible species (Devreux et al. 1959). Its geographical range and mating system, as well as the impact from past climates, have resulted in wide genetic variability within the species (Gomez et al. 2009). Another consequence of its mating system is that the majority of cultivated *C. canephora* is still made up of unselected populations obtained from open-pollinated seeds (Eskes & Leroy 2012: 82).

A major breakthrough in our understanding of the genetics of *C. canephora* was made in the 1980s by Berthaud (1984) who identified two main diversity groups by analysing isozyme polymorphism. Following the chorological system established by White (1979), Berthaud proposed to classify the wild populations growing in Upper Guinea as "Guinean", and those from Lower Guinea and Congolia as "Congolese". Further studies based on the use of DNA markers (RFLP

and SSR) validated this classification in two major genetic groups and led to the recognition of five sub-groups (SG1, SG2, B, C, and UW) within the Congolese group (Dussert et al. 2003; Gomez et al. 2009; Musoli et al. 2009; Cubry et al. 2013b; Leroy et al. 2014).

In Upper Guinea, wild *C. canephora* was recorded for the first time in forests of Guinea and Côte d'Ivoire by Auguste Chevalier, in 1905 and 1909 respectively (Chevalier 1905, 1920: 336). In both countries, the cultivation of robusta began in the second decade of the twentieth century, mainly on the initiative of a few European planters (Portères 1939, 1962; Fréchou 1955; Cordier 1961). Although these planters may have used seeds and seedlings taken from the forests close to their plantations, they mainly imported robusta seeds from Gabon and the former Belgian Congo (now the Democratic Republic of Congo) to compensate for a shortage of plants and, from 1921, for losses caused by attacks by the coffee berry borer *Hypothenemus hampei* (Ferrari) (Portères 1939). Massive quantities of robusta seeds belonging to the Congolese group were again introduced after the outbreak of tracheomycosis, a disease caused by *Gibberella xylarioides* R.Heim & Saccas, which affected West African plantations from 1948 onwards (Cordier 1961). Thus, *C. canephora* plants of the Congolese group coexisted with local *C. canephora* in the plantations. In Côte d'Ivoire, Berthaud (1984) analysed the genetic diversity of several clones previously selected for their outstanding agronomic value. He demonstrated that natural hybridization occurred between the two genepools and showed that a majority of these selected clones were Guinean-Congolese F_1 hybrids. In the following years, a breeding programme based on reciprocal recurrent selection was launched in Côte d'Ivoire to create hybrids from the germplasm of the two genetic groups (Leroy et al. 1993, 1997).

The genetic resources conserved in *ex situ* gene banks were the basis of that breeding programme. The largest gene bank of *C. canephora* is located at the Centre National de Recherche Agronomique (CNRA) research station near Divo, Côte d'Ivoire, where about 1900 clonal accessions are conserved in the field. Many of these accessions have been genetically characterized over the past 30 years using isozymes, RFLP, or SSR markers, e.g. by Montagnon et al. (1992a, 1993), Dussert et al. (2003), Gomez et al. (2009), and Cubry et al. (2013b). Only 205 accessions (about 10%) of *C. canephora* germplasm in the CNRA gene bank belong to the Guinean group. In Guinea itself, a field gene bank of 101 accessions is currently maintained by the Institut de Recherche Agronomique de Guinée (IRAG) at the Sérédou research station near Macenta, Guinea Forest Region. It consists only of clones and progeny of planting material imported in the past from the former Belgian Congo. Other worldwide collections of robusta germplasm contain mainly accessions belonging to the Congolese group (Anthony et al. 2007). Only two French institutes, the Institut de Recherches pour le Développement (IRD) and the Centre de Coopération Internationale en Recherche Agronomique pour le Développement (CIRAD) conserve a few accessions of the Guinean group in their biological resource centres in Réunion and in French Guiana respectively.

Consequently, although *C. canephora* native to Upper Guinea may be an important source of diversity for the genetic improvement of the species, its germplasm is underrepresented in *ex situ* gene banks. While some populations from Côte d'Ivoire have been recently collected and characterized, very little is known about germplasm from Guinea. Since the mid-20th century, the natural forests of Côte d'Ivoire and Guinea have been highly threatened by deforestation (Brou et al. 2000), reinforcing the need to assess the existence of potentially valuable and largely untapped genetic resources for current and future breeding schemes.

To overcome the limitations of living collections and expand the geographical and temporal range of data on *C. canephora* belonging to the Guinean group, we managed to get access to old herbarium specimens. Recent advances in genomic technology have revealed the potential of herbarium specimens for population genetics and phylogeographic studies (Lister et al. 2010). Genotyping studies on herbarium samples using SSR markers have been successfully conducted on various species, e.g. on 100 year-old emmer wheat (Lister et al. 2008) and on sweet potato samples dating back from the eighteenth century (Roullier et al. 2013). To carry out our study, we used the resources of two herbaria located in the MNHN in Paris. The Paris vascular plant herbarium, registered in Index Herbariorum as P (Thiers 2019), conserves 91 herbarium sheets of *C. canephora* collected in Côte d'Ivoire and Guinea corresponding to ca. 50 distinct plant specimens, including the oldest specimens collected in 1905 by Chevalier (1905). Roland Portères was another prolific collector of African cultivated plants (Leroy 1974). The second herbarium (acronym PAT for “Paris, Agronomie Tropicale” according to Index Herbariorum), currently managed by the Laboratoire d'Éco-Anthropologie (<http://ecoanthropologie.cnrs.fr>), conserves a thousand specimens of *Coffea* spp. collected by Portères, a quarter of which are wild and cultivated *C. canephora* collected in Guinea and Côte d'Ivoire between 1929 and 1961. Our study is restricted to these two countries, which represent only a part of the Upper Guinea subcentre, but cover most of its west – east extension. Large herbarium and germplasm collection programmes were conducted in these former French colonies, where coffee cultivation was particularly developed with active administration support to research and extension services.

The aim of our study is to provide an updated overview of the geographical distribution and genetic diversity of *C. canephora* originating in Guinea and Côte d'Ivoire. First, we provided a detailed overview about the history of collection missions in both countries and how the coffee herbaria were assembled. Then, we investigated the pattern of genetic diversity of 126 herbarium specimens, supplemented by 36 genotypes used as controls, with a set of 23 polymorphic nuclear markers (SSRs) and a combination of principal coordinates analysis (PCoA), neighbor-joining (NJ) analysis, and model-based Bayesian analysis. We discussed the results in the light of historical documentation. Lastly, we gave an overview of deforestation trends and of other factors that could hamper further collections of native *C. canephora* in both countries.

REVIEW OF HERBARIUM DATA

Guinea

About three thousand kilometres from the type locality in Gabon, the first botanical record of wild *C. canephora* of West Africa was made by Chevalier in 1905 in Guinea, on the southern edge of the Fouta Djallon plateau (fig. 1) near the village of Bilima, 20 km west of Mamou (Chevalier 1905). Chevalier named it *Coffea maclaudii* (as *Maclaudi*) after Charles Maclaud, a colonial military doctor, who had reported the occurrence of coffee plants in the upper valley of the Konkouré River a few years before (Maclaud 1899). With Octave Caille, a head gardener of the Jardin des Plantes in Paris, Chevalier (1914) collected several samples from these coffee trees and sent them to P [Apr. 1905: *Chevalier 12181, 12330* (type specimen), *12331* (specimen lost), *12332, 12333*; May 1905: *12332bis*; 8 Sep. 1905: *14893*]. Caille made an additional collection in June 1913. From his monograph (*Les Caféiers du Globe*) onwards, Chevalier named it *C. canephora* var. *maclaudii* (Chevalier 1929: 83).

Another important contributor to the P herbarium was Pierre Barthe, an agronomist who was head of the agricultural extension service of the *cercle* of Macenta (a *cercle* was the smallest unit of administration in French Colonial Africa), Guinea Forest Region from 1934 to 1936 (Anonymous 1934). In 1936, he collected coffee in forests and plantations in the region, established a collection plot at the Macenta farm school and sent samples to P.

In 1939, the Sérédou experimental station for coffee and *Cinchona officinalis* L. was established near the village of Sérédou in the Ziama massif, 35 km southeast of Macenta, with Portères as director (Tourte 2005a: 185). Portères planted several collection plots mainly with clones or seeds transferred from Côte d'Ivoire, but originally imported from the Lula and Yangambi research stations of the Institut National pour l'Étude Agronomique du Congo belge (INEAC). He also collected material from local plantations and from Barthe's collection (Sérédou centre annual report 1958, unpublished report). Portères sent some samples from his collections to PAT in 1958 and 1959.

In 1962, Portères published a review of *Coffea* species cultivated in Guinea, which remains the most detailed and comprehensive work on the topic to date (Portères 1962). He classified the *C. canephora* specimens that “grew spontaneously” in Guinea in three groups:

- (i) var. *maclaudii* A.Chev.: from Bilima near Mamou.
- (ii) cultivar Gamé (*sensu stricto*): from the village of Bambaradou, a few kilometers from Macenta, and named after Gamé Guilavogui, the chief of the *canton* (a territorial subdivision gathering several villages) of Kolibirima Toma in the early 1930s (JOGF 1935) who promoted Gamé cultivation with the support of the colonial administration. Portères reported that 50% of coffee area in the *cercle* of Macenta was planted with Gamé cultivar in the late 1950s. He considered this cultivar more productive than any other material introduced in Guinea and its cup quality as “excellent”. In addition to Gamé s.s., he defined a “Gamé geographical-racial complex” (or *Géo-Gamé*) as a group of coffee populations that he thought native and growing over a vast area that

extends beyond the border of the *cercle* of Macenta, from Kissidougou (to the north-west) to Nzérékoré and Beyla (to the east).

- (iii) cultivar Gouecké: originally from the vicinity of Gouecké in the north of the *cercle* of Nzérékoré.

When Guinea became independent in 1958, the Sérédou coffee collection contained coffee trees belonging to the Congolese group and only eight clones of Gamé (Sérédou centre annual report 1958, unpublished report). In 1990, IRAG research staff started to set up a new collection of clonal material with remaining plants from Portères' collection plots and a few clones collected from plantations. At the time of writing (June 2018), the IRAG gene bank comprised 101 accessions of *C. canephora*, most of which are clones or progenies of the Congolese group. No cv. Gamé or var. *maclaudii* genotypes remain.

Côte d'Ivoire

Chevalier made the first record of native *C. canephora* in Côte d'Ivoire in 1909, in the forests near Assikasso in the Indénie Region, 45 km northeast of Abengourou, in the far east of the country [specimens: 18 Dec. 1909: Chevalier 22589, 22590, and 22602 (P)]. Based on morphological similarities

with samples collected in Bilima (Guinea), more than 1000 km to the west, he identified it as conspecific with his *Coffea maclaudii* A.Chev. from Guinea (Chevalier 1920: 336). In 1912, Arsène Dellabonin, an employee of the Assikasso agricultural station, collected similar populations in the same area, near the village of Kongodia, and sent samples to P (Portères 1937b). This coffee was widely cultivated from 1920 in eastern Côte d'Ivoire under the name of *Café Petit Indénie*, in comparison with the species *C. liberica* found by Chevalier in the same area and called *Café Gros Indénie* because of the large size of the bean (Portères 1939).

In 1912, while hunting near Koro, northeast of Touba (Western Côte d'Ivoire), the planter Jules Landré found wild *C. canephora* in gallery forests along the Irama-ba River and two tributaries of the Sassandra River: the Férédougouba and the Boa-ba. This coffee, and other populations gathered in the same area by Le Campion in 1913, were called *Café Landré* or *Café Touba* (Portères 1937b). In 1915, it was introduced into the collection of the agricultural station of Bingerville near Abidjan (Portères 1934) and later duplicated and sent to the station of the Institut d'Enseignement et de Recherches Tropicales (IDERT) near Adiopodoumé, from where Portères sampled it for PAT.

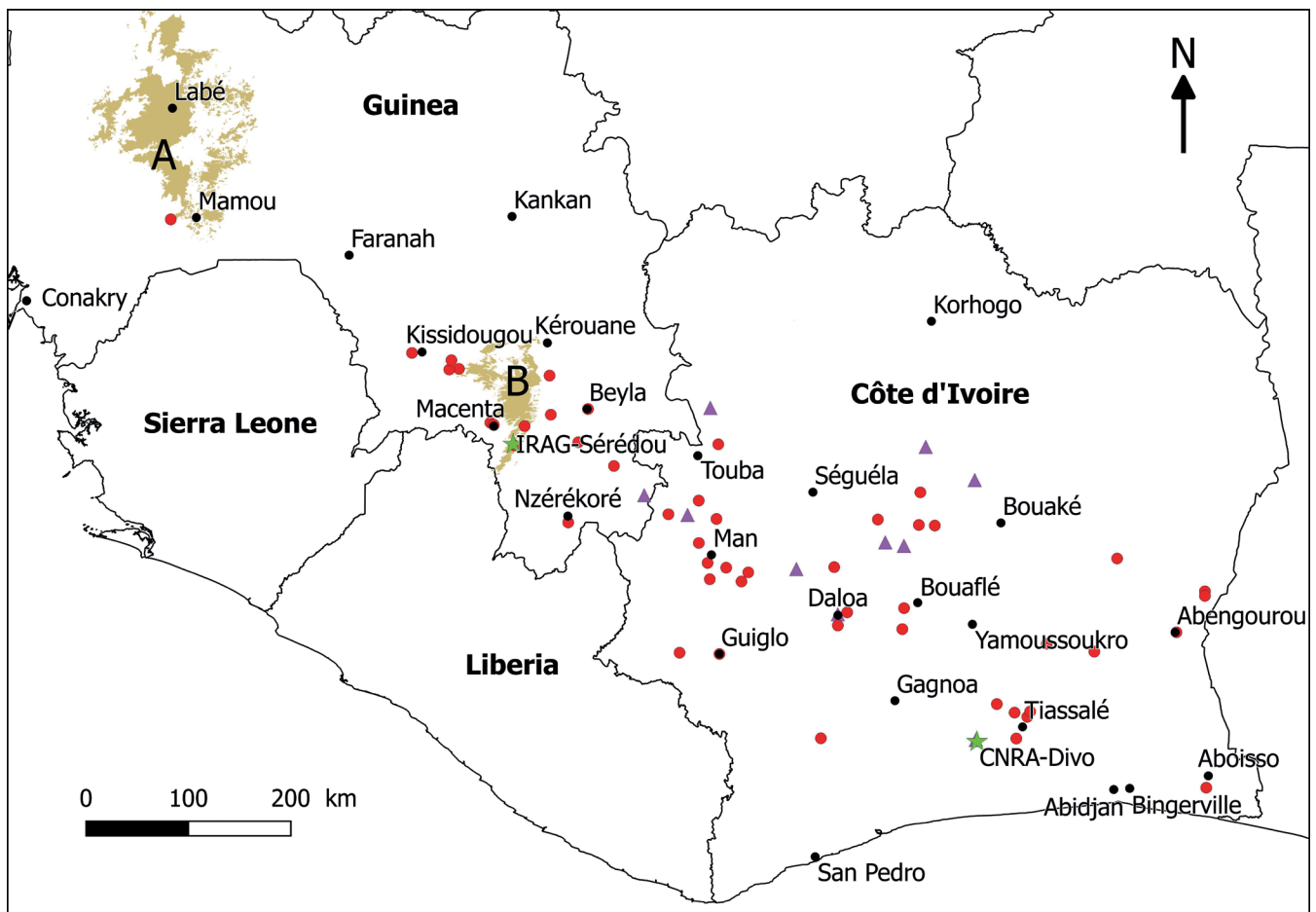


Figure 1— Collection points of the *Coffea canephora* herbarium specimens (red dots) and genotypes used as controls (purple triangles), main towns (black dots), and IRAG and CNRA research stations mentioned in the text (green stars). Two main geographic features of Guinea (elevation > 800 m a.s.l.), the Fouta Djallon plateau (A) and the Ziamia massif with its northern extension (B), are shown in brown.

In 1929, Portères founded the agricultural station of Man, in western Côte d'Ivoire, and set up collection plots and comparative trials of various species of *Coffea*, including native and imported *C. canephora* (Poupart 1938; Tourte 2005a: 193). From 1929 to 1931, Portères collected samples around Man, Touba, and Béoumi. In 1937, he published a first review of *Coffea* spp. in Côte d'Ivoire with their geographical distribution and ecology, the maps of the collection sites in the above-mentioned areas, and the main morphological characteristics of the coffee plants (Portères 1937a, 1937b, 1937c).

Later, in 1958 and again in 1962, Portères and Louis Cordier, a geneticist at the Centre de Recherches Agronomiques (CRA) in Bingerville, assembled the largest coffee herbarium collection ever made in Côte d'Ivoire. These collections included coffee plants from various origin: local or imported, from forests, smallholders' plantations, estates, and agricultural research stations. Duplicates were sent to PAT. There is no published inventory of this herbarium but handwritten documents by Portères give detailed information on the location, origin and morphological traits of the specimens.

Beginning in 1957, the coffee germplasm preserved in the first research stations (Bingerville, Akandjé, Abengourou) and agricultural stations (Man, Gagnoa) was gradually transferred to a new site, near the town of Divo (fig. 1), which became the main research station of the Institut Français du Café et du Cacao (IFCC 1960; Tourte 2005b: 221). This station and its coffee gene bank are now managed by the CNRA.

When Côte d'Ivoire became independent, Portères (1959) and Cordier (1961) traced the history of coffee cultivation and made an inventory of the native and imported coffee material in the country. On the basis of morphological traits, they classified the populations of *C. canephora* into two groups: "Kouilou" and "Robusta". On average, "Kouilou" seeds, flowers, leaves, and the tree itself are smaller than those of "Robusta" (Portères 1959). Portères and Cordier classified most of the populations of *C. canephora* native to Côte d'Ivoire (Petit Indénié, Touba, Bandama, Tos, Kouibly, Agbo, Dianlé, etc.) as belonging to the "Kouilou" group. Only the Ébobo population, which Portères thought originated near Aboisso in the extreme southeast of Côte d'Ivoire, was classified as "Robusta". In fact, there was already a *Café du Kouilou* (or *Café Kouilou*) that was named after the Kouilou River in the Republic of the Congo, along which it was discovered in the 1880s (Chevalier 1929: 85–86). Based on similar morphological characteristics, Chevalier used the term Kouilou in 1909 to describe the *C. macclaudii* that he had found in Guinea in 1905 (Chevalier 1909: 260). Kouilou is also spelled *Kouillou*, *Quillou*, or *Kwilu*. By orthographic distortion, the term became *Conilon* in Brazil and is still applied to the first *C. canephora* populations introduced at the beginning of the twentieth century in that country (Ferreira 2007). The use of the terms Kouilou and Robusta has caused some confusion in the classification of *C. canephora*, an allogamous species with great variability in morphological traits, even within a single population. Berthaud's genetic works clarified the nomenclature. He found that: (i) all populations of "Robusta" (including the Ébobo population), as

well as the true *Café du Kouilou* originating in the Republic of the Congo, belong to the Congolese group and (ii) all the "Kouilou" populations in Côte d'Ivoire belong to the Guinean group (Berthaud 1984: 111).

Following Berthaud's findings, major germplasm collecting efforts resumed from 1975, mainly in Côte d'Ivoire (Berthaud 1984: 12; Le Pierrès et al. 1989; Couturon & Montagnon 1991). Twenty-one populations of *C. canephora* were collected from forest sites in Côte d'Ivoire, and one population from the forest of Piné in Guinea. These represent the largest part of the Guinean coffee germplasm currently preserved in the CNRA gene bank (Dussert et al. 2003: 242). Among these populations, Cubry (2008) and Cubry et al. (2013b) genotyped seven populations using microsatellite markers with fresh samples provided by the CNRA gene bank. We used DNA of a few genotypes taken from these populations as controls for the diversity study.

MATERIAL AND METHODS

Material

We genotyped 162 samples of which 126 were herbarium specimens and 36 were obtained from previous works and used as controls (table 1). More details on the material are available in supplementary file 1 (worksheets A1 to A4).

We focused on herbarium specimens reported to be "spontaneous", or "wild", or "local" by the collectors. The collection dates range from 1905 to 1993. The collection points (red dots in fig. 1) are rather regularly distributed along a north-east – south-west axis from Bilima near Mamou, Guinea to Assikasso near Abengourou, Côte d'Ivoire, two locations 1000 km apart, with a gap in the Faranah Region, Guinea. In total, 40 herbarium specimens were collected from 15 different sites in Guinea and 86 herbarium specimens from 34 sites in Côte d'Ivoire.

As controls (purple triangles in fig. 1), we used DNA of *C. canephora* already extracted, characterized in previous studies (Cubry et al. 2013a, 2013b; Leroy et al. 2014), and stored at -20°C at CIRAD, Montpellier, France. We added leaves of a few clones maintained in the CIRAD coffee field gene bank in French Guiana (information available at: <http://florilege.arcad-project.org/fr/collections/collection-cafeiers-guyane>), collected in 2014, and stored in silica gel. The material used as controls (36 individuals) can be briefly characterized as follows:

Guinean group:

- Côte d'Ivoire (locations: Daloa, Divo, Fourougbankoro, Gohitafla, Ira, Kaloufla, Logbonou, Mouniandougou, and Pelezi): wild material collected from the end of the 1970s to 1991 (Berthaud 1984: 12; Le Pierrès et al. 1989; Couturon & Montagnon 1991).
- Guinea (location: Piné): population collected in 1987 in the forest of Piné on the Guinea-Côte d'Ivoire border (Le Pierrès et al. 1987).

Congolese group (not shown on the map):

- INEAC: population originally imported from the INEAC Lula and Yangambi research stations, Democratic Republic of the Congo. INEAC population was massively imported

Table 1 – List of the collection points, geographic coordinates and number of samples of *Coffea canephora* genotyped in the study. The samples are ordered according to the longitude coordinate (from the West to the East). Details are in supplementary file 1. (*) Exact collection point unknown (coordinates estimated); (**) Gamé cultivar collected in Sérédou gene bank by Portères (exact origin unknown).

Collection points	Number of samples	Sample ID	Latitude	Longitude
Total	162			
Herbarium specimens	126			
Guinea	40			
Bilima	10	1–10	10.35	-12.31
Massakoundou	2	11–12	9.18	-10.19
Makolo	2	13–14	9.04	-9.86
Wondékéré	2	15–16	9.12	-9.84
Boundo	2	17–18	9.04	-9.78
Bambaradou	1	19	8.57	-9.50
Macenta	1	20	8.55	-9.47
Sérédou_Gamé (**)	8	21–28	8.37	-9.29
Cercle_Macenta (*)	1	29	8.54	-9.20
Milo	1	30	8.98	-8.98
Koréla	5	31–35	8.64	-8.97
Nzérékoré	1	36	7.69	-8.82
Yarankadougou	1	37	8.40	-8.73
Beyla	2	38–39	8.69	-8.64
Foumbadougou	1	40	8.19	-8.42
Côte d'Ivoire	86			
Bangopolé	1	45	7.77	-7.94
Gueya	1	46	6.55	-7.84
Bafing	1	51	7.89	-7.67
Kiélé	2	52–53	7.51	-7.67
Dianlé	1	54	7.34	-7.59
Kodopleu	1	55	7.20	-7.58
Digoualé	1	60	7.73	-7.52
Touba	7	61–67	8.38	-7.50
Guiglo	1	68	6.54	-7.49
Kékobly	1	69	7.30	-7.43
Bahibly	2	70–71	7.18	-7.30
Kouibly	7	72–78	7.26	-7.24
Soubré	1	83	5.80	-6.60
Déma	1	84	7.30	-6.48
Gbétitapia	6	86–91	6.79	-6.45
Zaprégouhé	1	92	6.91	-6.37
Toubalo	1	93	7.72	-6.10
Prozibanfla	3	95–97	6.76	-5.89
Tos	6	99–104	6.94	-5.87
Asséyakro	7	105–111	7.67	-5.74
Totokro	1	112	7.96	-5.73
Kongonouassou	2	118–119	7.67	-5.60
Goudy	2	123–124	6.10	-5.06
Mbrimbo	2	125–126	6.03	-4.90
Mopri	1	127	5.80	-4.89
Kassa I	4	128–131	5.99	-4.79
Kassa II	2	132–133	6.04	-4.76
Sagassou	2	134–135	6.63	-4.62
Agbo	4	136–139	6.56	-4.20
Nzinzebelekro	2	140–141	7.38	-4.00
Abengourou	3	142–144	6.73	-3.48

Table 1 (continued) – List of the collection points, geographic coordinates and number of samples of *Coffea canephora* genotyped in the study.

The samples are ordered according to the longitude coordinate (from the West to the East). Details are in supplementary file 1. (*) Exact collection point unknown (coordinates estimated); (**) Gamé cultivar collected in Sérédou gene bank by Portères (exact origin unknown).

Collection points	Number of samples	Sample ID	Latitude	Longitude
Total	162			
Herbarium specimens	126			
Côte d'Ivoire	86			
Assikasso	4	145–148	7.09	-3.23
Kongodia	2	149–150	7.05	-3.23
Ébobo	3	151–153	5.37	-3.22
Controls	36			
A. Guinean group				
Guinea	4			
Piné	4	41–44	7.93	-8.15
Côte d'Ivoire	23			
Ira	4	47–50	7.76	-7.77
Mouniandougou	4	56–59	8.70	-7.57
Pélézi	4	79–82	7.29	-6.82
Daloa	1	85	6.89	-6.45
Kalouffa	1	94	7.52	-6.04
Gohitafla	1	98	7.49	-5.87
Fourougbankoro	5	113–117	8.36	-5.68
Logbonou	2	120–121	8.07	-5.25
Divo	1	122	5.79	-5.24
B. Congolese group				
DR Congo	7			
Luki (sub-group SG1)	3	154–156	N/A	N/A
INEAC (sub-group SG2)	4	159–162	N/A	N/A
Gabon via Togo/Bénin	2			
Niaouli (sub-group SG1)	2	157–158	N/A	N/A

into Côte d'Ivoire from 1935 on and to Guinea from 1952 on (Portères 1959, 1962). It belongs to the SG2 sub-group.

- Luki: genotypes collected in forests and plantations in the Mayumbe region, Democratic Republic of the Congo. They belong to the SG1 sub-group.

- Niaouli: population mainly cultivated in Togo and Benin, but originating from the Atlantic Coast of Gabon (Adibolo & Bertrand 1988). It was introduced into Côte d'Ivoire in 1914 (Portères 1934), and later to Guinea (Portères 1962). Niaouli coffee belongs to the SG1 sub-group.

Methods

DNA extraction – All the samples were processed at the Grand Plateau Technique Régional (GPTR) facility in Montpellier, France (<http://www.gptr-lr-genotypage.com>). Genomic DNA was extracted from approximately 20 mg of dried leaf tissue according to the MATAB (Mixed Alkyl Trimethylammonium Bromide) protocol described by Risterucci et al. (2000). The DNA concentration was estimated with a Fluoroskan Ascent Microplate Fluorimeter (ThermoFisher Scientific, Waltham, Massachusetts, USA) with a bisbenzimidazole DNA intercalator (Hoechst 33258) and by comparison with known standards of DNA.

Microsatellite markers – We used SSR markers defined from *C. canephora* and *C. arabica* genomes and already described (Poncet et al. 2004, 2007; Leroy et al. 2005; Cubry et al. 2008, 2013a). Because DNA is highly fragmented in herbarium specimens, we selected short size loci (less than 200 bp) as recommended by Särkinen et al. (2012) in order to increase the PCR success rates. Thirty-three markers were initially tested, but 10 were removed from the final analysis due to a high level of missing data or low genotyping quality (table 2).

PCR amplification – PCR reactions were performed as simplex experiments in a volume of 10 µl containing 1 µl of reaction buffer 10X (100 mM Tris-HCl, pH 9.0; 100 mM KCl 80 mM (NH₄)₂SO₄; X-100 1% Triton), 1 µl of 2 mM dNTP (Jena Bioscience GmbH, Jena, Germany), 0.3 µl of 50 mM MgCl₂, 0.08 µl of 10 µM forward primer with a M13 tail at the 5'-end, 0.1 µl of 10 µM reverse primer, 0.1 µl of fluorescently labelled M13-tail (6-FAM, NED, VIC or PET from Applied Biosystems, Foster City, California, USA), 0.12 µl of 5U/µl Taq DNA Polymerase (Taq'Ozyme OZYA001 from Ozyme, Montigny-le-Bretonneux, France), 5 µl of template DNA (1 ng/µl), and 2.3 µl of water. A touchdown cycling program was used as follows: an initial denaturation at 94°C for 5 min, followed by 10 cycles at 94°C for 30 s,

Table 2 – Main characteristics and references of the 23 SSRs markers used in the study.(*) accession number in EMBL-EBI database (European Bioinformatics Institute, <http://www.ebi.ac.uk>); (**) in SGN database (Sol Genomics Network, <http://solgenomics.net>).

TropGeneDB ID	Forward primer sequence	Reverse primer sequence	PCR product size	Repeat motif	Fluoro-chrome	PCR pool	Accession number	Reference
Me_BE_gA14_DL024	CGAGCGTAGAGGAACA	ACATGATACAAAGAATGGAGA	133	(AAT)4	VIC	M2	AJ871903*	Leroy et al. 2005
Me_ssr120291	GCTGCTAAAGCATCCGAAAC	CTTGGTAAACCTCAACCCGA	159	(AAG)8	6-FAM	M3	SGN-U614007**	Cubry et al. 2008
Me_ssr120568	TTGAGATTTTCAGCCGAAGG	AAACTGAATCCAAACGCATCC	117	(GA)16	6-FAM	M4	SGN-U619826**	Cubry et al. 2008
Me_ssr123106	GAGACGTGGTTCTGTGCTGTA	GTAAATCGCAGGCTAAAGCG	185	(AGA)5	NED	M4	SGN-U619140**	Cubry et al. 2008
Me_ssr123682	TCTCTTCCAATCCCAATGGC	TCCATAGTCCCCAAAAGCTG	126	(GAA)7	PET	M1	SGN-U613225**	Cubry et al. 2008
Mg_CMA0041	GCCTCTGTCGATCTTTCTA	GAGTCTTGTCGCCACCTTC	117	(ATG)8	NED	M3	AJ308802*	Poncet et al. 2004
Mg_M257	GACCATACATTTACACAC	GCAATTTGTGACACTGTA	103	(AC)8	6-FAM	M1	AJ250257*	Poncet et al. 2004
Mg_M312	TTGCGTAAAGAGAACGAGCA	GTGTGGAATAATCACCTGAGC	106	(TG)9 (AAAT)4	VIC	M5	AM408740*	Poncet et al. 2007
Mg_M344	CTGAGATGAAAAGGAAAAGG	TGAAAAAGAGCCAATCAAGT	132	(GT)9	6-FAM	M2	AM408766*	Poncet et al. 2007
Mg_M348	TGGCTATGGTGATGAAGAACT	GATGCGATCATATTCATTGT	92	(AC)10	NED	M5	AM408770*	Poncet et al. 2007
Mg_M356	TGAAGTCAACCTGAATACCAGA	ACGCACGCAGCAATG	175	(TG)14	NED	M5	AM231553*	Poncet et al. 2007
Mg_M364	AGAAAGATGAAGACGAAACACA	TAACGCCTGCCATCG	90	(A)21	6-FAM	M1	AM231556*	Poncet et al. 2007
Mg_M370	ATTGCCAGAGCATTGAAA	TGCAGTTGTAGTTAATCTCTCCT	93	(CA)8 (GA)5	NED	M2	AM408601*	Poncet et al. 2007
Mg_M441	TTAGCAATACAAATAGCCTTCC	GCATTTAGCCAGTTCGTGT	189	(AC)9	6-FAM	M2	AM408664*	Poncet et al. 2007
Mg_M449	GGTGGAACCGAAGATTGAG	CTTTGCCCTCTAAAGTTCTTGAAAC	99	(AC)8 (CA)9	VIC	M1	AM408670*	Poncet et al. 2007
Mg_M486	GGTATCGGTGGAGAAAAGTATG	CCAGCTATCAATGAAGACGA	136	(AC)8 (TA)5	NED	M2	AM408700*	Poncet et al. 2007
Mg_M494	ACTTTGTGGAGTAATTGACAGG	CCATTTTGGATTCATACGG	102	(AC)8	NED	M1	AM408708*	Poncet et al. 2007
Mg_M497	TGTTTGAATAATCCCATATCATAGA	TGGGCTCACATACGTTAATTG	101	(TG)7	6-FAM	M3	AM408710*	Poncet et al. 2007
Mg_M524	TCATGTCTGAATTTCTATTCTCTT	ACTTCTTTCCCTCTTATCACC	115	(TG)10	NED	M4	AM408735*	Poncet et al. 2007
Mg_M742	GGCTTCTTGGGTGCTGTGT	CCATTGGCTTTGTATTCTGG	110	(TG)7 (AG)6	VIC	M3	AJ308742*	Poncet et al. 2004
Mg_M779	TCCCCCATCTTTTCTTTCC	GGGAGTGTTTTTGTGTGCTT	116	(TG)17	6-FAM	M5	AJ308779*	Poncet et al. 2004
Mg_M783	CCAACTTCGTATGGTTGTCTG	TGATAGGAGGCACCTTGACACA	121	(TA)5 (TG)17	PET	M3	AJ308783*	Poncet et al. 2004
Mg_M850	AAGCCGTTTCAAGCCATA	GCAATTTCTGCTTCTGTCTC	159	(TA)6 (TG)11	6-FAM	M4	AJ308850*	Poncet et al. 2004

55°C for 60 s (0.5°C decrease at each cycle), 72°C for 1 min, followed by 25 cycles at 94°C for 45 s, 50°C for 1 min, 72°C for 1 min and a final extension at 72°C for 30 min.

Fluorescently labelled PCR products were then organized in five pools for electrophoresis, using respectively 2 µL of products labelled with 6-FAM, 2 µL of those with VIC, 2.5 µL of those with NED and 3.5 µL of those with PET, and completed at 20 µL with high purity water. Table 2 shows which fluorochrome was used to label each marker and how PCR pools were composed. 2 µL of this solution was taken and added to 10 µL of Hi-Di formamide and 0.12 µL of GeneScan 600 LIZ size standard (Applied Biosystems). Migration of PCR products was made on an ABI 3500xL Genetic Analyzer (Life Technologies, Carlsbad, California, USA). Alleles were scored using GeneMapper v.4.1 software (Applied Biosystems).

For some low-quality samples resulting in poor PCR amplification, we used a modified version of this protocol. After extraction, genomic DNAs were purified with Agencourt AMPure XP (Beckman Coulter, Brea, California, USA) magnetic beads with 1 volume of DNA for 1.8 volume of beads, then standardized at 0.5 or 1 ng/µL with high purity water. Then, we performed PCR in 10 µL reaction volume using Qiagen Type-it Microsatellite PCR Kits with 5 µL of PCR MasterMix (HotStarTaq Plus DNA polymerase, PCR buffer, dNTP Mix), 2 µL of purified DNA (0.5–1 ng), 0.2 µL of primer mix (0.1 µL of 10 µM forward primer with an M13 tail at the 5'-end, 0.1 µL of 10 µM reverse primer), 0.2 µL of fluorescently labelled M13-tail (6-FAM, NED, VIC or PET from Applied Biosystems), and 2.7 µL of high purity water. The PCR amplification was conducted with an initial denaturation at 95°C for 5 min, followed by 10 cycles of 95°C for 30 s, 55°C for 1 min 30s (0.5°C decrease at each cycle), 72°C for 30 s, followed by 25 cycles at 95°C for 30 s, 50°C for 1 min 30 s, 72°C for 30 s and a final extension at 60°C for 30 min. Dilution, migration of PCR products, and allele scoring were made as previously described.

Structure analysis – To identify the pair-wise genetic relationships between the individual genotypes, we computed a genetic dissimilarity matrix using simple matching index with DARwin v.6 software (Perrier & Jacquemoud-Collet 2006). An overall representation of the structure of genetic diversity was obtained by a factorial analysis (Principal Coordinates Analysis, PCoA) using distance matrices, while genetic relationships between individuals were assessed using the neighbor-joining (NJ) method (Saitou & Nei 1987), as implemented in DARwin v.6.

In order to test for sample clustering and to estimate admixture proportions for each individual, we used the model-based Bayesian approach implemented in STRUCTURE v.2.3.4, (Pritchard et al. 2000). The parameters were set to a burn-in period of 100,000 with 200,000 iterations. We performed 20 independent runs for each K, K varying from 1 to 10. We used the online servers CLUMPAK (Kopelman et al. 2015) and STRUCTURE HARVESTER (Earl & vonHoldt 2012) to analyse the STRUCTURE outputs and the method of Evanno et al. (2005) to identify the number of genetic clusters corresponding to the uppermost hierarchical level of genetic partitioning between populations.

Firstly, we analysed the data of all the samples of both countries, assigned each sample to Congolese or Guinean groups and estimated admixture rates. Then, we removed all the Congolese samples as well as admixed ones (we chose a threshold of 95% for membership coefficient) and performed the same analyses in order to reveal a possible structure within the Guinean group.

Diversity parameters – We calculated several descriptors of genetic diversity: allele number per marker, observed heterozygosity (H_o), and expected heterozygosity (H_e), and performed a test of Hardy-Weinberg (HW) equilibrium, using the R package STRATAG (Archer et al. 2017). These statistics were computed for different sets of genotypes defined according to the structure analysis. We used the R package HIERFSTAT (Goudet 2005) to calculate allelic richness, corrected for the different sample size of populations, using a rarefaction method as recommended by El Mousadik & Petit (1996). We also computed the number of private alleles for the different sets of genotypes. Lastly, three different measures of population differentiation (Weir's F_{st} , G_{st} , and Jost's D), as well as their significance, were calculated between each genetic group or sub-groups using the R package STRATAG: Congolese group vs. Guinean group and between all Guinean sub-groups as defined by the structure analysis.

RESULTS

Structure analysis

Whole sample – The genotyping data matrix of the 162 samples may be found in supplementary file 1 (worksheet B). The results of the factorial analyses of the SSR-based dissimilarity matrix are represented in fig. 2. The first axis enabled a clear separation of the Guinean and Congolese groups with 19.68% of the global inertia. The second axis (5.88% inertia) and the third axis (4.75% inertia) separated the SG1 and SG2 Congolese sub-groups. Most of the herbarium specimens were distributed in the Guinean group. However, some herbarium specimens were grouped with the Congolese controls, and others occupied an intermediate position between Guinean and Congolese groups, which allowed us to classify them as admixed genotypes. The NJ tree (supplementary file 2) confirmed that classification. Similarly, Bayesian analysis using STRUCTURE clearly differentiated two clusters corresponding to Guinean and Congolese groups (fig. 3 and supplementary file 3).

Examining these results in detail and by collecting country (supplementary files 2 & 3), we found that:

- (i) in Guinea, eight herbarium specimens (numbers 11 to 18) collected in the villages around Kissidougou grouped together with the Congolese genotypes (SG1 or SG2 sub-groups) used as controls, while all other samples were assigned to the Guinean group (membership coefficient $Q > 95\%$);
- (ii) in Côte d'Ivoire, three herbarium specimens from Ébo-bo (numbers 151, 152, and 153) collected by Portères near Aboisso grouped together with the Niaouli and Luki populations (SG1 Congolese sub-group);
- (iii) one herbarium specimen (number 142), collected in 1930 near Abengourou and stated to be “spontaneous” by

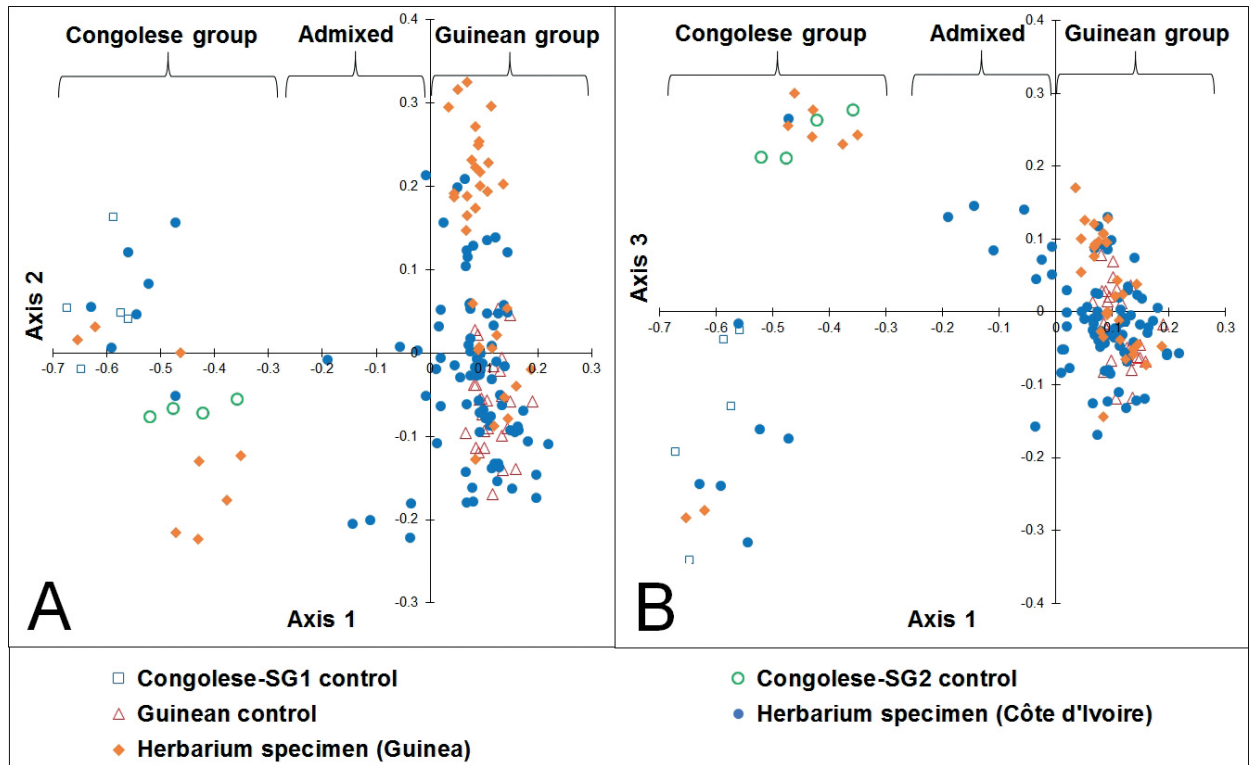


Figure 2 – Factorial analysis (PCoA) of the simple matching distance matrix of the 162 samples of *Coffea canephora*: 126 herbarium specimens collected in Guinea and Côte d'Ivoire and 36 genotypes used as controls. Representation of the first three axes: axis2/axis1 (A); axis3/axis1 (B).

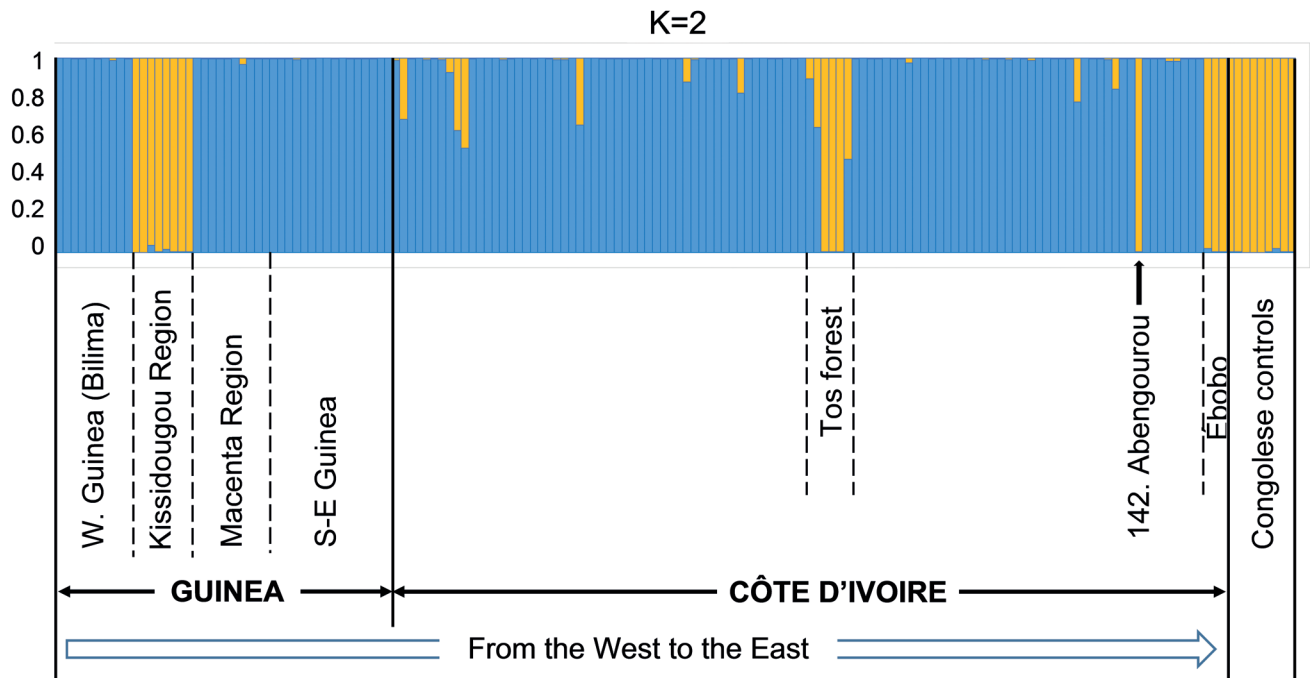


Figure 3 – Results of STRUCTURE analysis for the 162 samples of *Coffea canephora*: 126 herbarium specimens collected in Guinea and Côte d'Ivoire and 36 genotypes from previous studies used as controls. Population assignment for K = 2. Each individual is represented by a thin vertical line, which is partitioned into two coloured segments that represent the individual's estimated membership fractions in the two clusters.

Chevalier, clearly belongs to the SG1 Congolese sub-group ($Q > 99.5\%$).

(iv) unlike the herbarium specimens collected in Guinea, some of those collected in Côte d'Ivoire are Guinean-Congolese admixed genotypes. The results of STRUCTURE program ($K = 2$) showed that 12 herbarium specimens have a membership coefficient that varies from $Q = 47.9\%$ (probable F_1 hybrid between Congolese and Guinean genotypes) to $Q = 95\%$. For example, Portères collected six plants “in a plantation established from local seeds and seedlings collected in the forest of Tos” near Daloa (Centre West of Côte d'Ivoire). Among them, we found three pure Congolese ($Q > 99\%$ for numbers 101, 102, and 103), one putative F_1 hybrid ($Q = 52\%$ for number 104), and two admixed ($Q = 10.8\%$ and 35.7% for numbers 99 and 100 respectively).

Within Guinean group – Once we removed the Congolese, putative Guinean-Congolese F_1 hybrids, and other admixed genotypes ($Q < 95\%$), there were 126 samples of probable Guinean origin left (99 herbarium specimens and 27 controls) that we reanalysed with the same tools. The results of factorial analysis of the matrix of the Guinean group samples did not reveal a strong structure (fig. 4). However, the first axis (9.98% of the global inertia) enabled the separation of all the specimens of *C. canephora* var. *maclaudii* near Bilima and Gamé cultivar near Macenta from more eastern origins. The second axis (5.37%) separated *maclaudii* specimens from Gamé specimens. In accordance with the factorial analysis, the results of the Bayesian assignment for $K = 2$ (best K according to Evanno's test) reflected the separation of the *maclaudii* and the Gamé specimens from more eastern origins (fig. 5 and supplementary file 4). From $K = 5$, the *maclaudii* and the Gamé specimens were assigned to two distinct ancestry sub-groups with little admixture. The other specimens collected from South East Guinea to East Côte d'Ivoire were distributed into three other clusters with a significant level of admixture and no strong geographical support.

Genetic diversity of the Guinean and Congolese groups

Once we removed the admixed individuals based on STRUCTURE assignment threshold defined above (i.e. removal of individuals having $Q < 95\%$), we calculated the diversity parameters of the 126 individuals of the Guinean group and the 24 individuals of the Congolese group. We found the Congolese group more diverse than the Guinean one with higher expected and observed heterozygosities (mean $He = 0.67$ and mean $Ho = 0.51$ for Congolese while mean $He = 0.48$ and mean $Ho = 0.34$ for Guinean, supplementary file 5, worksheet 2) and higher rarefied allelic richness per marker (mean of 6.87 for Congolese and mean of 4.21 for Guinean, supplementary file 5, worksheet 10). Several markers were found not to be at HW equilibrium within both groups (16 and 21 out of 23 markers for Congolese and Guinean groups respectively). We also found a high number of private alleles differentiating the two groups, with 79 alleles found within the Congolese group and 54 within the Guinean group (supplementary file 5, worksheet 9). The three complementary differentiation statistics computed were consistently high and highly significant between the two groups (supplementary file 5, worksheet 11).

Genetic diversity and measure of population differentiation between STRUCTURE-defined Guinean genetic sub-groups

We assigned genotypes to sub-groups based on within Guinean group STRUCTURE analysis at $K = 5$ by considering a threshold of 75% of ancestry. This left us with 91 samples (out of 126) clustered into five different Guinean sub-groups (sgG1 to sgG5) of 10, 9, 18, 29, and 25 genotypes respectively, the remaining 35 samples having admixed ancestry (supplementary file 5, worksheet 3). Mean number of alleles per marker ranged from 1.61 (sgG1) up to 4.04 (sgG3). Observed heterozygosity and expected heterozygosity remained rather high, with values ranging from 0.16 to 0.41 (Ho) and from 0.16 to 0.5 (He) depending on sub-groups. The number of markers found not to be at HW equilibrium within sub-groups ranged from 1 to 11. The number of private alleles detected was moderate, ranging from 2 to 19 depending on the sub-group considered (supplementary file 5, worksheets 4 to 9). Genetic diversity assessed through rarefied allelic richness ranged from 1.64 to 3.38, with the lowest value for sgG1 as compared to the other groups (supplementary file 5, worksheet 10). As for the measures of genetic divergence between sub-groups, they were rather high for F_{st} and G_{st} and moderate for Jost's D but all were highly significant, whichever the pairwise comparison considered (supplementary file 5, worksheet 11).

Based on the results of the STRUCTURE analysis of the Guinean group detailed above, we represented the geographic

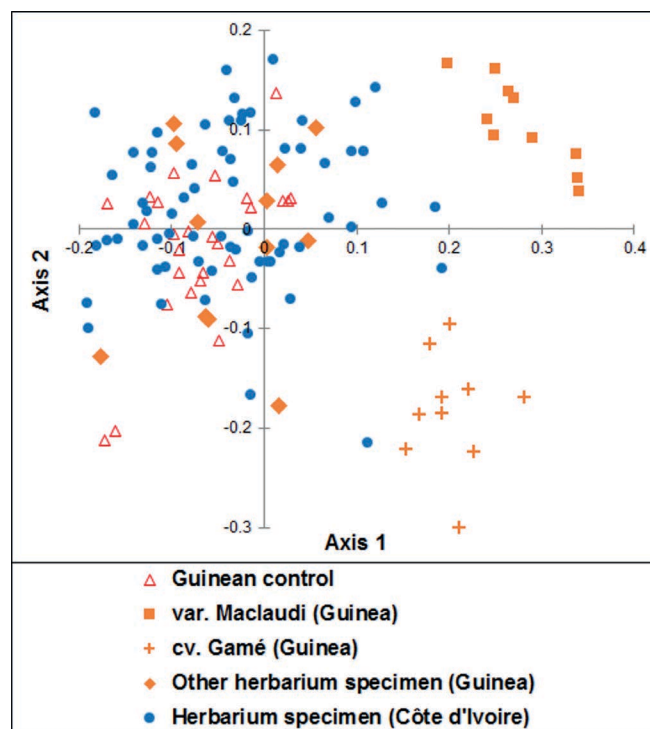


Figure 4 – Factorial analysis (PCoA) of the simple matching distance matrix of 126 samples (99 herbarium specimens and 27 controls) of *Coffea canephora* collected in Guinea and Côte d'Ivoire and belonging to the Guinean group. Representation of the first two axes.

distribution of the five Guinean sub-groups on a map (fig. 6). The first sub-group sgG1, located in the south-west of Fouta Djallon, corresponds to *Coffea canephora* var. *maclaudii* with 10 specimens grouped into a single cluster with very little admixture from other groups. The second sub-group sgG2 corresponds to the Gamé cultivar located near Macenta. The last sub-groups, less differentiated, are located in the south of Côte d'Ivoire near the towns of Divo and Tiassalé (sgG3), and in the west and centre west of Côte d'Ivoire (sgG4 and sgG5, respectively).

DISCUSSION

The most recent comprehensive reviews of *C. canephora* in Guinea and Côte d'Ivoire, by Portères, date back to the end of the 1950s (Portères 1959, 1962). He classified coffee plants as local or introduced based on historical records (e.g. gene bank entry books), habitat description, phenotypic observations, and by questioning planters and villagers during his collection missions. By genotyping old herbarium specimens with nuclear microsatellites and appropriate controls, we were able to determine whether each specimen was native or introduced, and, if applicable, to estimate the level of admixture. In the light of our results, Portères' observations appear to be remarkably precise and correct. Two notable exceptions are specimens collected around Kissidougou in Guinea and those collected near Aboisso (Ébobo population) in Côte d'Ivoire, which both grouped with Congolese genotypes whereas their collectors considered them native to the country.

Pattern of diversity of genotypes belonging to the Guinean group

Within the Guinean group, our analyses revealed five sub-groups distributed over the two countries (fig. 6) and variable levels of admixture between them.

In Guinea, two sub-groups were clearly distinguished from the others, suggesting restricted gene flow probably due to their geographical isolation. The first sub-group (sgG1) corresponds to *Coffea canephora* var. *maclaudii* determined by Chevalier in 1905 near Bilima. The 10 analysed specimens (numbers 1 to 10), collected between 1905 and 1993, grouped into a single cluster with very little admixture from other groups. This is the most western and the most northern natural habitat of *C. canephora* ever found in Africa. The high plateaus of the Fouta Djallon form a natural barrier separating the Southern Rivers basin from the upper basin of Bafing-Sénégal (Boulvert 2003: 43). Bilima, located west of Fouta Djallon, is exposed to rains (about 2000 mm per year) coming from the coast and has a climate that is suitable for coffee. The only known site where coffee was observed is a forest island, approximately 1400 m long and 350 m wide, which stretches down over the western slope of the Bilima-Hénéré plateau, south of Bilima village, and is surrounded by cultivated fields and fallow land. This diversity group also had the lowest allelic richness observed in our study, suggesting both its peculiarity and vulnerability.

A second Guinean sub-group (sgG2) corresponds to the Gamé cultivar described by Portères (1962). Bilima and Macenta are separated by the southern prolongation of the Fouta

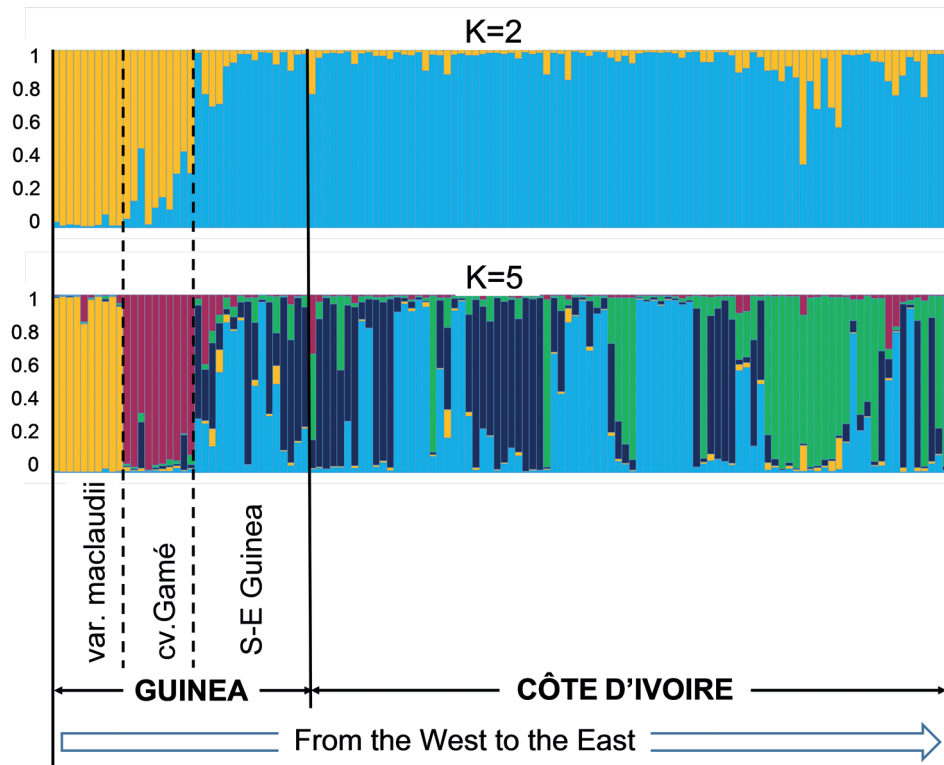


Figure 5 – Results of STRUCTURE analysis for 126 samples of *Coffea canephora* belonging to the Guinean group. Estimated population structure for $K = 2$ and $K = 5$. Each individual is represented by a thin vertical line, which is partitioned into K coloured segments that represent the individual's estimated membership fractions in K clusters.

Djallon to Mamou and by a forest-savanna transition zone around Faranah and Kissidougou over a distance of about 300 km. According to Portères, the area of origin of Gamé is limited to the plain near Macenta and the surrounding hills, west of the Ziama massif. The only herbarium specimen (number 19) taken from the sub-group's natural habitat was collected by Barthe in 1936 in Bambaradou, a village located 3 km from Macenta, west of the Ziama massif, and which was considered by Portères as the birth place of Gamé. The other specimens were collected by Portères in 1958 either from Barthe's collection established in Macenta (number 20) or from the plantation that Fora Camara, an assistant of Portères, set up near the village of Sérédou, east of the Ziama massif (8 genotypes of original collection whose location is unknown).

Separated from the Gamé sub-group in the west by the Ziama massif, the other samples are distributed over a vast area extending from the south of the prefecture of Kérouane, Guinea to the far east of Côte d'Ivoire. The factorial and Bayesian analyses showed low differentiation and weak grouping among these individuals. Several factors could account for this situation. In terms of ecology and topography,

there is no significant natural barrier that could restrict gene flow between coffee populations in that vast area. Moreover, from the first quarter of the twentieth century, there was a dramatic expansion of the plantations and intense movements of planting material throughout both French colonies from eastern Côte d'Ivoire to the Guinea Forest Region. This movement may have included coffee with a wide range of admixture levels in locations characterized by plantations and forests situated in proximity. However, the geographical distribution of the sub-groups shown in fig. 6 suggests that one sub-group (sgG3) could be located in the south of Côte d'Ivoire near the towns of Divo and Tiassalé. The last two less differentiated groups (sgG4 and sgG5) are located in the west and centre west of Côte d'Ivoire.

For the very first time, despite the limited number of markers (23) we used, compared with the number used in previous studies, e.g. 108 SSR markers in Cubry et al. (2013a), we found significant evidence for population genetic structure within the Guinean group of *C. canephora*. The structure detected is well organized spatially, especially in Guinea, and may reflect diverse local adaptation or a long history of isolation of the populations. Further investigations are needed

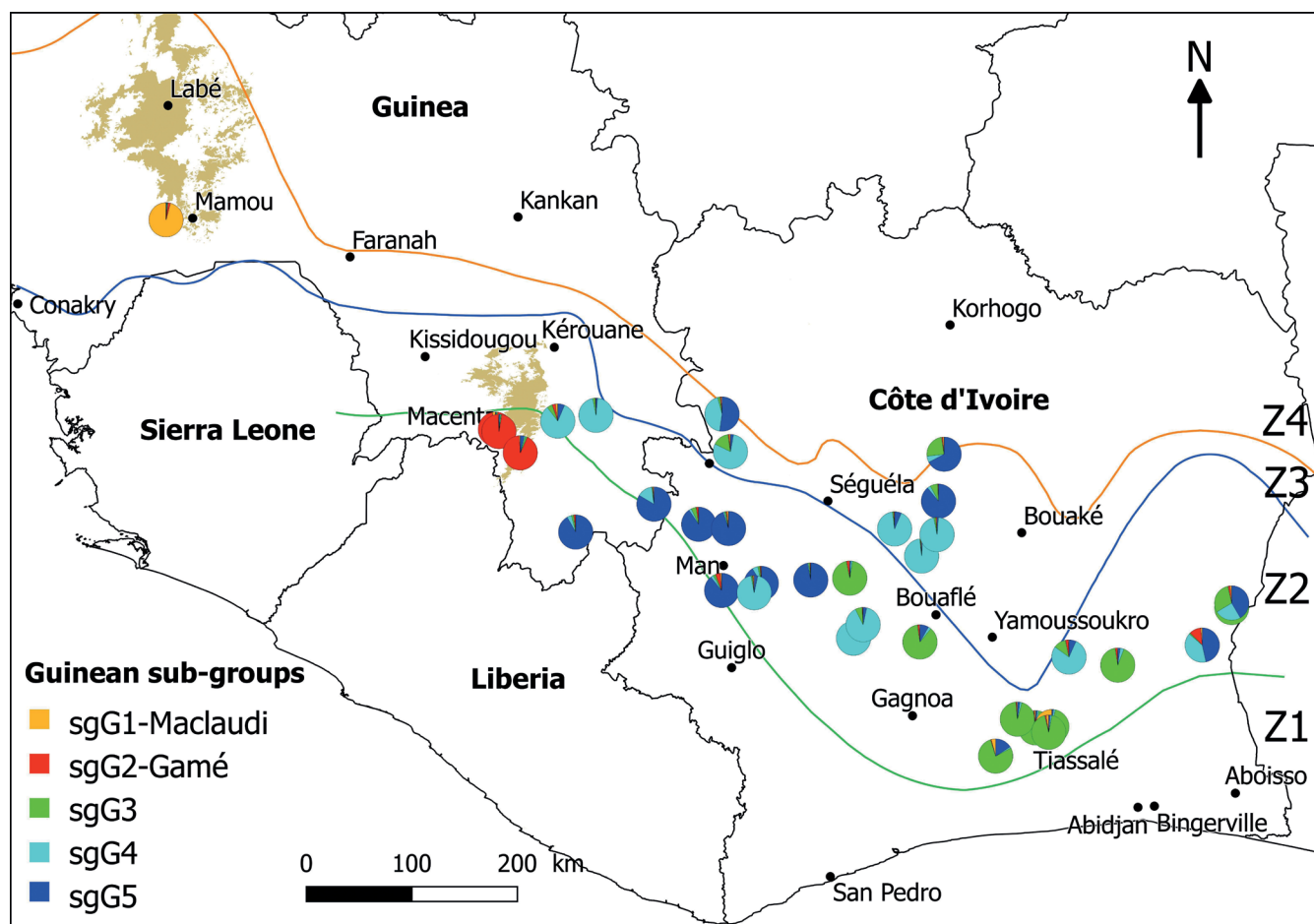


Figure 6 – Geographical distribution of Guinean sub-groups of *Coffea canephora* (pie charts). Assignment in STRUCTURE analysis at $K = 5$ assumed groups ($Q \geq 75\%$ membership). The four main vegetation zones are delimited by coloured lines: Guinean domain (Z1: moist evergreen forest, widely degraded; Z2: moist semi-deciduous forest); Z3: transition zone; Z4: Sudanian domain. The lines for vegetation zones are drawn after distribution data presented in maps by Guillaumet & Adjanohoun (1971) for Côte d'Ivoire and Boulvert (2003) for Guinea.

to confirm these results and to elucidate the precise origin of the structure (isolation by distance, gene flow barriers, past climatic history, etc.).

Admixture between Guinean and Congolese groups

We focused on herbarium specimens reported to be “spontaneous”, or “wild”, or “local” by their collectors. However, we found several Guinean-Congolese admixed genotypes, mainly among the specimens collected by Portères in Côte d’Ivoire at the end of the 1950s, e.g. the ones collected from Tos plantation (see supplementary file 3). At that time, Congolese germplasm was predominant in the plantations after massive imports from the 1930s. The results of our genotyping study underline the interest of the large herbaria collected by Portères to validate information related to the history of introductions and movements of coffee in West Africa we retrieved in our literature review.

The case of two populations of the Congolese group

Robusta Kissidougou – Passing the barrier of Fouta Djallon and progressing to the southeast, coffee is cultivated in the prefecture of Kissidougou in the forest-savanna transition zone, mainly in the island forests that surround the villages. We found that the eight herbarium specimens of our study collected in 1959 from four villages around Kissidougou (Boundo, Makolo, Wendékéré, Massakoundou) belong to the Congolese group (SG1 and SG2 sub-groups), although Portères considered them to derive from local material. According to a handwritten note kept at MNHN, he located the discovery of a local *C. canephora* by Sagna Camara, an agriculture field assistant, between 1927 and 1930, in a forest near Boundou and the dissemination of that material only from the end of the 1940s. However, our study showed that the coffee plants sampled by Portères most probably belong either to the population Niaouli (Massakoundou specimens) or to the population that he called *Robusta Congo Belge*. The latter was introduced for the first time in 1914 in Côte d’Ivoire from the botanical garden of Eala (Democratic Republic of the Congo). This population, which is genetically close to the INEAC population, was widely disseminated in Côte d’Ivoire from 1922 and since the end of the 1920s in Guinea, especially in the Kissidougou region (Portères 1934, 1959, 1962; Cordier 1961). According to Fairhead & Leach (1996) “coffee became more central to Kissidougou’s economy from the 1930s. In the context of the high coffee prices of the late 1940s and 1950s farmers increased their coffee production dramatically, filling virtually all available forest island space with plantations and extending forest islands to harbour new plantations where possible.” Therefore, the existence of coffee of the Guinean group originating in the region of Kissidougou, as stated by Portères, remains to be demonstrated.

Robusta Ébobo – In the extreme south east of Côte d’Ivoire, the Plantation *Ébobo* was created by the Société des Plantations de la Tanoé in 1925 near the village of the same name on the Abi lagoon. According to Portères’ handwritten notes conserved at MNHN, this plantation sent seeds to the Plantation de La Bia, an estate belonging to the Société des Plantations d’Elima (SPE) located at the mouth of the La Bia

River, south of the town of Aboisso (Lemery 1937). In 1958, Portères collected 20 herbarium specimens of robusta Ébobo in the Plantation de La Bia and sent them to PAT. According to Portères (1959) and Cordier (1961), the robusta Ébobo may derive from plants native to the Montagnes Bleues (70 to 100 m a.s.l.) located between the villages of Ébobo and Elima (Laplante & Rougerie 1949). In another handwritten note, Portères mentioned a plot of robusta Ébobo established in 1935 on the Plantation *Ébobo* and inferred 1933–1934 as the likely date of discovery of its natural habitat. We genotyped three herbarium specimens (numbers 151, 152, and 153) and found them to belong to the Congolese group. They are genetically close to the Niaouli and Luki populations that belong to the SG1 sub-group. These results are in accordance with those of previous studies on several clones of robusta Ébobo by Berthaud (1984), Adibolo & Bertrand (1988), Montagnon et al. (1992b), and Dussert et al. (2003), Dussert using the term “Aboisso” instead of “Ébobo”. Therefore, there is little evidence to support Portères’ hypothesis concerning a local origin of robusta Ébobo. This coffee may have been introduced earlier from Lower Guinea, possibly by Beynis, a European planter who established his plantation near Aboisso using seeds of Niaouli and *Café du Kouilou* imported from Gabon around 1910 (Cordier 1961). Bodard (c. 1960, CRA Bingerville, Côte d’Ivoire, unpubl. res.) mentioned other introductions of *Café du Kouilou* from Gabon between 1917 and 1927 that were mainly disseminated in the coastal region, particularly to the Plantation de La Bia.

The same early imports could explain why the herbarium sample number 142, collected in 1930 by Chevalier near Abengourou, although labelled as “spontaneous”, is actually a specimen of the Niaouli population.

Potential sites for further collection of Guinean germplasm

As shown in fig. 6, most of the native *C. canephora* populations collected in the past in Côte d’Ivoire and Guinea were found in the vegetation zone (Z2) characterized by semi-deciduous moist forests, according to the classification introduced during the Conference of Yangambi in 1956 (Aubréville 1957; Guillaumet & Adjanooun, 1971: 190). As already mentioned by Portères (Portères 1937a, 1937b) and later by Berthaud (1984: 140), a number of populations were also found in a transition zone (Z3) where forest and savanna co-exist in a complex mosaic of forest islands and gallery forests along the rivers.

However, since the middle of the twentieth century, population pressure and immigration, illegal or excessive logging, land encroachment and extension of coffee and cocoa plantations, and inefficient public forest administration have led to the destruction or fragmentation of large forest areas. This historical trend has mainly been documented in Côte d’Ivoire (Brou et al. 1998, 2005b). According to the conservative estimate by Fairhead & Leach (1998: 40), in 1900, Côte d’Ivoire may have had a forest cover of 7 to 8 million hectares that remained stable until 1955, but only 2.7 million in 1990. The U.S. Geological Survey Earth Resources Observation and Science (USGS EROS) Centre (<http://eros.usgs.gov/westaf-rica/>) has mapped the land cover of West Africa for three pe-

riods in time (1975, 2000, and 2013) using many hundreds of Landsat images. By 2013, Côte d'Ivoire had lost nearly 60% of the 3.7 million hectares of dense tropical forests that existed in 1975. Similarly, degraded forest decreased by 28% and woodland area declined by 48%. In Guinea, during the same period, dense forest areas decreased by about 33% to only 0.4 million hectares. Recent studies have started to document the impact of the political-military crisis faced by Côte d'Ivoire between 2001 and 2013, on the 182 protected forests ("forêts classées" in French) of that country. Before the conflict, the forest of Haut-Sassandra, in which a coffee population (Pélési, numbers 79 to 82) was collected in 1986, was one of the best protected forest reserves in Côte d'Ivoire with 93% of forest cover. In 2015, the same forest cover has decreased to less than 28%, mainly due to cocoa planting and tree logging that followed the withdrawal of the forest ranger force (Barima et al. 2016).

Both countries have also been affected by periods of severe drought since the end of the 1960s, in step with what has been observed in the Sahel. The trend intensified during the 1980s and 1990s before a remission in the 2000s (Lubès et al. 1995; Brou et al. 2000, 2005a). Moreover, recent studies investigated the potential impact of global climate change on the natural habitats and cultivation areas of several species of *Coffea* (Davis et al. 2012, 2019; Bunn et al. 2015). In a few decades, changes in climate will reduce the global area suitable for coffee, leading to severe stress and a high risk of extinction. Even if the species *C. canephora* (as a whole, including the Congolese group) is not considered as critically endangered, the populations of the Guinean group growing in the transition zone between savanna and forest (Z3) in Guinea and Côte d'Ivoire will likely be the first affected. Lastly, as important as forest degradation, the dilution of the Guinean genetic pool resulting from hybridization with coffee belonging to the Congolese group, which constitutes the major part of the coffee plantations in Guinea and Côte d'Ivoire, could also hamper new collections of native coffee plants in both countries.

Despite those factors, a few populations of coffee trees belonging to the Guinean group were successfully collected in the last decades of the twentieth century in Côte d'Ivoire and Guinea (Le Pierrès et al. 1989; Couturon & Montagnon 1991; Montagnon et al. 1993). Berthaud (1984: 140) mentioned the presence of wild coffee in "sacred forests" that are preserved to host initiation ceremonies. These sites often contain vast biodiversity with abundant fauna and flora (Yao et al. 2013). Recent studies on farming systems in the Guinea Forest Region have highlighted the expansion of agroforests surrounding the villages (Fairhead & Leach 1996; Camara et al. 2009; Correia et al. 2010). In these agroforests, coffee, cocoa and kola are associated with various forest tree species. However, the genetic identity of the coffee populations found in these agroforests remains unknown.

Our findings on the genetic structure of *C. canephora* in Guinea and Côte d'Ivoire will guide further surveys in both countries. The use of recent aerial and satellite images will provide preliminary information on the current state of vegetation at past collection points. On the spot, in addition to coffee sampling, it is advisable to record detailed information about the ecology and associated flora, as well as a descrip-

tion of the coffee populations (age structure, sanitary status, etc.) and of the environmental and social factors that could affect the future of the forest. For sampling, priority should be given to germplasm that is not or insufficiently represented in *ex situ* gene banks i.e. var. *maclaudii*, cultivar Gamé, and populations from South-East Guinea, and, in Côte d'Ivoire, forest remnants in the south (Toumodi, Kassa, and Mopri) and in the east (Abengourou and Agnibilekrou). The gene banks belonging to CNRA and IRAG, once supplemented with new material from the Guinean group, will be major resources for the reactivation of coffee breeding activities, especially given the increasing need for adapted germplasm in the context of climate change. As stated by Davis et al. (2019), wild variants of coffee species will be of primary importance in the future, especially with the increasing incidence and duration of drought and the emergence or spread of diseases and pests. This is particularly the case for populations located at the margins of the environmental range of the species, like var. *maclaudii* in the extreme northwest and populations located in the transition zone (Z3). These particular locations might be associated with specific local adaptations of potential interest for breeding purposes. A more in-depth analysis including phenotypic evaluation of the genotypes from these locations should be undertaken to detect particular traits associated with their specific history and ecological growing conditions.

CONCLUSIONS AND FURTHER STUDIES

The analysis of genetic data of herbarium specimens, combined with information extracted from historical records, herbarium documentation and scientific literature, allowed us to validate (or, in a very few cases, invalidate) previous observations made by botanists regarding the history of introductions and movements of robusta coffee germplasm in Guinea and Côte d'Ivoire.

We explored the genetic diversity of *C. canephora* in these two countries by genotyping old herbarium specimens with 23 nuclear microsatellites and appropriate controls. We were able to determine the genetic group (Congolese or Guinean) of each specimen, and, if applicable, to estimate the level of admixture. We identified several genotypes belonging to the Congolese group, introduced with the development of commercial coffee plantations from 1920, and diverse Guinean-Congolese hybrids resulting from natural hybridization between local and introduced genetic stocks. The results of factorial and Bayesian analyses provided significant evidence for population genetic structure within the Guinean group of *C. canephora*. The Guinean genotypes can be assigned to five sub-groups with distinct spatial distribution, especially in Guinea where two sub-groups, corresponding to the var. *maclaudii* and the cultivar Gamé, are characterized by a low level of admixture, which suggests restricted gene flow probably due to geographical distances and the natural barriers formed by the Fouta-Djallon and the Ziama massifs.

Because of past and current trends in forest degradation, there is an urgent need to describe and protect potential collection sites in West Africa and to evaluate the genetic diversity of the coffee populations they contain. Collecting *C.*

canephora genotypes of the Guinean group will serve as a valuable resource for future genetic improvement of robusta coffee. Due to its limited genetic diversity compared to other Guinean sub-groups, the population of *C. canephora* var. *maclaudii* near Bilima may be considered vulnerable and a priority target for conservation.

The following step could be to extend genetic analysis to a larger number of specimens taken from MNHN herbaria (P and PAT) and other European herbaria (e.g. herbarium LISU, Lisbon, Portugal and BR, Meise Botanic Garden, Belgium). With recent DNA technologies like next generation sequencing methods (chloroplast sequencing, gene capture, etc.), already successfully applied to herbarium specimens (e.g. Staats et al. 2013; Hart et al. 2016; Suchan et al. 2016), it is now possible to conduct in-depth investigation of diversity within a species, at reasonable cost. Provided the herbarium specimens are accurately documented, a detailed and comprehensive study of the species *C. canephora*, including Congolese and Guinean groups and their sub-groups, becomes an achievable goal.

SUPPLEMENTARY FILES

Five supplementary files are associated to this paper:

Supplementary file 1: detailed information on material used and genotyping data matrix of the whole sample (Excel file)
<https://doi.org/10.5091/plecevo.2020.1584.2033>

Supplementary file 2: neighbor-joining tree of the whole sample (pdf)
<https://doi.org/10.5091/plecevo.2020.1584.2035>

Supplementary file 3: results of STRUCTURE analysis of the whole sample (pdf)
<https://doi.org/10.5091/plecevo.2020.1584.2037>

Supplementary file 4: results of the STRUCTURE analysis of the Guinean group samples (pdf)
<https://doi.org/10.5091/plecevo.2020.1584.2039>

Supplementary file 5: genetic diversity parameters and measures of population differentiation (Excel file)
<https://doi.org/10.5091/plecevo.2020.1584.2041>

ACKNOWLEDGMENTS

The authors thank the Muséum National d'Histoire Naturelle, Paris and Myriam Gaudeul for providing the specimen samples from the P herbarium, and Serge Bahuchet, Julien Blanc, and Simon Juraver (UMR 7206 Eco-Anthropologie, <http://ecoanthropologie.cnrs.fr>) for giving access to Portères' coffee specimens and documentation at the PAT herbarium. We also thank Pierre Charmetant for providing samples from Cirad gene bank in French Guiana. We are most grateful to Pierre Mournet and Hélène Vignes at the GPTR platform in Montpellier, France for help in genotyping and to Xavier Perrier and Fabien De Bellis in data analyses. We also thank the reviewers and the editors for their thorough editing and their helpful comments. H.A. Nguyen's scholarship was supported by a grant from the CULTIVAR project funded by Agropolis Fondation (ID 1504-007) through the Investissements

d'Avenir program (Labex Agro:ANR-10-LABX-0001-01) under the frame of I-SITE MUSE (ANR-16-IDEX-0006).

AUTHOR CONTRIBUTIONS

J.-P. Labouisse and P. Cubry conceived and designed the research. R. Rivallan, H.A. Nguyen, and J.-P. Labouisse performed the laboratory works. J.-P. Labouisse, P. Cubry, and F. Austerlitz analysed and interpreted the data. J.-P. Labouisse, P. Cubry, F. Austerlitz, and R. Rivallan wrote the manuscript.

REFERENCES

- Adibolo Y., Bertrand B. (1988) Étude de l'origine de la variété Niaouli au Togo et Bénin. *Café Cacao Thé* 32(4): 293–298.
- Anonymous (1934) Nouvelles de l'école et des anciens élèves. Nominations et mutations. *L'Agronomie coloniale* 197: 154.
- Anthony F., Dussert S., Dulloo E. (2007) Coffee genetic resources. In: Engelmann F., Dulloo M.E., Astorga C., Dussert S., Anthony F. (eds) Conserving coffee genetic resources: complementary strategies for *ex situ* conservation of coffee (*Coffea arabica* L.) genetic resources: a case study in CATIE, Costa Rica: 12–22. Rome, Bioversity International.
- Archer F.I., Adams P.E., Schneiders B.B. (2017) StrataG: An R package for manipulating, summarizing and analysing population genetic data. *Molecular Ecology Resources* 17(1): 5–11. <https://doi.org/10.1111/1755-0998.12559>
- Aubréville A. (1957) Accord à Yangambi sur la nomenclature des types africains de végétation. *Bois et Forêts des Tropiques* 51: 23–27.
- Avelino J., Cristancho M., Georgiou S., Imbach P., Aguilar L., Bornemann G., Läderach P., Azueto F., Hruska A.J., Morales C. (2015) The coffee rust crises in Colombia and Central America (2008–2013): impacts, plausible causes and proposed solutions. *Food Security* 7(2): 303–321. <https://doi.org/10.1007/s12571-015-0446-9>
- Barima Y.S.S., Kouakou A.T.M., Bamba I., Sangne Y.C., Godron M., Andrieu J., Bogaert J. (2016) Cocoa crops are destroying the forest reserves of the classified forest of Haut-Sassandra (Ivory Coast). *Global Ecology and Conservation* 8: 85–98. <https://doi.org/10.1016/j.gecco.2016.08.009>
- Berthaud J. (1984) Les ressources génétiques pour l'amélioration des caféiers africains diploïdes : évaluation de la richesse génétique des populations sylvestres et de ses mécanismes organisateurs. Conséquences pour l'application. PhD thesis, Université de Paris-Sud, Orsay, France. Available at <http://www.documentation.ird.fr/hor/fdi:16623> [accessed 7 Nov. 2017].
- Boulvert Y. (2003) Carte morphopédologique de la République de Guinée à 1:500 000 en feuilles Ouest et Est. Paris, IRD.
- Brou Y.T., Akindès F., Bigot S. (2005a) La variabilité climatique en Côte d'Ivoire : entre perceptions sociales et réponses agricoles. *Cahiers Agricultures* 14(6): 533–540. Available at <http://revues.cirad.fr/index.php/cahiers-agricultures/article/view/3054816623> [accessed 7 Jun. 2018].
- Brou Y.T., Oszwald J., Bigot S., Servat E. (2005b) Risques de déforestation dans le domaine permanent de l'Etat en Côte d'Ivoire : quel avenir pour ces derniers massifs forestiers ? *Télédétection* 5(1–3): 236–275.
- Brou Y.T., Servat E., Paturel J.-E. (1998) Activités humaines et variabilité climatique : cas du sud forestier ivoirien. In: Servat E., Hughes D., Fritsch J.-M., Hulme M. (eds) Water resources

- variability in Africa during the 20th century. International Conference, Abidjan (CIV), 1998/11/16-19, Publication - AISH: 365–373. Wallingford, UK, AISH.
- Brou Y.T., Servat E., Paturel J.-E. (2000) Évolution du couvert forestier ivoirien sur la période 1950-1990, en relation avec la variabilité du climat et les activités anthropiques. In: Servat M., Servat-Vildary S. (eds) Dynamique à long terme des écosystèmes forestiers intertropicaux: 57–62. Paris, IRD, UNESCO.
- Bunn C., Läderach P., Ovalle Rivera O., Kirschke D. (2015) A bitter cup: climate change profile of global production of Arabica and Robusta coffee. *Climatic Change* 129(1–2): 89–101. <https://doi.org/10.1007/s10584-014-1306-x>
- Camara A., Dugué P., Cheylan J.-P., Kalms J.-M. (2009) De la forêt naturelle aux agroforêts en Guinée forestière. *Cahiers Agricoles* 18(5): 425–432. <https://doi.org/10.1684/agr.2009.0325>
- Chevalier A. (1905) Les caféiers sauvages de la Guinée française. Séance du 29 mai 1905. *Comptes rendus hebdomadaires des séances de l'Académie des Sciences de Paris* 140: 1472–1475.
- Chevalier A. (1909) Les Hauts plateaux du Fouta Djallon. *Annales de Géographie* 18(99): 253–261.
- Chevalier A. (1914) Sudania. Énumération des plantes récoltées en Afrique tropicale et équatoriale par M. Aug. Chevalier de 1905 à 1907. Tome II (N° 12001 à 20214). Paris, Auguste Challamel.
- Chevalier A. (1920) Exploration botanique de l'Afrique occidentale française. Tome I: Énumération des plantes récoltées. Paris, Lechevalier.
- Chevalier A. (1929) Les caféiers du Globe. Fascicule I. Généralités sur les caféiers. Paris, Lechevalier.
- Cordier L. (1961) Les objectifs de la sélection caféière en Côte d'Ivoire. *Café Cacao Thé* 5(3): 147–159.
- Correia M., Diabaté M., Beavogui P., Guilavogui K., Lamanda N., de Foresta H. (2010) Conserving forest tree diversity in Guinée Forestière (Guinea, West Africa): the role of coffee-based agroforests. *Biodiversity and Conservation* 19(6): 1725–1747. <https://doi.org/10.1007/s10531-010-9800-6>
- Couturon E., Montagnon C. (1991) Rapport de prospection de caféiers *C. canephora* spontanés dans la région de Touba en Côte d'Ivoire du 4/11/1991 au 7/11/1991. IRCC Côte d'Ivoire.
- Cubry P. (2008) Structuration de la diversité génétique et analyse des patrons de déséquilibre de liaison de l'espèce *Coffea canephora* Pierre ex. Froehner. PhD thesis, Université Montpellier 2, France.
- Cubry P., De Bellis F., Avia K., Bouchet S., Pot D., Dufour M., Legnate H., Leroy T. (2013a) An initial assessment of linkage disequilibrium (LD) in coffee trees: LD patterns in groups of *Coffea canephora* Pierre using microsatellite analysis. *BMC Genomics* 14(1): 10. <https://doi.org/10.1186/1471-2164-14-10>
- Cubry P., De Bellis F., Pot D., Musoli P., Leroy T. (2013b) Global analysis of *Coffea canephora* Pierre ex Froehner (Rubiaceae) from the Guineo-Congolese region reveals impacts from climatic refuges and migration effects. *Genetic Resources and Crop Evolution* 60(2): 483–501. <https://doi.org/10.1007/s10722-012-9851-5>
- Cubry P., Musoli P., Legnaté H., Pot D., De Bellis F., Poncet V., Anthony F., Dufour M., Leroy T. (2008) Diversity in coffee assessed with SSR markers: structure of the genus *Coffea* and perspectives for breeding. *Genome* 51(1): 50–63. <https://doi.org/10.1139/g07-096>
- Davis A.P. (2011) *Psilanthus mannii*, the type species of *Psilanthus*, transferred to *Coffea*. *Nordic Journal of Botany* 29(4): 471–472. <https://doi.org/10.1111/j.1756-1051.2011.01113.x>
- Davis A.P., Chadburn H., Moat J., O'Sullivan R., Hargreaves S., Lughadha E.N. (2019) High extinction risk for wild coffee species and implications for coffee sector sustainability. *Science Advances* 5(1): eaav3473. <https://doi.org/10.1126/sciadv.aav3473>
- Davis, A.P., Gole T.W., Baena S., Moat J. (2012) The impact of climate change on indigenous arabica coffee (*Coffea arabica*): Predicting future trends and identifying priorities. *PLoS ONE* 7(11): e47981. <https://doi.org/10.1371/journal.pone.0047981>
- Davis A.P., Govaerts R., Bridson D.M., Stoffelen P. (2006) An annotated taxonomic conspectus of the genus *Coffea* (Rubiaceae). *Botanical Journal of the Linnean Society* 152(4): 465–512. <https://doi.org/10.1111/j.1095-8339.2006.00584.x>
- Devreux M., Vallaey G., Pochet P., Gilles A. (1959) Recherches sur l'autostérilité du caféier robusta (*Coffea canephora* Pierre). Bruxelles, Publications de l'Institut National pour l'Étude Agronomique du Congo Belge (INEAC).
- Dussert S., Lashermes P., Anthony F., Montagnon C., Trouslot P., Combes M.-C., Berthaud J., Noirot M., Hamon S. (2003) Coffee (*Coffea canephora*). In: Hamon P., Seguin M., Perrier X., Glaszmann J.-C. (eds) Genetic diversity of cultivated tropical plants: 239–258. Enfield (NH), USA & Montpellier, France, Science Publishers, Inc. & CIRAD.
- Earl D.A., vonHoldt B.M. (2012) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources* 4(2): 359–361. <https://doi.org/10.1007/s12686-011-9548-7>
- El Mousadik A., Petit R.J. (1996) High level of genetic differentiation for allelic richness among populations of the argan tree [*Argania spinosa* (L.) Skeels] endemic to Morocco. *Theoretical and Applied Genetics* 92(7): 832–839. <https://doi.org/10.1007/BF00221895>
- Eskes A.B., Leroy T. (2012) Coffee selection and breeding. In: Wintgens J.N. (ed.) Coffee: growing, processing, sustainable production: a guidebook for growers, processors, traders, and researchers: 61–90. Weinheim, Germany, Wiley-VCH Verlag GmbH & Co. KGaA.
- Evanno G., Regnaut S., Goudet J. (2005) Detecting the number of clusters of individuals using the software structure: a simulation study. *Molecular Ecology* 14(8): 2611–2620. <https://doi.org/10.1111/j.1365-294X.2005.02553.x>
- Fairhead J., Leach M. (1996) Enriching the landscape: social history and the management of transition ecology in the forest-savanna mosaic of the Republic of Guinea. *Africa: Journal of the International African Institute* 66(1): 14–36. <https://doi.org/10.2307/1161509>
- Fairhead J., Leach M. (1998) Reframing deforestation: global analyses and local realities with studies in West Africa. London, Routledge.
- Ferrão R.G. (2007) Café Conilon. Vitória, Brazil, Incaper-Instituto Capixaba de Pesquisa, Assistência Técnica e Extensão Rural.
- Fréchou H. (1955) Les plantations européennes en Côte d'Ivoire. *Cahiers d'Outre-Mer* 29: 56–83.
- Froehner A. (1897) Übersicht über die Arten der Gattung *Coffea*. *C. canephora* Pierre. *Notizblatt des Königl. Botanischen Gartens und Museums zu Berlin* 7: 230–238. Available at <https://www.biodiversitylibrary.org/page/28795736#page/272/mode/1up> [accessed 27 Apr. 2019].
- Gomez C., Dussert S., Hamon P., Hamon S., de Kochko A., Poncet V. (2009) Current genetic differentiation of *Coffea canephora* Pierre ex A.Froehn in the Guineo-Congolian African zone: cumulative impact of ancient climatic changes and recent human

- activities. *BMC Evolutionary Biology* 9(1): 167. <https://doi.org/10.1186/1471-2148-9-167>
- Goudet J. (2005) HIERFSTAT, a package for r to compute and test hierarchical F-statistics. *Molecular Ecology Notes* 5(1): 184–186. <https://doi.org/10.1111/j.1471-8286.2004.00828.x>
- Guillaumet J.-L., Adjanohoun E. (1971) La végétation de la Côte d'Ivoire. In: Le milieu naturel de la Côte d'Ivoire, Mémoires ORSTOM: 161–263. Paris, ORSTOM. Available at <http://www.documentation.ird.fr/hor/fdi:16372> [accessed 13 Sep. 2018].
- Hart M.L., Forrest L.L., Nicholls J.A., Kidner C.A. (2016) Retrieval of hundreds of nuclear loci from herbarium specimens. *Taxon* 65(5): 1081–1092. <https://doi.org/10.12705/655.9>
- ICO (2018) Coffee market report. October 2018. International Coffee Organization. Available at <http://www.ico.org/documents/cy2018-19/cmr-1018-e.pdf> [accessed 28 Nov. 2018].
- IFCC (1960) Institut Français du Café et du Cacao. Annual report 1960. Paris.
- JOGF (1935) Actes du gouvernement local. Décision du Lieutenant-Gouverneur fixant les soldes annuelles des chefs de canton du cercle de Macenta. *Journal Officiel de la Guinée française. 1^{er} septembre 1935*. Available at <http://gallica.bnf.fr/ark:/12148/bpt6k9776901t> [accessed 13 June 2018].
- Kopelman N.M., Mayzel J., Jakobsson M., Rosenberg N.A., Mayrose I. (2015) Clumpak: a program for identifying clustering modes and packaging population structure inferences across K. *Molecular Ecology Resources* 15(5): 1179–1191. <https://doi.org/10.1111/1755-0998.12387>
- Laplante A.C., Rougerie G. (1949) Étude et carte pédologique provisoire de la région littorale située à l'Est de la Côte d'Ivoire. *Journal d'agriculture traditionnelle et de botanique appliquée* 29: 24–33. <https://doi.org/10.3406/jatba.1949.6201>
- Le Pierrès D., Charmetant P., Gnesio T., Bamba A., Konate M., Yapo A. (1987) Les ressources génétiques naturelles des cafés en Guinée : rapport sur la mission de prospection effectuée par une équipe ORSTOM/IRCC du 1^{er} février au 2 mars 1987. Adiopodoumé, Côte d'Ivoire, ORSTOM. Available at <http://www.documentation.ird.fr/hor/fdi:27766> [accessed 22 February 2017].
- Le Pierrès D., Charmetant P., Yapo A., Leroy T., Couturon E., Bontems S., Téhé H. (1989) Les cafés sauvages de Côte d'Ivoire et de Guinée : bilan des missions de prospection effectuées de 1984 à 1987. In: Treizième colloque scientifique international sur le café, Païpa, Colombie, 1989/08/21–25: 420–428. Paris, ASIC.
- Lemery A. (ed.) (1937) Le café en Côte d'Ivoire. La Société des Plantations d'Elima. *La Revue des ambassades et des questions diplomatiques et coloniales* Janvier 1937: 36. Available at <https://bibliotheque-numerique.diplomatie.gouv.fr/ark:/12148/bpt6k6473705h/f29.image.texteImage> [accessed 22 Jan. 2019].
- Leroy J.-F. (1974) Roland Portères : Professeur au Muséum National d'Histoire Naturelle (1906–1974). *Journal d'agriculture tropicale et de botanique appliquée* 21(10–12): 384–384. Available at https://www.persee.fr/doc/jatba_0021-7662_1974_num_21_10_6948 [accessed 10 Jan. 2018].
- Leroy T., De Bellis F., Legnate H., Kananura E., Gonzales G., Pereira L.F., Andrade A.C., Charmetant P., Montagnon C., Cubry P., Marraccini P., Pot D., de Kochko A. (2011) Improving the quality of African robustas: QTLs for yield and quality-related traits in *Coffea canephora*. *Tree Genetics & Genomes* 7(4): 781–798. <https://doi.org/10.1007/s11295-011-0374-6>
- Leroy T., De Bellis F., Legnate H., Musoli P., Kalonji A., Looor Solórzano R.G., Cubry P. (2014) Developing core collections to optimize the management and the exploitation of diversity of the coffee *Coffea canephora*. *Genetica* 142(3): 185–199. <https://doi.org/10.1007/s10709-014-9766-5>
- Leroy T., Marraccini P., Dufour M., Montagnon C., Lashermes P., Sabau X., Ferreira L.P., Jourdan I., Pot D., Andrade A.C., Glaszmann J.-C., Vieira L.G.E., Piffanelli P. (2005) Construction and characterization of a *Coffea canephora* BAC library to study the organization of sucrose biosynthesis genes. *Theoretical and Applied Genetics* 111(6): 1032–1041. <https://doi.org/10.1007/s00122-005-0018-z>
- Leroy T., Montagnon C., Charrier A., Eskes A.B. (1993) Reciprocal recurrent selection applied to *Coffea canephora* Pierre. I. Characterization and evaluation of breeding populations and value of intergroup hybrids. *Euphytica* 67(1–2): 113–125. <https://doi.org/10.1007/BF00022734>
- Leroy T., Montagnon C., Cilas C., Yapo A., Charmetant P., Eskes A.B. (1997) Reciprocal recurrent selection applied to *Coffea canephora* Pierre. III. Genetic gains and results of first cycle intergroup crosses. *Euphytica* 95(3): 347–354. <https://doi.org/10.1023/A:1003074716379>
- Lister D.L., Bower M.A., Howe C.J., Jones M.K. (2008) Extraction and amplification of nuclear DNA from herbarium specimens of emmer wheat: a method for assessing DNA preservation by maximum amplicon length recovery. *Taxon* 57(1): 254–258.
- Lister D.L., Bower M.A., Jones M.K. (2010) Herbarium specimens expand the geographical and temporal range of germplasm data in phylogeographic studies. *Taxon* 59(5): 1321–1323.
- Lubès H., Paturel J.-E., Servat E., Masson J.M., Kouamé B. (1995) Analysis of annual rainfall series in Ivory Coast in order to detect climatic variations. In: Sixth international meeting on statistical climatology: 455–458. Galway, Ireland, University College. Available at <http://www.documentation.ird.fr/hor/fdi:010060227> [accessed 12 Dec. 2018].
- Maclaud C. (1899) À travers la Guinée et le Fouta-Djallon. La mission Maclaud. *Bulletin du Comité de l'Afrique française* Septembre 1899: 293–298.
- Montagnon C., Leroy T., Yapo A. (1992a) Diversité génotypique et phénotypique des quelques groupes de cafés (*Coffea canephora* P.) en collection. Conséquence sur leur utilisation en sélection. *Café Cacao Thé* 36(3): 187–198.
- Montagnon C., Leroy T., Yapo A. (1992b) Étude complémentaire de la diversité génotypique et phénotypique des cafés de l'espèce *Coffea canephora* en collection en Côte d'Ivoire. In: Quatorzième colloque scientifique international sur le café, San Francisco, USA, 1991/07/14–19: 444–450. Paris, ASIC.
- Montagnon C., Leroy T., Yapo A. (1993) Caractérisation et évaluation des cafés *Coffea canephora* prospectés dans les plantations de Côte d'Ivoire. *Café Cacao Thé* 37(2): 115–119.
- Musoli P., Cubry P., Aluka P., Billot C., Dufour M., De Bellis F., Pot D., Bieysse D., Charrier A., Leroy T. (2009) Genetic differentiation of wild and cultivated populations: diversity of *Coffea canephora* Pierre in Uganda. *Genome* 52(7): 634–646. <https://doi.org/10.1139/g09-037>
- Perrier X., Jacquemoud-Collet J.-P. (2006) DARwin software Version 6. Available at <http://darwin.cirad.fr/darwin> [accessed 20 Nov. 2018].
- Poncet V., Dufour M., Hamon P., Hamon S., de Kochko A., Leroy T. (2007) Development of genomic microsatellite markers in *Coffea canephora* and their transferability to other coffee species. *Genome* 50(12): 1156–1161. <https://doi.org/10.1139/g07-073>
- Poncet V., Hamon P., Minier J., Carasco C., Hamon S., Noirot M. (2004) SSR cross-amplification and variation within cof-

- fee trees (*Coffea* spp.). *Genome* 47(6): 1071–1081. <https://doi.org/10.1139/g04-064>
- Portères R. (1934) Observations et sélection sur les caféiers à la station de Bingerville. I. Conduite de la sélection du caféier à la station de Bingerville. Bingerville, Côte d'Ivoire. Service de l'agriculture. Gouvernement de Côte d'Ivoire.
- Portères R. (1937a) Étude sur les caféiers spontanés de la section des *Eucoffea*. Partie 1 : Répartition et habitat. *Annales Agricoles de l'Afrique Occidentale Française et étrangère* 1: 68–91.
- Portères R. (1937b) Étude sur les caféiers spontanés de la section des *Eucoffea*. Partie 2 : Espèces, variétés et formes. *Annales Agricoles de l'Afrique Occidentale Française et étrangère* 1: 219–263.
- Portères R. (1937c) Étude sur les caféiers spontanés de la section des *Eucoffea*. Partie 3 : Mise en culture et sélection. *Annales Agricoles de l'Afrique Occidentale Française et étrangère* 1: 407–439.
- Portères R. (1939) Le choix des variétés et l'amélioration des caféiers en Côte d'Ivoire. *Revue de botanique appliquée et d'agriculture coloniale* 19: 18–29.
- Portères R. (1959) Valeur agronomique des caféiers des types Kouilou et Robusta cultivés en Côte d'Ivoire. *Café Cacao Thé* 3(1): 3–13.
- Portères R. (1962) Caféiers de la République de Guinée. *Café Cacao Thé* 6(1): 3–18.
- Poupart Y. (1938) La culture du *Coffea Robusta* dans l'Ouest de la Côte d'Ivoire. *Agronomie Coloniale* 27(252): 173–180.
- Pritchard J.K., Stephens M., Donnelly P. (2000) Inference of population structure using multilocus genotype data. *Genetics* 155(2): 945–959.
- Risterucci A.M., Grivet L., N'Goran J.A.K., Pieretti I., Flament M.H., Lanaud C. (2000) A high-density linkage map of *Theobroma cacao* L. *Theoretical and Applied Genetics* 101(5–6): 948–955. <https://doi.org/10.1007/s001220051566>
- Roullier C., Benoit L., McKey D.B., Lebot V. (2013) Historical collections reveal patterns of diffusion of sweet potato in Oceania obscured by modern plant movements and recombination. *Proceedings of the National Academy of Sciences of the United States of America* 110(6): 2205–2210. <https://doi.org/10.1073/pnas.1211049110>
- Saitou N., Nei M. (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution* 4(4): 406–425. <https://doi.org/10.1093/oxfordjournals.molbev.a040454>
- Särkinen T., Staats M., Richardson J.E., Cowan R.S., Bakker F.T. (2012) How to open the treasure chest? Optimising DNA extraction from herbarium specimens. *PLoS ONE* 7(8): e43808. <https://doi.org/10.1371/journal.pone.0043808>
- Staats M., Erkens R.H.J., van de Vossenberg B., Wieringa J.J., Kraaijeveld K., Stielow B., Geml J., Richardson J., Bakker F.T. (2013) Genomic treasure troves: complete genome sequencing of herbarium and insect museum specimens. *PLoS ONE* 8(7): e69189. <https://doi.org/10.1371/journal.pone.0069189>
- Suchan T., Pitteloud C., Gerasimova N.S., Kostikova A., Schmid S., Arrigo N., Pajkovic M., Ronikier M., Alvarez N. (2016) Hybridization capture using RAD probes (hyRAD), a new tool for performing genomic analyses on collection specimens. *PLoS ONE* 11(3): e0151651. <https://doi.org/10.1371/journal.pone.0151651>
- Thiers B. (2019) Index Herbariorum: A worldwide index of 3,400 herbaria and associated staff. New York Botanical Garden Steere Herbarium (continuously updated). Available at <http://sweetgum.nybg.org/science/ih/> [accessed 27 Apr. 2019].
- Tourte R. (2005a) Histoire de la recherche agricole en Afrique tropicale francophone. Volume V. Le temps des stations et de la mise en valeur 1918–1940/1945. Rome, FAO.
- Tourte R. (2005b) Histoire de la recherche agricole en Afrique tropicale francophone. Volume VI. De l'empire colonial à l'Afrique indépendante. 1945–1960. La recherche prépare le développement. Rome, FAO.
- Wallengren M. (2017) The arabica-robusta balance. Does it really matter? A special report by SpillingTheBeans for the Global Coffee Review's. Jan. 2013 issue. Available at <http://spilling-the-beans.net/special-report-the-arabica-robusta-balance-does-it-really-matter/> [accessed 2 Dec. 2018].
- White F. (1979) The Guineo-Congolian region and its relationships to other phytochoria. *Bulletin du Jardin botanique national de Belgique / Bulletin van de Nationale Plantentuin van België* 49(1): 11–55. <https://doi.org/10.2307/3667815>
- Yao A., Kpangui K., Kouao K., Adou L., Vroh B., N'Guessan K. (2013) Diversité floristique et valeur de la forêt sacrée Bokasso (Est de la Côte d'Ivoire) pour la conservation. *VertigO - la revue électronique en sciences de l'environnement* 13(1). <https://doi.org/10.4000/vertigo.13500>

Communicating Editor: Myriam Heuertz.

Submission date: 11 Feb. 2019

Acceptance date: 7 Oct. 2019

Publication date: 26 Mar. 2020