



Is the genetic integrity of wild *Coffea canephora* from Ivory Coast threatened by hybridization with introduced coffee trees from Central Africa?

Diane Estelle Gnapi · Désiré N'Da Pokou · Hyacinthe Legnate ·
Zhang Dapeng · Christophe Montagnon · Benoît Bertrand ·
Assonvo Simon-Pierre N'guetta

Received: 1 November 2021 / Accepted: 10 March 2022
© The Author(s) 2022

Abstract Robusta coffee (*Coffea Canephora*) is an important source of income and employment, contributing significantly to the economies of many developing countries. This species is split into two genetic groups: Guinean and Congolese. Côte-d'Ivoire is the primary diversification center of the Guinean pool, which offers great potential for genetic improvement. However, genetic erosion of this group is currently a threat in this country because of the massive introduction of Congolese accessions done in the previous decades. In this work, we analyzed the progenies of isolated population of *C. canephora* composed of a mix of Guinean and Congolese using SNP markers. The analysis is focus on the male parents'

contribution in open pollination. The results show that the introduction of cultivars from the Congolese gene pool threat the genetic integrity of wild populations from the Guinean gene pool. Indeed, this analysis showed that Congolese pollens have more ability to pollinate the Guinean group than the Guinean pollen perform on Congolese trees. We recommend that a program be put in place to ensure the conservation of wild population of *C. canephora* through in situ or ex situ conservation.

Resume Le café Robusta (*Coffea Canephora*) est une source importante de revenus et d'emplois, contribuant de manière significative aux économies de

D. E. Gnapi · A. S.-P. N'guetta
Unité de Formation et de Recherche Biosciences,
Laboratoire de Génétique, Université Félix Houphouët
Boigny, 22 BP 582, Abidjan 22, Ivory Coast

D. E. Gnapi · D. N. Pokou (✉)
Central Biotechnology Laboratory, Centre National
de Recherche Agronomique, 01 BP 1740, Abidjan 01,
Ivory Coast
e-mail: desire.pokou@gmail.com

D. E. Gnapi · H. Legnate
Station de Recherche Café Cola, Centre National de
Recherche Agronomique, Man, Ivory Coast BP440

Z. Dapeng
United State Department of Agriculture-Agricultural
Research Service (USDA-ARS), 10300 Baltimore Avenue
Bldg 001 Barc-west, Beltsville, MD 20705, USA

C. Montagnon
RD2 Vision, 60 rue du Carignan, 34270 Valflaunes, France

B. Bertrand
Centre International pour la Recherche Agronomique
et le Développement (CIRAD), Unité Mixte de
Recherche DIADE, CIRAD-IRD 911 av Agropolis,
34398 Montpellier, France

B. Bertrand
Université de Montpellier UMR DIADE, CIRAD, IRD,
34398 Montpellier, France

nombreux pays en voie de développement. La Côte d'Ivoire est le centre de diversification primaire du pool Guinéen qui offre un grand potentiel d'amélioration génétique. L'érosion génétique de ce groupe est actuellement une menace dans ce pays. Cependant, ce pool génétique court un risque plus élevé d'érosion génétique en raison de l'extinction locale. Dans ce travail, Nous montrons comment l'introduction des cultivars du pool génétique Congolais, menace l'intégrité génétique de populations sauvages du pool génétique Guinéen. En effet, les analyses ont montré que les pollens des arbres Congolais sont plus aptes à féconder les arbres Guinéens que les Guinéens le sont pour les Congolais. Ce qui nous fait suggérer la mise en place d'un programme de toute urgence pour la conservation in situ et ex situ de ce pool génétique.

Keywords *Coffea canephora* · Genetic resources · Genetic erosion · Crop wild relative

Mots clés Ressources génétique · *Coffea canephora* · érosion génétique · Caféiers sauvages

Introduction

Coffee (*Coffea* sp) is a perennial plant originated from the tropical and intertropical areas of Africa, Comoros, Mascarene Islands (Mauritius and Reunion Island), Asia and Australia (Davis et al. 2006). Coffee trees are split into 130 species (Davis and Rakotonasolo 2021), all belonging to the genus *Coffea* and the family Rubiaceae (Davis et al. 2011). In all producing countries, the economic and social interest of coffee is no longer in question. Indeed, the income generated by its cultivation are estimated at over 200 billion dollars (US) for the entire coffee value chain (ICO 2019). This represent a major export product, with more than 13 million tons of green coffee produced or 169.6 million bags of 60 kg green coffee for the year 2020/2021 (ICO 2021), and its cultivation covers more than 10 million hectares of cultivable land in 80 countries (FAOSTAT 2019). In addition, coffee production and processing employs over 125 million people in the production chain (Van Der Vossen 2016). In Côte d'Ivoire, the coffee cultivation area represents over 400,000 hectares (FAOSTAT 2018).

Coffee therefore presents an important economic and social challenge for producing countries.

Among the 130 described species, *Coffea canephora* is the main species cultivated in Côte d'Ivoire because of its adaptation to the local environment (Coste 1989). This species is diploid ($2n=2x=22$), strictly allogamous and self-incompatible (Berthaud 1986; Devreux et al. 1959; Coulibaly et al. 2002). Phylogenetic studies carried out within the species *C. canephora*, have allowed a structuring into two distinct genetic groups with contrasting morphological traits for architecture, resistance to orange rust, bean size and caffeine content (Berthaud 1986; Leroy et al. 1993, 1997; Montagnon et al. 2008). These are the early flowering Guinean pool, originating from West Africa, especially the mountainous western area of Ivory Coast and East Guinea as are considered to be the primary center of diversification; while the late flowering Congolese pool is originated to Central Africa (essentially the Congo Basin). The Congolese pool is subdivided into two subgroups named SG1 and SG2 (Montagnon et al. 1992a, b).

In order to improve the productivity, numerous accessions from SG1 and SG2 were introduced in Côte-d'Ivoire at the beginning of the cultivation of *C. canephora* (Berthaud 1986; Montagnon et al. 1992b). Thus, due to the contrasting characters of the Guinean and Congolese groups, they were used in a reciprocal recurrent selection breeding scheme (Leroy et al. 1993, 1997; Montagnon 2000). The genetic resources involved in this selection scheme were conserved as ex situ genebanks. Most of these resources are located at the Centre National de Recherche Agronomique (CNRA) station in Divo (Côte d'Ivoire). This gene bank held about 1900 accessions of *C. canephora*, from various origins (Labouisse et al. 2020).

The geographic range and reproductive system of *C. canephora*, as well as the impact of past climates, have resulted in high genetic variability within the species and the formation of unselected populations obtained from seeds from open pollinations (Gomez et al. 2009; Eskes and Leroy 2012). In addition, other factors such as the lack of pollens and female flowers, flowering synchronicity, number and density of trees, and their spatial distributions, act to alter the genetic diversity and structure of natural populations (oddou-Muratorio et al. 2011; Sagnard et al. 2011). In addition, population size and the levels of gene flow in *C. canephora*, may have long-standing effects on

the genetic composition of current gene pools in the wild. These threats are to be considered with the impact of climate change (Bunn et al. 2015), bushfires, and heavy deforestation of natural forests in Côte d'Ivoire and Guinea (Brou et al. 2000). Besides, the introduction of the Congolese pool into Côte d'Ivoire has helped foster genetic exchanges with the Guinean group in local plantations (Montagnon et al. 1993). Is the shift between early and late flowering a sufficient barrier to prevent mass pollination of the wild populations by pollens from the cultivated trees?

The objective of this study is to evaluate the impact of the introduction of the Congolese pool on the genetic integrity of the Guinean pool endemic to Côte d'Ivoire, and mostly distributed in natural areas where coffee growing is important.

Material and methods

Site of study

The study was conducted at the Centre National de Recherche Agronomique (CNRA) research station of Divo, located between 5°46' north latitude and 5°17' west longitude. The relief is characterized by plains with an altitude of between 100 and 200 m. The soils are ferralitic, deep, acidic and desaturated in exchangeable bases (Kassin 2009). The Average annual temperatures varies between 24 and 28 °C, with an average rainfall of 1218 mm/year (Ehounou 2021). Rainfall is the most important limiting climatic factor after the outside temperature. There are two notions to be considered: the height of monthly and annual rainfall and their distribution, mainly weekly. This distribution will explain the duration and the severity of the dry season, a period of vegetative rest and floral induction: limited to three-four months for *Canephora*. Coffee trees generally thrive in regions reaching 1500 to 1800 mm/year, with some less rainy months favorable to fruit ripening. Below 800 to 1000 mm of rainfall per year, coffee growing can become problematic. *Canephora* does well with 1,200 to 2,000 mm of rainfall, well distributed with less than five consecutive months of dry period. The flowering of the coffee trees is triggered by a rainfall of at least 10 mm and when this quantity is not reached, the coffee trees turn pink, resulting in a star-like flowering. This starry bloom does not allow the

release of pollen from the anther which will lead to a bad bloom and even a bad coffee production at the end of the year.

Plant material

The plant material is composed of 94 genotypes among which, 34 genotypes from the Guinean group, 44 genotypes from the Congolese group and 16 hybrid genotypes between the Guinean and Congolese groups. Each genotype has been pollinated by a mix of pollens from others genotypes. The progenies were planted in an experimental plot named Ci1/11/93 which is split into six blocks. Each block is composed by 51 lines of 15 trees. In each line, the genotype is planted at location 8 and the others location in the line are composed by its progenies. The plot Ci1/11/93 is established in 1993 and isolated by an area of primary forest to prevent any involvement of genotypes from other plots in the pollination. All the 94 genotypes and their planted progenies were considered as mother trees and were subject of open pollination. Offsprings from block 1 were used in the present study. As regard the study of flowering phenology, only block 1 was considered. The plant material in this block is composed of 62 genotypes of 34 from the Congolese group or 55%, 25 from the Guinean group or 44% and 3 Guinean-Congolese inter-group hybrids or 1% of the population.

Plant material used for the phenology study

Flowering phenology parameters were monitored on 139 trees out of the 276 trees in block 1. The 139 are composed of 79 trees from 12 Guinean genotypes and 60 trees from 11 Congolese genotypes.

Plant material used for the paternity study

All 94 parental genotypes of 34 genotypes from the Guinean group, 44 genotypes from the Congolese group, and 16 hybrids genotypes from the cross between the Guinean and Congolese were considered for paternity analysis. The progenies are composed of 1361 seedlings obtained from open pollination of 79 mother trees in Block I. The paternity of the different genetic groups (Guinean, Congolese and hybrids) was determined and the contribution of the paternal

parents in the progeny under open pollination conditions was evaluated (Table 1).

Methods

Phenology study

In coffee plants under the climatic conditions of Côte d'Ivoire, flowering is generally triggered by a rainfall of at least 10 mm (Louarn 1992) during the dry season, which extends from December to March in Côte d'Ivoire. The study of flowering phenology consisted of noting the time between the rainfall trigger and anthesis on the one hand, and evaluating the flowering intensity of the plants on the other, using the following rating scale:

- 0: no flowering;
- 1: very weak flowering (at least 20% of the crown covered with flowers, about ten);
- 2: weak (between 20 and 50% of the crown covered with flowers, about twenty);
- 3: medium (at least 50% of the crown covered with flowers). From this score, flowering is considered satisfactory (Maalouf et al. 2011);
- 4: abundant (between 50 and 80% of the crown covered with flowers);
- 5: very abundant (more than 80% of the crown covered with flowers).

Paternity study

Discovering SNP markers through data mining

SNP data mining was performed using sequence data of 14 *C. canephora* genotypes which were deposited in the NCBI Sequence Read Archive (SRA) database. These SRA reads were downloaded from the database

and mapped on the *C. canephora* reference genome (Denoeud et al. 2014) using BWA program. The Genome Analysis Toolkit (GATK) package v 3.520 was used for SNP calling using HaplotypeCaller with default parameter.

Then the hard filters (parameters: $QD < 2.0 \parallel FS > 60.0 \parallel MQ < 40.0 \parallel MQRankSum < -12.5 \parallel ReadPosRankSum < -8.0$) were applied to the filter low-quality alleles. To select high-quality SNPs for experimental validation, any SNPs that had other possible adjacent SNP sites 80 bp upstream or 80 bp downstream were eliminated. From the discovered putative SNPs, a subset of 288 putative SNPs was selected for validation test using the nanofluidic array genotyping system (Fluidigm Co, South San Francisco, CA). The primers of the selected 288 SNPs were designed by Fluidigm and applied on the selected jujube cultivars for validation. Based on the validation result, the top 200 SNPs with high repeatability were used for further analysis of genetic diversity. then the 90 most polymorphic SNPs were used for paternity analysis.

DNA isolation and SNP genotyping

SNP genotyping DNA was isolated from 3 discs of dry leave collected from each tree in the field and each progeny seeding. SNP genotyping was performed at USDA-ARS, Sustainable Perennial Crops Lab, Beltsville, MD, USA, using the Fluidigm 96.96 Dynamic Array™ (Fluidigm, San Francisco, CA). Each 96.96 Dynamic Array can run 96 samples against 96 SNP assays generating a total of 9.216 data points in a single experiment. One key feature of this protocol is the inclusion of a specific targeted amplification (STA) reaction (Wang et al. 2009), which allows the enrichment of template molecules for each individual Integrated Fluidic Circuit® (IFC) reaction that facilitates the multiplexing during genotyping. An advantage to STA is that it allows the use of limited or low-quality

Table 1 Genotypes used for the study of phenology and paternity

Study	Guinean genotypes	Congolese genotypes	Hybrides genotypes	Total relative
Phenology	12 [79]	11 [60]		23 [139]
Paternity	34 (501)	44 (656)	1 (202)	79 (1361)

[]: Number of plants per genetic pool

(): Number of offspring per genetic pool

DNA samples and reduces bias that may occur when samples are loaded to the 96 sample wells of the IFC. Since coffee leaf tissues contain high levels of polysaccharides and polyphenolic compounds that can potentially inhibit PCR amplification, the STA step was strongly recommended. The STA reaction was performed as described in the Fluidigm SNP Genotyping User Guide, PN 68,000,098 Rev I1 (Fluidigm 2013). The STA master mix consisted of 2.5 µL of TaqMan® Taq polymerase (Life Technologies, Carlsbad, CA), PreAmp Master Mix (2X), 1.25 µL of Pooled assay mix (0.2X), and 1.25 µL of genomic DNA for a total reaction volume of 5.0 µL.

PCR was performed with an initial denaturation step of 95 °C for 10 min, followed by 14 cycles of a 2-step amplification profile consisting of 15 s at 95 °C and 4 min at 60 °C. The resulting amplified DNA was then diluted 1:5 in TE buffer in order to reduce the concentration of any remaining PCR by-products. Samples were then genotyped using the nanofluidic 96.96 Dynamic Array™ IFC (Integrated Fluidic Circuit; Fluidigm Corp.). The 96.96 Dynamic Array IFC for SNP genotyping was described by Wang et al. (2009). End-point fluorescent images of the 96.96 IFC were acquired on an EP1™ imager (Fluidigm Corp.). The data was recorded with Fluidigm Genotyping Analysis Software (Fluidigm, San Francisco, CA).

Data analysis

Paternity analysis

CERVUS 3.0.6 software (Marshall et al. 1998; Kalinowski et al. 2007) was used to assign relatedness with SNPs genotype data. The method used is based on the maximum likelihood approach (Thompson 1976; Meagher 1986). For each descendant, a measure of the probability that each potential father is the actual father is calculated as a logarithmic score LOD (logarithm of odds). In CERVUS, the male parents of each offspring are determined using the "sex unknown" option. The list of candidate parents is composed of all genotypes in the trial. Paired parent analysis was applied to assign the two parents of each individual. The female parent serves as a control for the applied test. In CERVUS, the proportion of input errors is used as the error rate in the likelihood calculation during the simulation that precedes the parent

pair assignment operation. An error rate of 0.2% was used for this analysis. The simulation was done with 10,000 cycles of replicates with discount assuming that 80% of the potential parents were listed. The critical likelihood value (LOD score) was considered with a percentage of 95% in the strict sense and 80% in the broad sense during the simulation, respectively.

The contribution of a male progenitor or group of male progenitors to the pollination of a tree or group of trees in the plot is determined by the percentage of the number of paternal progenies observed for the tree or group of trees. The Mann–Whitney U test was used to compare the contribution values of Guinean and Congolese progeny. This test was performed using the XLSAT 2014 software.

Phenology analysis

Flowering data were analyzed with R v 3.6.2 software. The normality of the data and the homogeneity of the Fisher variances were verified by the Shapiro–Wilk test (Shapiro and Wilk 1965). Thus, non-parametric one-criteria classification tests were performed to compare the genetic groups and genotypes of canephora studied. For the twenty-three genotypes, the Kruskal–Wallis test was used and the Wilcoxon–Mann–Whitney test for the 2 genetic groups. These tests allowed multiple comparison of means at the 5% threshold when significant differences were observed.

Pollen dispersal in the plot

The distances between pollinated trees and pollinator trees are calculated by the Pythagore theorem. The relationship between the pollination distance and the importance of the contribution to pollination is estimated by linear regression.

Results

Spreading of the flowering

Flowering onset was synchronous between the Guinean and Congolese genetic groups over the four years of observation. However, analysis of the flowering spread curves shows a shift between the anthesis peaks of each group. Thus:

- In 2018, the peak of anthesis was observed on the 6th day after which 54.43% of the trees flowered in the Guineans. The peak for the Congolese was observed on day 7, with 81.96% of the trees flowering. From these two peaks, a continuous inflection of flowering in both genetic groups was observed until day 8 (Fig. 1a).
- In 2019, 97.68% of the trees in the Guinean group flowered on day 6, showing their peak anthesis. The peak anthesis of the Congolese trees was observed one day later with 51.66% of the trees flowering. From these two peaks, a progressive decrease in flowering of both genetic groups was observed until day 8 (Fig. 1b).
- Observations of the February 2020 flowering indicate that the peak of anthesis of the Guinean was reached on day 6 with 51.41%. One day later, the peak of the Congolese group was reached with 59.32% of trees in bloom. A continuous inflection of flowering in both groups was observed on day 8 (Fig. 1c).
- In January 2021, the peak of anthesis of Guinean trees was observed on day 6 with 46.62% of trees in flower. The Congolese peak was observed 2 days later with 91.66% of trees in bloom. From these two peaks, a gradual decrease in flowering of the two genetic groups was observed until day 9 (Fig. 1d).

Average flowering time in Guinean and Congolese coffee trees

The results of the average flowering time recorded over these four years showed that there was a significant difference between the flowering of the Congolese and Guinean coffee groups after the onset of the first rainfall following the long dry period (Table 2). Thus, the Guinean genotypes flowered earlier on the 6th day after the first rainfall following the long dry season, while the Congolese cultivars flowered one or two days later on the 7th or 8th day in 2018, 2019 and

Fig. 1 Flowering spread of all trees studied in Block I in February 2018, February 2019, February 2020 and February 2021

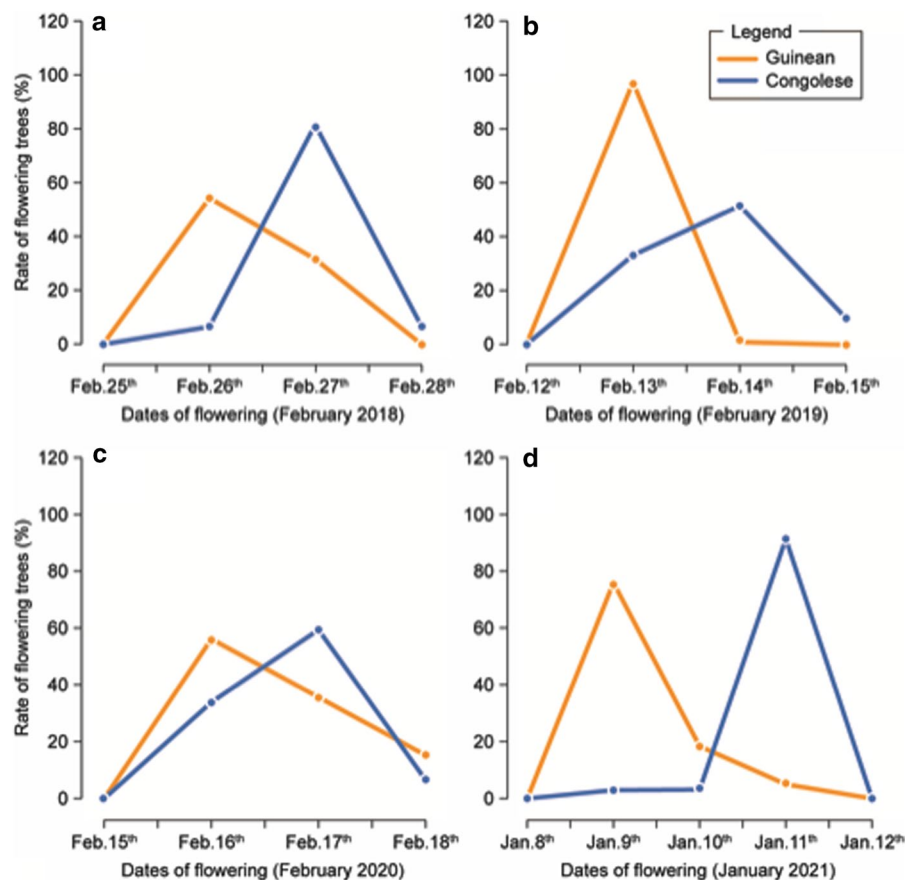


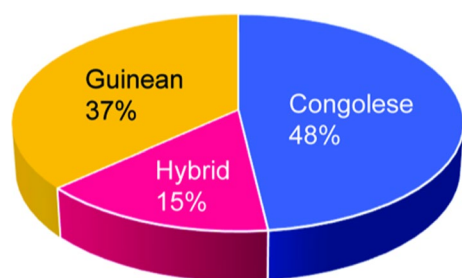
Table 2 Average values of flowering time in Guinean and Congolese genotypes during 2018, 2019, 2020 and 2021

Genetic group	Average day of flowering			
	2018	2019	2020	2021
Guinean	6.57 b	6.01 b	6.71 a	6.44 b
Congolese	7.04 a	6.83 a	6.77 a	7.8 a
Average	6.81	6.42	6.74	7
P	<0.0001	<0.0001	0.737	<0.0001

Values with different letters are statistically different: “a” is high and “b” is low value

(*i*): amount in mm of rain triggering flowering

2018: 13 mm; 2019: 14.07 mm; 2020: 99 mm; 2021: 34.4 mm

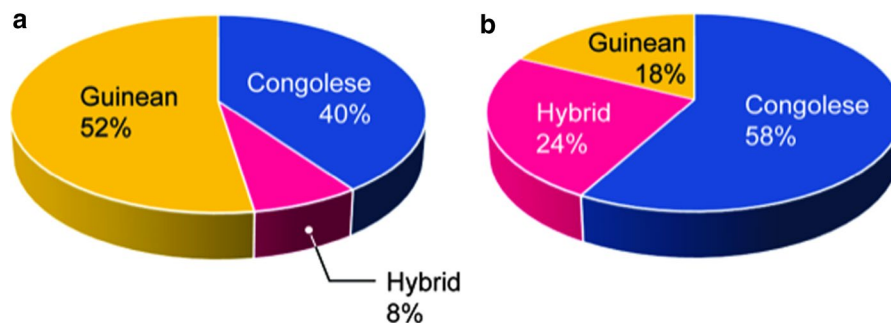
**Fig. 2** Paternal contribution of genetic groups in the offspring

2021. However, observations in 2020 showed synchronized flowering of the two genetic groups on day 7.

Paternity of parents

Paternal contributions in the descendants

After open pollination, progeny analysis revealed that Congolese sires were the major contributors to progeny with 48%. Guinean progenitors contributed 37% of the progeny and hybrids contributed 15% (Fig. 2).

Fig. 3 a Contribution of Congolese, Guinean and hybrid progenitors to the progeny of Guinean mother plants; **b** Contribution of Congolese, Guinean and hybrid progenitors to the progeny of Congolese parent plants

In the free-living progeny, there were 501 plants of Guinean progenitors, 656 plants of Congolese progenitors and 202 plants of inter-group hybrid progenitors. Thus, there was a very significant increase ($p < 0.001$) in the proportion of intergroup hybrids in the progeny, from 1 to 15%.

Contribution of Guinean, Congolese and hybrid male genotypes in the progeny of Guinean and Congolese feet.

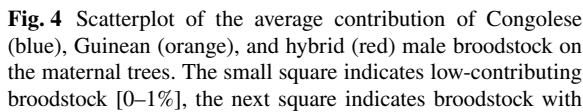
The sum of the contributions of the Guinean spawners to the Guinean broodstock set was 52% compared to 40% for the sum of the Congolese spawners (Fig. 3a). The Mann–Whitney test ($p = 0.053$) showed that there was no significant difference between the two proportions.

Among all Congolese mother plants, the sum of the contributions of Congolese broodstock was 58% compared to 18% for the sum of the contributions of Guinean mother plants (Fig. 3b). The Mann–Whitney test ($p = 0.002$) applied to the contributions of the Guinean and Congolese broodstock showed a significant difference.

Hybrids contributed significantly to the pollination of Congolese mother trees (24%) and less so to Guinean mother trees (8%). Thus, the hybrids appear to be more compatible with the Congolese than with the Guineans.

Average individual contribution of Congolese, Guinean and hybrid progenitors to the progeny of Congolese and Guinean mother plants.

The average individual contribution of the progenitors to the progeny of the genetic groups is shown in Fig. 4, ranging from 0 to 5.1% in the Guinean mother trees and from 0 to 4.5% in the Congolese mother trees. A first group of trees with a low contribution between 0 and 1% can be seen. This group represents 43.52% of the trees that made a low contribution to the pollination of female trees. An identical proportion of 43.52% shows a



Pollen dispersal in the plot

a contribution of 1–3%, and the largest square indicates broodstock that contributed the most to the pollination of female trees [3–5%]

Discussion

Gene flow between genetic groups of the same species plays a major role in the spatial organization of genetic diversity and is a major evolutionary force (Futuyma 1998). In plants, gene flow results from two major mechanisms, pollination and seed dispersal.

This study contributes to the analysis of gene flow between two genetically distant populations of the same species (*C. canephora*) (Guinean and Congolese) that co-exist in the same isolated plot. In the case of the two genetic groups studied here, this is particularly relevant because these two groups are known to flower at different dates (Louarn 1992; Akaffou 2013). Our study shows differences between the flowering peaks of the Guinean and Congolese genetic groups in the years 2018, 2019, 2020 and 2021. The anthesis period of a species or hybrid is the day when more than 50% of individuals flower (Maalouf et al. 2011). We observed a flowering lag between the Guinean and Congolese genetic groups with an

Table 3 Minimum and maximum distance of pollen dispersal

Offspring	Minimum distance of pollen dispersal (m)	Maximum distance of pollen dispersal (m)
8821 L2 A3	2	94.13
8183 L26 A15	2	42.57
8293 L7 A7	2	72
8245 L9 A9	4	76
8069 L49 A3	3.16	90.75
8583 L9 A3	2.23	76
8668 L31 A4	8	60.03
8518 L47 A13	4	88
9477 L44 A5	3	80
8189 L42 A13	6	80
8159 L30 A12	4.12	58.13
8410 L29 A6	2	54
8056 L37 A7	11.18	62.9
8464 L19 A3	2	60.2
8411 L3 A6	2	88
8121 L5 A14	4	90.19
8105 L18 A15	2	62.39
8214 L21 A7	2	56
8387 L25 A10	4.12	54.62
8477 L27 A3	2	52.95
8416 L43 A2	10	82.93
8159 L45 A14	8.48	78.10
8392 L50 A7	1.41	90.13

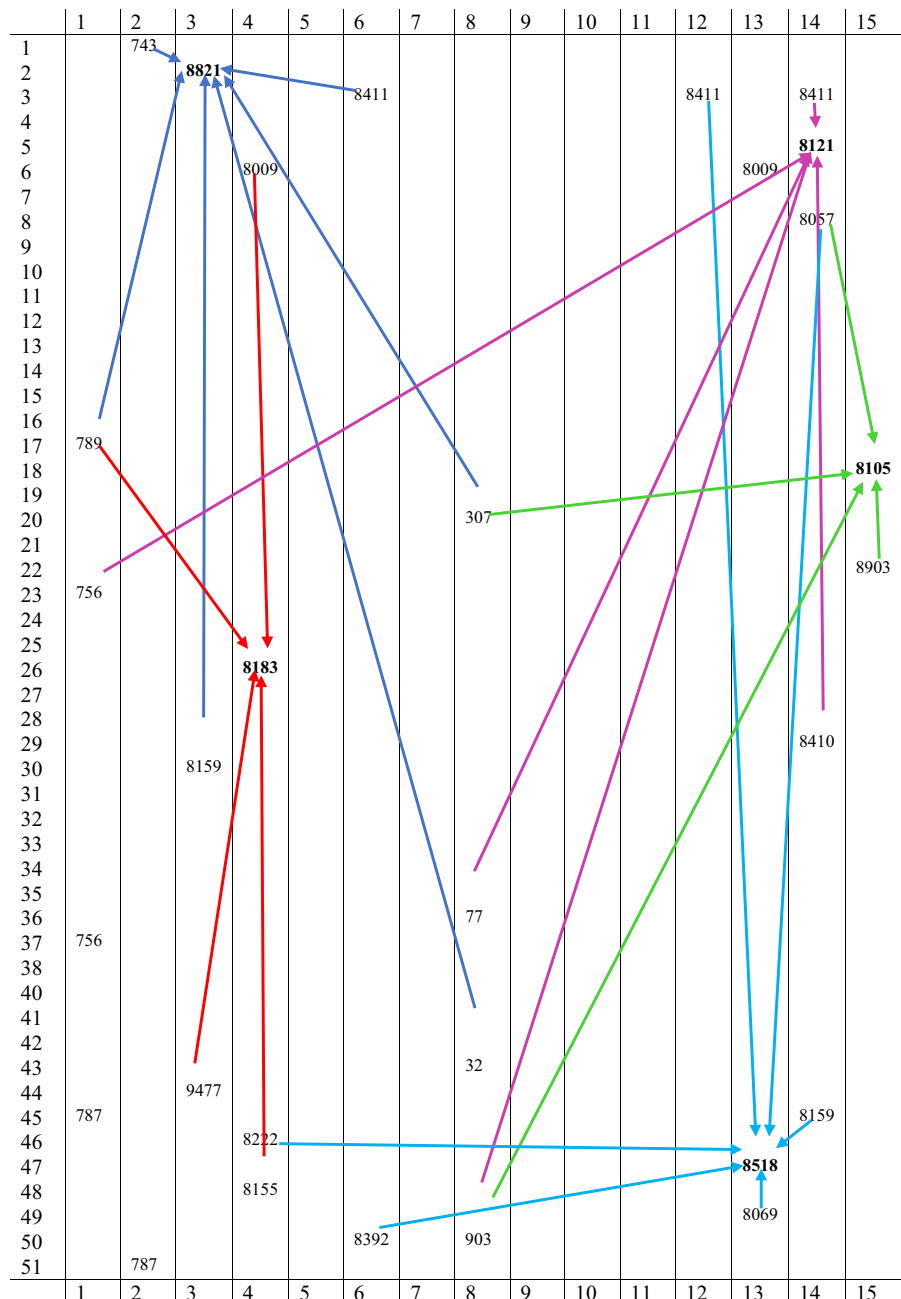
earliness of the Guineans and a tardiness of the Congolese of about 1 day. Our results are consistent with those of Berthaud (1986) and Leroy et al. (1993). However, this time lag seems to vary according to the year. In 2020, this time lag does not exist.

The genetic mixing study was conducted on 1361 offspring during 2018, or the average lag between the two groups was greater than ½ day. Despite this flowering lag, genetic mixing was observed between the two genetic groups. Our results revealed an increase in the number of hybrids (from 1% in the parental population to 15% in the progeny). This increase is accompanied by a 7% reduction in the frequency of Congolese (55 to 48%) and Guineans (44 to 37%). The significant increase in the frequency of hybrids is evidence of inter-group mixing. This inter-group mixing seems to depend mainly on a small number of progenitors (about 12.94%) that contributed to more than 28% of the pollinations. We have verified

that these "super-spawners" interbreed more with Guineans (19%) than with Congolese (9%) and this even when they are Congolese. This imbalance in the contribution of spawners does not seem to be related to their position in the plot. It does not seem to be clearly related to the number of plants either. For example, the Guinean "super-spawner" 789 is represented only once in the plot.

Assuming that the integrity of the concomitant genetic groups would be maintained, the pollen involved in the fertilization of both groups would have to come exclusively (100%) from the same group. We show in this study that this is not the case. The anthesis mismatch is not a sufficient barrier for cross-pollination. Paternity analyses showed that the contribution of Guinean and Congolese progenitors in the progeny of Guinean mother trees are equivalent. However, in the progeny of the Congolese mother trees, the contribution of the Congolese gene pool is much higher than the Guinean pool. These differences can be explained by several factors including the founder effect, genetic erosion in the Guinean gene pool (endemic to Ivory Coast) and the type of data generated by the different markers used in this study. A recent study carried out on wild populations of Guinean in their area of diversification, many hybrids between Guinean and Congolese were found (Labouisse et al. 2020). Our results therefore confirm the recent findings in wild population in Guinea. Besides, preferential fathers were also found for the mother coffee plants of the different genetic groups. This would be the consequence of gametophyte incompatibilities between certain genotypes (Berthaud 1986; Coulibaly et al. 2002), discrepancies between flowering periods (Berthaud 1986; Leroy et al. 1993) and the viability of the pollen produced by each progenitor. Furthermore, the introduction of genotypes from the Congolese pool into coffee populations in Ivory Coast most likely represents an evolutionary force. This is supported by the increase of (1–15%) hybrids in the progeny population which may be facilitated by gene flow through pollen (Berthaud 1984). Gene flow from these domesticated crops to wild Guinean species represents a threat to the genetic integrity of the Guinean pool. Indeed, repeated occurrences of hybridization can lead to the loss of genetic integrity of the wild species, which becomes assimilated to the cultivar (Aerts et al. 2013). Therefore, to conserve the integrity and genetic diversity of the Guinean

Fig. 5 Pollen dispersal in the plot



pool, in situ and ex situ conservation strategies are essential.

Conclusion and recommendations

The success of Ivorian coffee growing is largely due to the introduction of the Congolese pool. This pool

naturally hybridized with endemic Guinean coffee trees in Ivory Coast to form fertile and sometimes very productive intergroup hybrids that were empirically selected by breeders. However, in addition to deforestation, bush fires and climate change, the massive introduction of genes from the Congolese could have consequences for the maintenance of the genetic integrity of wild coffee trees representing the endemic Guinean

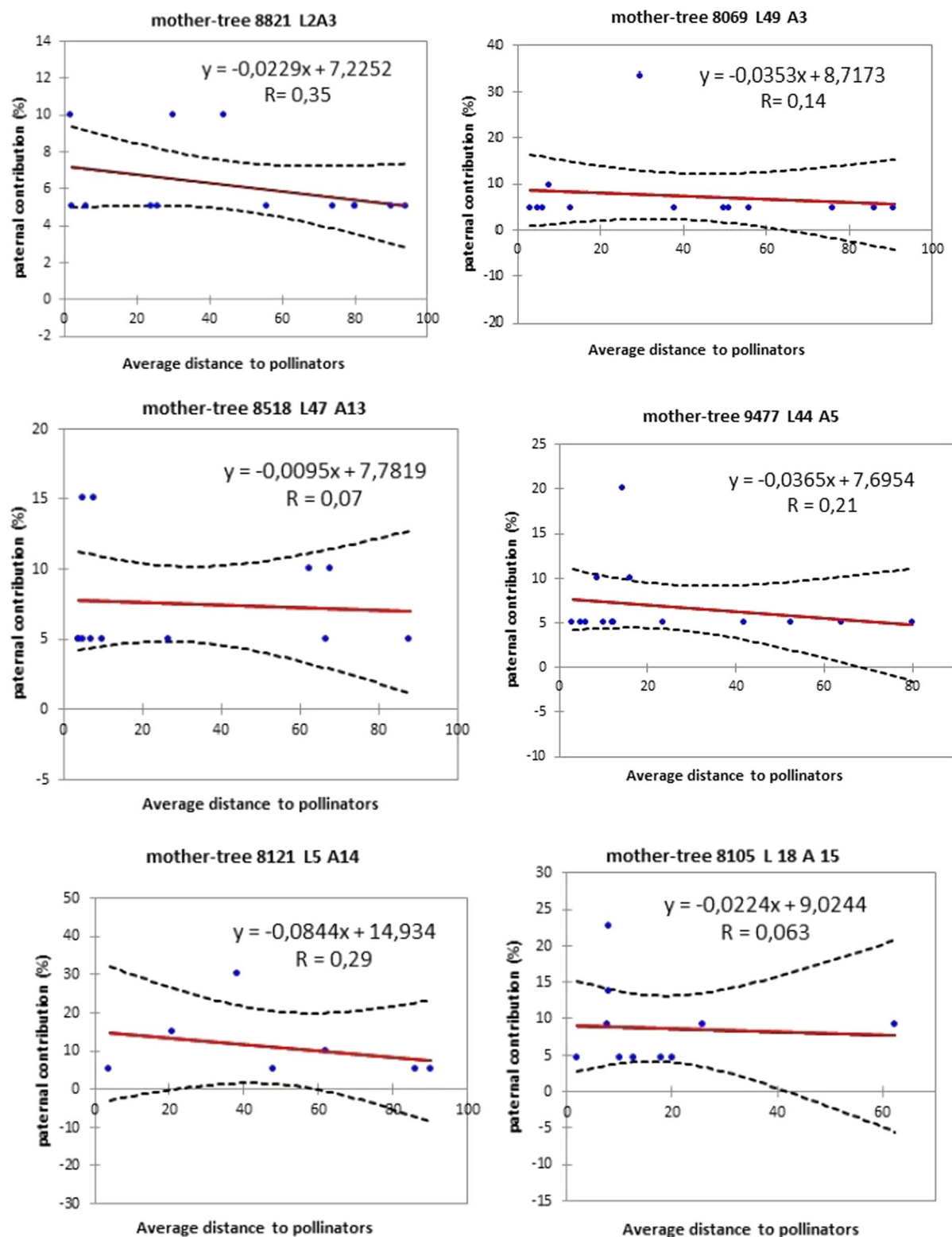


Fig. 6 Relationship between paternal contribution and pollinators distance

group in Ivory Coast. Indeed, in our system, flowering time differences between the two groups do not constitute a sufficient barrier to prevent massive hybridization. It is important to verify if this contamination is already present in the wild coffee trees of Ivory Coast. It is also necessary to verify that seeds could not have been dispersed by wildlife in the forests hosting the Guinean genetic resources. And finally, a comprehensive assessment of existing germplasm should be undertaken, based on which a comprehensive conservation strategy should be developed that includes both in situ and ex situ conservation and addresses the need for duplication of germplasm in multiple locations.

Author contributions Diane Estelle Gnapi: executed the experiment, collected field data, analyzed the data and wrote the paper; Désiré N'Da Pokou: designed the experiment, supervised the experiment and analyzed the data; Hyacinthe Legnaté: Designed the experiment, supervise the field data collection; Zhang Dapeng: generated the molecular data and reviewed; Benoît Bertrand: reviewed; Christophe Montagnon: Analyzed the data (field and molecular) and reviewed; Simon-Pierre Assonvo N'Guetta: supervised the work. Manuscript has been read and approved for submission by all the authors.

Funding This work is funded by the national agricultural research and extension fund “Firca”. The genotyping service is partially support by the USDA-ARS.

Declarations

Conflicts of interest I declare that there is no conflicts of interest associated with this publication. I confirm that the manuscript has been read and approved for publication in Euphytica by all authors.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Aerts R, Berecha G, Gijbels P, Hundera K, Van Glabeke S, Vandepitte K, Honnay O (2013) Genetic variation and risks of introgression in the wild *Coffea arabica* gene pool in south-western E thiopian montane rainforests. *Evol Appl* 6(2):243–252
- Akaffou DS (2013) Étude et cartographie génétique du croisement *Coffea pseudozanguebariae* BRIDSON X *Coffea canephora* PIERRE. Thèse de Doctorat d'état en sciences naturelles, spécialité: Génétique et Amélioration des plantes; Soutenue à l'Université Nangui Abrogoua (Côte d'Ivoire) 146 p.
- 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. <http://www.documentation.ird.fr/hor/fdi:16623>. Accessed 7 Nov 2017
- Berthaud J (1986) 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. Travaux et documents de l'ORSTOM 188:379
- Brou YT, Servat E, Paturel JE (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. IRD, UNESCO, Paris, pp 57–62
- Bunn C, Läderach P, Ovalle Rivera O, Kirschke D (2015) A bitter cup: climate change profile of global production of Arabica and Robusta coffee. *Clim Change* 129(1–2):89–101. <https://doi.org/10.1007/s10584-014-1306-x>
- Coste R (1989) Caféiers et cafés. G P Maisonneuve et Larose, Paris, p 373
- Coulibaly I, Noirot M, Lorieux M, Charrier A, Hamon S, Louarn J (2002) Introgression of self-compatibility from *Coffea heterocalyx* to the cultivated species *Coffea canephora*. *Theor Appl Genet* 105(6–7):994–999
- Davis AP, Rakotonasolo F (2021) Six new species of coffee (*Coffea*) from northern Madagascar. *Kew Bull* 76:497–511. <https://doi.org/10.1007/s12225-021-09952-5>
- Davis AP, Govaerts R, Bridson DM, Stoffelen P (2006) An annotated taxonomic conspectus of the genus *Coffea* (Rubiaceae). *Bot J Linn Soc* 152(4):465–512
- Davis AP, Tosh J, Ruch N, Fay MF (2011) Growing coffee: *Psilanthus* (Rubiaceae) subsumed on the basis of molecular and morphological data; implications for the size, morphology, distribution and evolutionary history of *Coffea*. *Bot J Linn Soc* 167(4):357–377
- Denoeud F, Carretero-Paulet L, Dereeper A, Droc G, Guyot R, Pietrella M, Lashermes P (2014) The coffee genome provides insight into the convergent evolution of caffeine biosynthesis. *Science* 345(6201):1181–1184
- Devreux M, Vallaeys G, Pochet P, Gilles A (1959) Recherches sur l'autostérilité du caféier robusta (*Coffea canephora* Pierre). Publications de l'Institut National pour l'Étude Agronomique du Congo Belge (INEAC), Bruxelles
- Ehounou JN (2021) Manifestations des changements climatiques en zone forestière de Côte d'Ivoire et impacts sur la disponibilité en eau pour une cacaoculture durable: cas de trois secteurs agroécologiques. Thèse de Doctorat

- unique, UFR STRM, Université de Cocody, Abidjan, Côte d'Ivoire, p 287
- Eskes AB, Leroy T (2012) Coffee selection and breeding. In: Wintgens JN (ed) Coffee: growing, processing, sustainable production: a guidebook for growers, processors, traders, and researchers. Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, pp 61–90
- FAOSTAT (2018) Food and Agriculture Organization of the United Nations, Statistics Division. <http://faostat.fao.org/>
- FAOSTAT (2019) Culture. <http://www.fao.org/faostat/fr/#data/QC>, consulté le 21/03/2019
- Fluidigm (2013) Fluidigm® SNP genotyping analysis software user guide. Rev I1, PN 68000098, South San Francisco, CA, Fluidigm Corporation. <https://www.wmsciencecomau/upload/pages/fluidigmtech/fluidigm-snp-genotyping-user-guide-151112pdf>. Accessed 10 July 2021
- Futuyma DJ (1998) Evolutionary Biology, 3rd edn. Sinauer Associates, Sunderland
- 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 Evol Biol 9:167–185
- ICO (2019) Coffee Development Report 2019 “Growing for Prosperity – Economic Viability as the Catalyst for a Sustainable Coffee Sector”. ICO FR-01- 19e, International Coffee Organization: London, UK
- ICO (2021) Coffee market report. September 2018. International Coffee Organization. <http://www.ico.org/documents/cy2020-21/cmr-0921-f.pdf>. Accessed 22 Feb 2022
- Kalinowski ST, Taper ML, Marshall TC (2007) Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. Mol Ecol 16:1099–1106
- Kassin KE (2009) Etudes de conditions pédoclimatiques pour la replantation cacaoyère au centre-ouest de la Côte d'Ivoire: cas des départements de Divo et Gagnoa. Thèse de Doctorat unique, UFR STRM, Université de Cocody, Abidjan, Côte d'Ivoire, p 112
- Labouisse JP, Curby P, Austerlitzs RR, Nguyen HA (2020) New insights on spatial genetic structure and diversity of *Coffea canephora* (Rubiaceae) in Upper Guinea based on old herbaria. Plant Ecol Evol 153(1):82–100. <https://doi.org/10.5091/plecevo.2020.1584>
- Leroy T, Montagnon C, Charrier A, Eskes AB (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
- Leroy T, Montagnon C, Cilas C, Yapo A, Charmettant P, Eskes AB (1997) Reciprocal recurrent selection applied to *Coffea canephora* Pierre. III. Genetic gains and results of first intergroups crosses. Euphytica 95:347–354
- Louarn J (1992) La fertilité des hybrides interspécifiques et les relations génomiques entre caféiers diploïdes d'origine africaine (Genre *Coffea* L sous-genre *Coffea*). Montpellier II, Sciences et Techniques du Languedoc, Montpellier, p 135
- Maalouf F, Khalil S, Ahmed S, Akintunde AN, Kharrat M, Shama'a KE, Hajjar S, Malhotra RS (2011) Yield stability of faba bean lines under diverse broomrape prone production environments. Field Crops Res 124:288–294
- Marshall TC, Slate J, Kruuk BE, Pemberton JM (1998) Statistical confidence for likelihood-based paternity inference in natural populations. Mol Ecol 7:639–655
- Meagher TR (1986) Analysis of paternity within a natural population of *Chamaelirium luteum*. I. Identification of most-likely male parents. Am Nat 128:199–215
- Montagnon C, Leroy T, Yapo A (1992a) Diversité génotypique et phénotypique de quelques groupes de caféiers (*Coffea canephora* Pierre) en collection. Conséquence Sur Leur Utilisation En Sélection Café Cacao Thé 36(3):187–198
- Montagnon C, Leroy T, Yapo A (1993) Caractérisation et évaluation de caféiers *Coffea canephora* prospectés dans des plantations de Côte-d'Ivoire. Café, Cacao, Thé 37:115–119
- Montagnon C, Leroy T, Cilas C, Legnaté H, Charrier A (2008) Heterozygous genotypes are efficient testers for assessing between-population combining ability in the reciprocal recurrent selection of *Coffea canephora*. Euphytica 160(1):101–110
- Montagnon C, Leroy T, Yapo A (1992b) Etude complémentaire de la diversité génotypique et phénotypique des caféiers de l'espèce *Coffea canephora* en collection en Côte d'Ivoire. XIV^e Colloque Scientifique International sur le Café. ASIC, San Francisco (USA), pp 444–450
- Montagnon C (2000) Optimisation des gains génétiques dans le schéma de sélection récurrente réciproque de *Coffea canephora* Pierre. Ecole Nationale Supérieure Agronomique de Montpellier, Thèse de Doctorat
- Oddou-Muratorio S, Klein EK, Vendramin G, Fady B (2011) Spatial vs. Temporal effects on demographic and genetic structures: the roles of dispersal, mating and differential mortality on patterns of recruitment in *Fagus sylvatica*. Mol Ecol 20:1997–2010
- Sagnard F, Oddou-Muratorio S, Pichot C, Vendramin GG, Fady B (2011) Effects of seed dispersal, adult tree and seedling density on the spatial genetic structure of regeneration at fine temporal and spatial scales. Tree Genet Genomes 7:37–48
- Shapiro SS, Wilk MB (1965) An analysis of variance test for normality (complete samples). Biometrika 52(3/4):591–611
- Thompson E (1976) Inference of genealogical structure. Soc Sci Inf 15:477–526
- Van Der Vossen H (2016) Narrowing the Yield Gap of Smallholder coffee Production as Sustainable Strategy to Meet Rising Demand for Quality Coffees. In: The programme and abstracts of the 26th International Conference on Coffee Science. November 13rd–17th 2016, Kunming, China
- Wang J, Lin M, Crenshaw A et al (2009) High-throughput single nucleotide polymorphism genotyping using nanofluidic Dynamic Arrays. BMC Genomics 10:561

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.