Plant Ecology and Evolution 158 (3): 358–366, 2025
https://doi.org/10.5091/plecevo.147801

RESEARCH ARTICLE

Development of new microsatellite markers for *Cola acuminata* (Malvaceae), a socio-economically important fruit tree species in Central Africa

Pablo Dupiol[1], Armel Chakocha[2], Marie-Louise Avana Tientcheu[2], Mohamed Mahamoud Charahabil[3], Cedric Mariac[1], Adeline Barnaud[1], Jérôme Duminil[1]

1 DIADE, IRD, University of Montpellier, CIRAD, Montpellier, France
2 Laboratory of Forestry, Faculty of Agronomy and Agricultural Sciences (FASA), University of Dschang, Dschang, Cameroon
3 Department of Agroforestry, University Assane Seck of Ziguincho, Ziguinchor, Senegal

Corresponding authors: Pablo Dupiol (pablo.dupiol@ird.fr) and Jérôme Duminil (jerome.duminil@ird.fr)

Academic editor: Katharina Budde ♦ **Received** 24 January 2025 ♦ **Accepted** 17 July 2025 ♦ **Published** 24 September 2025

Abstract

Background and aims – We developed a new set of nuclear microsatellite markers for *Cola acuminata* (Malvaceae), an important African food tree species commonly known as the kola nut. Probably originating from the tropical rainforests of the Congo Basin, *C. acuminata* is widely cultivated in the humid savannahs of the region where its nuts are sold throughout Central and West Africa for their stimulant properties. Nuclear microsatellite markers (SSRs) are well suited for assessing the genetic diversity and population structure of plant species due to their high variability.
Material and methods – Leaf samples were collected from 84 *C. acuminata* cultivated individuals across three sites in Cameroon, two in the savannah zone, one in the forest zone. SSR markers were developed by sequencing genomic DNA from two individuals using an Illumina HiSeq platform. Genetic diversity was assessed based on 14 SSR markers genotyped in 84 individuals, and marker transferability to the closely related species *Cola nitida* was tested.
Key results – Forty-eight new microsatellite loci were developed, of which 14 were polymorphic in *C. acuminata*. The results demonstrated a high level of genetic diversity with the presence of two to 33 alleles per locus (with an average of 14.14) across the three sites. The transferability of these markers was confirmed with 13 out of the 14 SSRs successfully amplifying in the closely related species, *Cola nitida*.
Conclusion – These newly developed SSRs will be useful for assessing genetic diversity, genetic differentiation, and gene flow patterns of *C. acuminata* in the tropical forests of Central Africa. Preliminary results suggest genetic similarity between the two savannah sites. However, these two sites were significantly differentiated from the site in the forest zone. This suggests that the propagation material introduced in the savannah zone did not originate from the forest in southern Cameroon.

Keywords

conservation, domestication history, food tree species, non-timber forest products, tree crop

INTRODUCTION

In Africa, an estimated 600 million people rely heavily on non-timber forest products (NTFPs) for their livelihood (Malleson et al. 2014). NTFPs include a wide range of resources like medicinal plants, fibres, resins, latex, fruits, seeds, construction materials, and bushmeat. All these products hold not only economical value but are also of medicinal, social, and spiritual significance, making them essential for the well-being and food security of local communities (Arnold and Pérez 2001). Despite their importance, NTFP species are threatened by several factors, such as changes in land use, non-sustainable harvesting practices, and climate change (Sunderland

Plant Ecology and Evolution is published by Meise Botanic Garden and the Royal Botanical Society of Belgium.

and Ndoye 2004). These pressures can erode genetic diversity by reducing the capacity of species to adapt to environmental changes. Conserving genetic diversity is vital for ensuring the resilience of species, maintaining the functions of ecosystems, and supporting the socio-ecological benefits provided by these resources (Cantwell-Jones et al. 2022).

The *Cola* Schott & Endl. genus of the Malvaceae family comprises around 125 species, including *Cola acuminata* (P.Beauv.) Schott & Endl., known as the kola tree, whose seeds (the kola nuts) hold a significant cultural value in many African communities for social and religious aspects (Lim 2012). Nuts are often presented as a gesture of respect and hospitality. Additionally, due to their high caffeine content, kola nuts are used to enhance alertness and combat fatigue (Sonibare et al. 2009). *Cola acuminata* is quite probably native to the tropical rainforests of the Congo Basin but is now widely cultivated in humid savannahs of Central Africa (Chevalier and Perrot 1911; Russell 1955). Despite its socio-economic importance (Starin 2013), there is limited information concerning the spatial distribution of its genetic diversity, which limits our understanding of its evolutionary and domestication history. To develop effective strategies of conservation, sustainable management, and utilization of its genetic resources, it is crucial to characterise its genetic diversity across different socio-ecological contexts (FAO 2014; Rimlinger et al. 2021).

So far, only one study investigated the genetic diversity of *C. acuminata* and *C. nitida* (Vent.) Schott & Endl, in Nigeria using Random Amplified Polymorphic DNA (RAPD) markers (Akinro et al. 2019). Given the limited reproducibility of RAPD markers (Penner et al. 1993), and the lack of genomic resources, we (i) developed a new set of polymorphic SSR markers in *C. acuminata*; (ii) evaluated the transferability of these markers to another socio-economically important *Cola* species, *C. nitida*, which is widely cultivated in West Africa; (iii) used 14 markers to characterize the distribution of genetic variability within and among cultivated populations of *C. acuminata* from Cameroon located either in the savannah zone (two populations), or in the forest zone (one population); (iv) used 13 markers to characterize the distribution of genetic variability within *C. nitida*, and between *C. acuminata* / *C. nitida* populations. The species most probably originates from the forest zone and has been introduced and subsequently cultivated in the savannah zone since several centuries (Chevalier and Perrot 1911). We hypothesised that the history of introduction of the species from the forest to the savannah zone has been accompanied by a genetic bottleneck (Doebley et al. 2006), and that the geographical proximity of the two cultivated savannah populations would facilitate the exchange of propagation material between them. Under these hypotheses, we expect: (i) relatively low genetic differentiation between the two cultivated savannah populations and higher genetic differentiation between these two populations and the cultivated population from the forest zone; (ii) cultivated populations from the savannah zone to exhibit reduced genetic diversity compared to the population from the forest region.

MATERIAL AND METHODS

Description of the species studied

The genus *Cola* contains two socio-economically important species known as kola trees, *Cola acuminata*, native to Central Africa, and *Cola nitida*, native to West Africa (Tachie-Obeng and Brown 2001). *Cola acuminata* occurs naturally in the rainforests of Central Africa, from Gabon to Nigeria. It is now cultivated throughout Central Africa, either in forested areas where it is spared during field preparation or planted voluntarily, mostly using local germplasm. It was introduced several centuries ago into the agroforests of the savannah zone using genetic material of unknown origin (Chevalier and Perrot 1911).

Cola acuminata can reach a height of 30 m and a diameter of about 30 cm (Tachie-Obeng and Brown 2001). The bark of the tree is grey, rough, and corky, dividing into squares (Russell 1955). The foliage is sparse and confined to the ends of the branches, leaves are dark green with elliptic shape and a long acumen compared to the cultivated sister species *Cola nitida* (Russell 1955). The flowers of *C. acuminata* are either male or hermaphrodites (Russell 1955). The fruits of *C. acuminata* are straight or slightly curved and rough to the touch and its leaves are curved, twisted, and acuminate on all the length of the leaf (Tachie-Obeng and Brown 2001). The tree produces fruits from April to June (Tachie-Obeng and Brown 2001). Its seeds are probably mainly dispersed by mammals, as observed in the sister species *C. lizae* N.Hallé (Tutin et al. 1991). The nut of *C. acuminata* is the main organ used by local people, either for its stimulant properties, as an aphrodisiac, or as a gift for special occasions like weddings (Lim 2012). The species is mainly propagated by its seeds, which are often exchanged between farmers. The seed has more than two cotyledons (three to five). In comparison, *C. nitida* seeds have only two cotyledons (Tachie-Obeng and Brown 2001). The seed of *C. acuminata* is commonly pink or red, rarely white, and tastes bitter in mouth (Russell 1955).

Study sites

We sampled leaf samples of cultivated individuals of *C. acuminata* from three sites in Cameroon, one in the forest zone and the other two in the savannah zone (Fig. 1). In each site, there were different ethnic groups: Bamoun people in Koupa-Matapit, Bamiléké people in Bametcha, and Boulou people in Meyo-Ville. Interviews were conducted at each of the sites on the uses and management practices of the species (data not shown), which will help us to interpret the results obtained on the distribution of the species' genetic diversity.

In the forest zone, trees were sampled in Meyo-Ville (3.55464°N, 11.93430°E; Mvila department, South region) and were corresponding either to wild individuals that were spared when the fields were prepared (slash-and-burn area), or were planted using local genetic material. In total, 28 individuals were sampled from four agroforest farms (corresponding to four different land owners).

In the savannah zone, trees were sampled in Koupa-Matapit, (5.76953°N, 10.80081°E; Noun department, West region) and in Bametcha (5.29322°N, 10.49694°E; Ndé department, West region). In both sites, 28 individuals were sampled from respectively 13 and nine agroforest plantations. The cultivated individuals sampled at these two sites were introduced several generations ago (probably several centuries ago) using genetic material of unknown origin.

Moreover, in order to test the transferability of SSR markers to the closely related *Cola nitida* species, 10 cultivated trees were collected from two different villages in Senegal, Adéane, and Agnack (Suppl. material 1).

Microsatellite development

DNA from two *C. acuminata* individuals from Koupa-Matapit village in Cameroon (AC051 and AC066; Suppl. material 1), was extracted following the protocol described by Mariac et al. (2006). For each sample, a genomic library was prepared as outlined in Mariac et al. (2014) and sequenced using paired-end sequencing (2 × 250 bp) on an Illumina HiSeq platform at Novogene (Munich, Germany). Sequencing yielded 1,326,072 and 1,551,002 reads for AC051 and AC066, respectively. Paired-end reads were merged using FLASH v.1.2.11 (Magoč and Salzberg 2011), resulting in 891,118 and 1,191,340 merged reads for AC051 and AC066, respectively. Microsatellite repeat motifs were detected and primers designed using the default parameters in the QDD software v.3.1.2 (Meglécz et al. 2010): PCR product size 90 to 300 bp; primer sizes 18 to 27 bp, with an optimum of 20; melting temperatures of 57°C to 63°C, with an optimum of 60°C and a max difference of 10°C; GC content ranges from 20% to 80%, with a 50% optimum. Among the 7501 and 6032 SSR-containing reads identified in AC051 and AC066 respectively, 48 were selected based on the following criteria: presence of perfect di- or trinucleotide motifs with at least ten repeat units and PCR primers located at a minimum of 20 bp distance from the microsatellite motif.

Microsatellite amplification

To enable multiplexed co-amplification of loci, we employed the three-primer PCR method described by Schuelke (2000) and further detailed by Culley et al. (2008). In this approach, each locus-specific forward

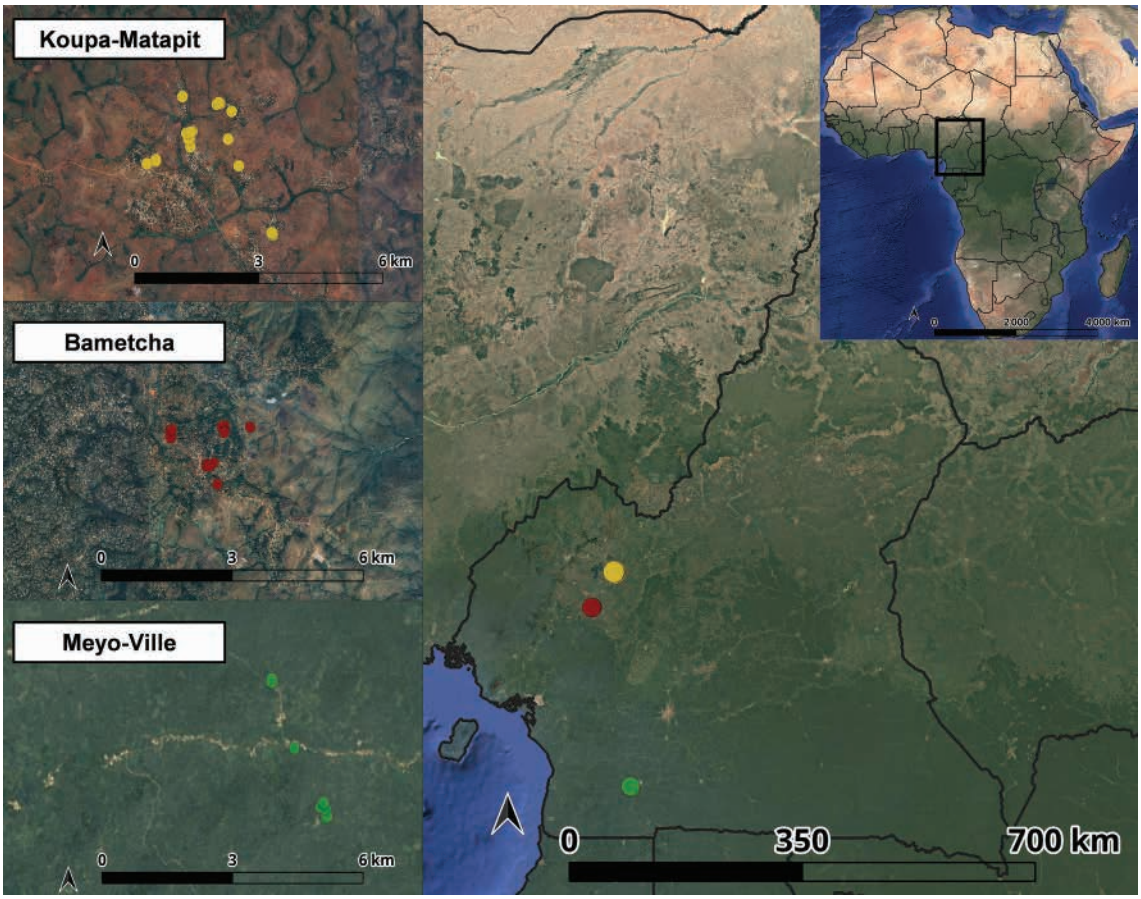

Figure 1. Sampling sites of *Cola acuminata* in Cameroon (Koupa-Matapit, Bametcha, and Meyo-Ville) and distribution of the samples in these sites.

Table 1. Characteristics of the 14 microsatellite markers developed for *Cola acuminata*. * Q1 = TGTAAAACGACGGCCAGT (Schuelke 2000), Q2 = TAGGAGTGCAGCAAGCAT, Q3 = ACTGCTTAGAGCGATGC, Q4 = CTAGTTATTGCTCAGCGGT (Q2–Q4 after Culley et al. 2008). ** Repeat number of di- or trinucleotide motif found in the microsatellite.

Microsatellite marker	Primer sequences (5'–3') F: forward, R: reverse	Tail and fluorescent label*	Motif and repeat number**	Allele size range (bp)	GenBank accession number
CA09	F: AGCCATAAGTGAAAGTGTGAAACT R: AGTTGCTGAAGTATGACTAAGCCA	Q1-6-FAM	$(AT)_{13}$	145–220	PP943415
CA18	F: TGGTAGAATTGCATAGGAGATTGA R: ACCAACCTCCAATCGCAACT	Q2-NED	$(AAT)_{17}$	140–170	PP943418
CA27	F: GCCAACACGATTCATTAACATATAGG R: ACCGGATATGAACTCTTTGCCA	Q3-VIC	$(AT)_{12}$	144–182	PP943421
CA36	F: AAGGGCCCTTGAAATGGTGG R: TGAGTAACGGTGTTGAGATCCT	Q3-VIC	$(AG)_{21}$	225–275	PP943423
CA05	F: TTTGTTAGCCACCTTCAAATCATT R: TGATTTCTTATTTCACTAGCAATGGC	Q4-PET	$(AG)_{10}$	135–175	PP943412
CA01	F: AGACGCAAGGAGTTCTTTCCA R: GGTTTCGTTGAATTGAATCAAACA	Q1-6-FAM	$(AT)_{10}$	116–146	PP943411
CA12	F: TTTCTTCCACGGCTAGATACAA R: TCAGTCTACAAAGAGGATAAGTAGG	Q1-6-FAM	$(AT)_{11}$	178–220	PP943416
CA23	F: CATAGGCTTCCTGGTGCACT R: TTGATGAGCTCAGATGGGCAT	Q2-NED	$(AAG)_{13}$	142–225	PP943420
CA06	F: TGTAGGTTTGGCTTGCCTCC R: CAAAGCATCCCAACAATCGCA	Q3-VIC	$(AT)_{10}$	135–175	PP943413
CA16	F: CCATCGATGGCTAGTTATGAGTT R: TGAGCCTTAGACTACTAGGGACA	Q4-PET	$(AT)_{12}$	105–265	PP943417
CA08	F: GTTTGTATTGAAGTGCAAACTGTT R: CACCTCCCTCCCTTCCCTTA	Q1-6-FAM	$(AT)_{16}$	145–180	PP943414
CA19	F: GTGCGTAAACACAGCCTCAG R: ACTTTAGTAAACCAACAGGAGGGA	Q2-NED	$(AGC)_{12}$	160–190	PP943419
CA34	F: TGAAATGGTCCTAAGTTACATCCA R: TCTAAACTTTGTGCGGGCCA	Q3-VIC	$(AT)_{10}$	195–228	PP943422
CA39	F: GCACTCAGAACGTTCTCCCT R: ACGAGATTCCTGTGCTGGTG	Q4-PET	$(AT)_{14}$	175–230	PP943424

primer was synthesised with one out of four unlabelled universal oligonucleotide tail added to its 5' end (see details in Table 1). A third primer, corresponding to the universal tail sequence but fluorescently labelled with either 6-FAM, NED, VIC, or PET, was then included in the PCR reaction. During amplification, the labelled universal primer binds to the tail sequence incorporated into the initial cycles, allowing for fluorescent labelling of the PCR products. Each of the 48 SSR was first tested individually for amplification (simplex) on four *C. acuminata* individuals, two from Koupa-Matapit and two from Meyo-Ville (Suppl. material 1). Ultimately, after excluding loci that showed poor amplification in PCR or lacked polymorphism, we retained only 14 microsatellite markers that met all filtering criteria, including thermodynamic suitability. The polymerase chain reactions (PCR) were conducted in a total volume of 10 µL, comprising the following components: 0.15 µL of the forward primer [10 µM], 0.1 µL of the reverse primer [10 µM], 0.15 µL of the corresponding fluorescently labelled primer [10 µM] (Q1, Q2, Q3, or Q4; see Table 1), 1 µL of DNA (10 ng/µL), 4 µL of PCR Master Mix (Type-it Microsatellite, Qiagen, Hilden, Germany), and the volume was completed to 10 µL with ultra-pure water (Invitrogen). PCR was carried out using a Veriti 96-Well thermal cycler (Thermo Fisher Scientific, Waltham, USA) with the following cycling conditions: initial denaturation at 95°C for 3 min followed by 30 cycles consisting of a denaturation step at 95°C for 30 s, annealing at 57°C for 90 s, elongation at 72°C for 30 s, followed by 10 similar cycles except for the annealing temperature that was set to 53°C, and a final extension at 60°C for 30 min. The genotyping was done on an ABI 3500 XL (Applied Biosystems, Thermo Fisher Scientific, Waltham, USA) at CIRAD

(Montpellier, France) using 1 µL of PCR product with 10 µL of Hi-Di Formamide and 0.12 µL of GeneScan 500 LIZ Size Standard. Electropherograms were visualised and scored in Geneious v.7.1.3. (Biomatters Ltd., Auckland, New Zealand).

The 14 polymorphic SSRs were then amplified for all 84 *C. acuminata* samples using three multiplex reactions (Multiplex 1: CA09-Q1, CA18-Q2, CA27-Q3, CA36-Q3, CA05-Q4; Multiplex 2: CA01-Q1, CA12-Q1, CA23-Q2, CA06-Q3, CA16-Q4; Multiplex 3: CA08-Q1, CA19-Q2, CA34-Q3, CA39-Q4). The multiplex PCR included 0.15 µL of each forward primer [10 µM], 0.10 µL of each reverse primer [10 µM], 0.15 µL of each tail (Q1–Q4) [10 µM], 1 µL of DNA, 7.5 µL of PCR Master Mix (Type-it Microsatellite, Qiagen), and was completed to a total volume of 15 µL with distilled water. PCRs and genotype scoring were carried out in the same way as for the simplex reactions above. The same protocol was used to test the transferability of the markers on 10 individuals of *C. nitida* coming from Senegal (Suppl. material 1).

Microsatellite diversity analysis

An analysis of genetic diversity was carried out on a dataset comprising 84 *C. acuminata* samples coming from three different Cameroonian populations (Suppl. material 1).

The SPAGeDi program v.1-5a (Hardy and Vekemans 2002) was used to perform genetic analysis on the Koupa-Matapit (N = 28), Bametcha (N = 28), and Meyo-Ville (N=28) individuals as well as to estimate preliminary genetic indicators for *C. nitida* individuals coming from Adéane (N = 9) and Agnack (N = 1) in Senegal. For each locus and each of the three populations of *C. acuminata*, we determined the number of alleles (*nA*), allelic richness for a subsample size of *k* (*k* = 30) gene copies (*AR*), the observed heterozygosity (H_o), the expected heterozygosity (H_e), and the inbreeding coefficient ($F_I = 1 - H_o / H_e$). We tested if F_I were significantly different from zero (test of Hardy-Weinberg equilibrium) after 999 permutations of gene copies among individuals. We also jointly estimated null allele frequencies (*fnull*) and the corrected inbreeding coefficient F_{IS} for each population under a population inbreeding model (PIM) accounting for inbreeding and for genotyping failures using INEST v.2.2 (Chybicki and Burczyk 2009).

We conducted a linkage disequilibrium test for each locus pair across all *C. acuminata* populations (Fisher's method) using Genepop v.4.7.5 (Raymond and Rousset 1995) with the following parameters: 1000 iterations per batch, 100 batches, and 1000 dememorization steps. Moreover, Micro-Checker v.2.2.3 was used to assess genotyping errors and null alleles (Van Oosterhout et al. 2004).

The genetic differentiation between populations (pairwise F_{ST}) was estimated using a nested ANOVA procedure following Weir and Cockerham (1984) in SPAGeDi. F_{ST} were calculated using the 14 SSR markers for estimations between *C. acuminata* populations, and using the 13 SSR markers in common for estimations between *C. acuminata* / *C. nitida* populations. We tested if the genetic differentiation between two populations was significant after 999 permutations of individuals between pairs of populations.

RESULTS

Microsatellites markers analyses in *C. acuminata*

We developed 14 new SSR markers in *C. acuminata* that proved to be highly polymorphic, with an average of 14.14 alleles per locus across all populations (Suppl. material 2). At the population level, the average number of alleles per locus was 9.14 in Meyo-Ville, 10.14 in Koupa-Matapit, and 7.36 in Bametcha (Table 2). The number of alleles per locus ranged from one (locus 6) to 19 (locus 16; Table 2). The multilocus average F_I, measuring inbreeding within populations, ranged from 0.228 in Bametcha to 0.305 in Meyo-Ville (Table 2). A number of loci were showing an excess of heterozygotes (F_I significantly superior to zero; see Table 2). After accounting for null alleles, the corrected F_{IS} values were 0.050 (95% highest posterior density interval [0.000–0.118]) for Koupa-Matapit, 0.038 (95% HPD [0.000–0.103]) for Bametcha, and 0.024 (95% HPD [0.000–0.068]) for Meyo-Ville.

We detected potential signs of scoring errors due to stuttering for four out of the fourteen loci (CA06, CA08, CA09, CA39) but we have not detected evidence of large allele dropout across all loci. Five out of the 91 locus pair comparisons showed signs of linkage disequilibrium (significant level of 0.05), but only one pair remained significant after applying a Bonferroni correction (loci 6 and 36).

Genetic differentiation between populations showed a relatively low F_{ST} value (0.023) between Bametcha and Koupa-Matapit, while differentiation between the two savannah sites and the forest site was higher (0.132 between Meyo Ville and Koupa Matapit and 0.196 between Meyo Ville and Bametcha; Suppl. material 3). All pairwise F_{ST} values were statistically significant (p value < 0.001).

Transferability of SSRs in *C. nitida*

Thirteen out of the fourteen SSR markers were transferable to *C. nitida*. Only CA12 was not transferable. Polymorphism was observed for all transferable loci except locus 9 (Suppl. material 4). The number of alleles per locus ranged from one (locus 9) to nine (locus 5). The allele size range in the *C. nitida* population was smaller compared to *C. acuminata* (Suppl. material 4). The F_{ST} value between *C. acuminata* populations and the *C. nitida* population was relatively high (F_{ST} > 0.250; Suppl. material 3).

Table 2. Intrapopulation genetic diversity of *Cola acuminata* in three Cameroonian populations at 14 newly developed microsatellite markers. N = number of individuals, *nA* = number of alleles, H_e = expected heterozygosity, H_o = observed heterozygosity, F_I = inbreeding coefficient with * p value < 0.05, ** p value < 0.01, *** p value < 0.001, *AR* = allelic richness, *fnull* = null allele frequency.

	Koupa-Matapit (N = 28)						Bametcha (N = 28)						Meyo-Ville (N = 28)					
Locus	***nA***	H_e	H_o	F_I	***AR (k = 30)***	***fnull***	***nA***	H_e	H_o	F_I	***AR (k = 30)***	***fnull***	***nA***	H_e	H_o	F_I	***AR (k = 30)***	***fnull***
CA01	9	0.8825	0.821	0.070	8.45	0.03	8	0.8487	0.750	0.118	7.30	0.06	10	0.8122	0.560	0.315***	8.71	0.06
CA05	13	0.9202	0.542	0.417***	11.42	0.00	12	0.7604	0.643	0.157	9.61	0.00	12	0.8998	0.500	0.450***	10.40	0.00
CA06	7	0.4682	0.148	0.688***	5.64	0.56	1	-	-	-	1.00	0.56	9	0.7409	0.571	0.232	7.13	0.56
CA08	4	0.5332	0.296	0.449*	3.36	0.33	4	0.5240	0.321	0.391*	3.32	0.24	6	0.7023	0.444	0.372**	5.53	0.26
CA09	2	0.1623	0.000	1.000**	1.99	0.00	2	0.2139	0.000	1.000**	2.00	0.00	2	0.3310	0.000	1.000**	2.00	0.00
CA12	13	0.9169	0.643	0.303***	11.48	0.18	10	0.8715	0.348	0.606***	9.05	0.51	11	0.8449	0.240	0.720***	9.03	0.51
CA16	19	0.9266	0.630	0.325***	14.55	0.21	13	0.8974	0.550	0.393***	11.61	0.38	16	0.9053	0.640	0.297***	12.98	0.38
CA18	9	0.8672	0.778	0.105	8.35	0.08	7	0.8041	0.680	0.157	6.53	0.16	5	0.7135	0.704	0.014	4.55	0.16
CA19	9	0.7305	0.679	0.072	6.97	0.04	5	0.7016	0.593	0.158	4.54	0.12	12	0.8422	0.786	0.068	9.33	0.03
CA23	8	0.8092	0.481	0.410***	7.48	0.00	5	0.6513	0.481	0.264*	4.54	0.00	5	0.7376	0.423	0.431***	4.92	0.00
CA27	11	0.8833	0.741	0.164*	10.15	0.00	8	0.8537	0.769	0.101	7.49	0.00	13	0.8945	0.926	-0.036	10.91	0.00
CA34	11	0.8718	0.600	0.316***	9.42	0.25	8	0.7968	0.857	-0.077	6.57	0.01	2	0.4832	0.000	1.000***	2.00	0.01
CA36	11	0.7792	0.714	0.085	8.16	0.00	7	0.6534	0.630	0.037	5.91	0.00	13	0.8279	0.857	-0.036	9.59	0.00
CA39	16	0.8903	0.750	0.160*	12.14	0.09	13	0.8700	0.704	0.194	10.19	0.09	12	0.8868	0.778	0.125	10.61	0.09
Multilocus average	10.14	0.760	0.559	0.269***	8.54	0.13	7.36	0.674	0.523	0.228***	6.40	0.15	9.14	0.758	0.531	0.305***	7.69	0.15

DISCUSSION

The high polymorphism observed in the 14 SSR markers developed for *C. acuminata* demonstrates their effectiveness in assessing the distribution of genetic diversity within this species. The presence of null alleles in several loci, such as in loci CA23, CA36, and CA09, leads to an underestimation of heterozygosity and high F_I values. The F_I values corrected for the presence of null alleles were close to zero, suggesting the absence of inbreeding or genetic substructure within the populations.

Our results provide preliminary insights into the evolutionary history of *C. acuminata* in Cameroon, particularly with regard to its cultivation history. The relatively low F_{ST} value (0.023) between the Bametcha and Koupa-Matapit populations suggests high gene flow and genetic similarity between them. This is consistent with observed local seed exchange practices and trade routes between these populations. The relatively higher genetic differentiation between the two cultivated populations from the savannah region and the cultivated population from the forest region, as reflected by higher F_{ST} values than those obtained between the two savannah populations, can be interpreted in different ways. As we have no information on the population of origin of planting material now present in the cultivated populations, and having no wild samples included in our study, our results need to be interpreted with caution. The cultivated population of Meyo-Ville is close to southern Cameroon rainforests, so that we can speculate that *C. acuminata* in this population represents remnant trees from a previously natural population and/or was introduced using wild material from southern Cameroon. If we now consider that the introduction of *C. acuminata* in the savannah zone also involved propagation material originating from southern Cameroon rainforests, the pattern of genetic divergence observed between the savannah and forest zone may be the result of a genetic bottleneck effect resulting from the introduction and subsequent cultivation process of *C. acuminata* in the savannah zone. This might also be due to the management practices employed in the two cultivated populations in the savannah region. In particular, the selection practices applied to these populations may have led to a loss of genetic diversity and genetic differentiation between the populations of origin in the forest and the introduced populations in the savannah. The genetic bottleneck effect and the effect of management practices are not mutually exclusive.

Alternatively, we could speculate that the original planting material used when the species was introduced to the savannah actually came from sources other than the rainforests in the south of Cameroon. Based on the ethnobotanical data that we have collected (data not shown), and from data obtained on other species (Yogom et al. 2020, 2023; Chakocha Ngandjui et al. 2025), cultivated material in the savannah region appears to originate from the rainforests of Cameroon's North West and South West regions or Nigeria. Given the extensive kola nut trade network, it is even possible that the source material originates from more distant areas. If the original planting material used to introduce the species to the savannah region came from rainforests other than those in southern Cameroon, the observed pattern of genetic differentiation between the forest and savannah zones might merely reflect the existing genetic differentiation between the wild populations that gave rise to the cultivated populations in these two zones. In any case, a much larger sample size including additional cultivated and wild populations of *C. acuminata* is required to better understand the species' cultivation history.

The levels of genetic diversity are relatively high in all three populations. Despite a low genetic differentiation between the two cultivated populations from the savannah zone, we observed differences in their levels of genetic diversity, with higher allelic richness in Koupa-Matapit than in Bametcha (ANOVA procedure, p value < 0.001). This result can be interpreted in light of social factors. During the sampling of the plant material, we conducted interviews with farmers in Koupa-Matapit, Bametcha and Meyo-Ville, addressing notably the origin of the genetic material used for plantation (data not shown). As discussed, the low level of genetic differentiation between the two savannah populations suggests that they share a common ancestry and/or experience extensive gene flow (i.e. exchange of planting material). However, differences in management practices between the two populations might have adversely influenced the levels of genetic diversity of *C. acuminata* in these two cultivated populations. In fact, farmers in Koupa-Matapit often source seeds for planting from local markets, whereas those in Bametcha tend to use their own seeds. Sourcing seeds locally in Bametcha, with no or very limited input of external genetic material, might reduce the genetic diversity of the species through genetic drift (reduction in population size). The selection of individuals with desirable traits by farmers can also result in a further reduction of genetic diversity in this population. These two processes can act in synergy.

The successful transferability of SSR markers to *C. nitida* highlights their potential utility for cross-species genetic studies within this genus. The reduced polymorphism and narrower allele size range observed in *C. nitida* may be due to the small size of the tested population. The high transferability rate of the SSR markers between the two species (93% based on our results) and the previous demonstration of hybridisation events between them (Adebola 2000) suggest that the two species are closely related.

CONCLUSION

Fourteen new polymorphic microsatellite markers were developed for the food tree *C. acuminata*. These markers enable to study the distribution of genetic diversity at different geographical scales, improving our

understanding of the domestication history and gene flow patterns within this species. The low level of genetic differentiation between the two cultivated populations in the savannah zone suggests a common origin and/or gene flow, probably facilitated by seed exchange and trade routes. In contrast, the higher genetic differentiation between the savannah and forest populations possibly indicated distinct source populations or could reflect differences due to the introduction and subsequent cultivation practices in the savannah zone. A higher level of genetic diversity was observed in the Koupa-Matapit population than in the Bametcha population, potentially due to differences in seed selection practices by the farmers. This demonstrates the importance of characterising the impact of farmers' management practices on the genetic diversity of cultivated species, in order to develop sustainable management strategies to maintain their genetic resources. The high transferability success rate of the newly-developed genetic markers to another *Cola* species, *C. nitida*, suggests that these markers may be useful for conducting genetic studies across these species.

DATA AVAILABILITY STATEMENT

The data [and related documentations] that support the findings of this study are openly available in DataSuds repository (IRD, France) at https://doi.org/10.23708/6FCSLF. Data reuse is granted under CC-BY license 4.0.

ACKNOWLEDGEMENTS

The authors acknowledge the ISO 9001 certified IRD i-Trop HPC (South Green Platform) at IRD Montpellier for providing HPC resources that have contributed to the research results reported within this paper. We are deeply grateful to the local farmers and communities that helped us collecting plant material in Cameroon and Senegal. We would like to acknowledge the help of two anonymous reviewers and Katharina Budde, the subject editor for Plant Ecology and Evolution, in improving the quality of the manuscript.

REFERENCES

Adebola PO (2000) Principal component analysis of variation pattern among *Cola nitida*, *Cola acuminata* (Sterculiaceae) and their interspecific hybrids. Feddes Repertorium 111(3/4): 183–188. https://doi.org/10.1002/fedr.20001110312

Akinro L, Adesoye A, Fasola T (2019) Genetic diversity in *Cola acuminata* and *Cola nitida* using RAPD primers. Genetika 51: 227–236. https://doi.org/10.2298/GENSR1901227A

Arnold JEM, Pérez MR (2001) Can non-timber forest products match tropical forest conservation and development objectives? Ecological Economics 39: 437–447. https://doi.org/10.1016/S0921-8009(01)00236-1

Cantwell-Jones A, Ball J, Collar D, Diazgranados M, Douglas R, Forest F, Hawkins J, Howes M-JR, Ulian T, Vaitla B, Pironon S (2022) Global plant diversity as a reservoir of micronutrients for humanity. Nature Plants 8: 225–232. https://doi.org/10.1038/s41477-022-01100-6

Chakocha Ngandjui AF, Avana Tientcheu ML, Momo Solefack MC, Rimlinger A, Meguem Mboujda FM, Duminil J (2025) Ethnoecological knowledge and management strategies for *Canarium schweinfurthii* (Burseraceae) in the forest and savannah zones of Cameroon. Economic Botany 79: 192–210. https://doi.org/10.1007/s12231-025-09633-5

Chevalier A, Perrot E (1911) Les Végétaux Utiles de l'Afrique Tropicale Française : Études Scientifiques et Agronomiques. Fascicule VI. Les Kolatiers et les Noix de Kola. A. Challamel, Paris, 1–483.

Chybicki IJ, Burczyk J (2009) Simultaneous estimation of null alleles and inbreeding coefficients. Journal of Heredity 100: 106–113. https://doi.org/10.1093/jhered/esn088

Culley TM, Weller SG, Sakai AK, Putnam KA (2008) Characterization of microsatellite loci in the Hawaiian endemic shrub *Schiedea adamantis* (Caryophyllaceae) and amplification in related species and genera. Molecular Ecology Resources 8: 1081–1084. https://doi.org/10.1111/j.1755-0998.2008.02161.x

Doebley JF, Gaut BS, Smith BD (2006) The molecular genetics of crop domestication. Cell 127 (7): 1309–1321. https://doi.org/10.1016/j.cell.2006.12.006

FAO (2014) Global plan of action for the conservation, sustainable use and development of forest genetic resources. https://openknowledge.fao.org/server/api/core/bitstreams/7dd342ac-4438-4b4a-b9a4-a97f5e11afa8/content

Hardy OJ, Vekemans X (2002) spagedi: a versatile computer program to analyse spatial genetic structure at the individual or population levels. Molecular Ecology Notes 2: 618–620. https://doi.org/10.1046/j.1471-8286.2002.00305.x

Lim TK (2012) *Cola acuminata*. In: Lim TK (Ed.) Edible Medicinal and Non Medicinal Plants. Volume 3, Fruits. Springer, Dordrecht, 168–174. https://doi.org/10.1007/978-94-007-2534-8_22

Magoč T, Salzberg SL (2011) FLASH: fast length adjustment of short reads to improve genome assemblies. Bioinformatics 27: 2957–2963. https://doi.org/10.1093/bioinformatics/btr507

Malleson R, Asaha S, Egot M, Kshatriya M, Marshall E, Obeng-Okrah K, Sunderland T (2014) Non-timber forest products income from forest landscapes of Cameroon, Ghana and Nigeria – an incidental or integral contribution to sustaining rural livelihoods? International Forestry Review 16: 261–277. https://doi.org/10.1505/146554814812572449

Mariac C, Luong V, Kapran I, Mamadou A, Sagnard F, Deu M, Chantereau J, Gerard B, Ndjeunga J, Bezançon G, Pham J-L, Vigouroux Y (2006) Diversity of wild and cultivated pearl millet accessions (*Pennisetum glaucum* [L.] R. Br.) in Niger assessed by microsatellite markers. Theoretical and Applied Genetics 114: 49–58. https://doi.org/10.1007/s00122-006-0409-9

Mariac C, Scarcelli N, Pouzadou J, Barnaud A, Billot C, Faye A, Kougbeadjo A, Maillol V, Martin G, Sabot F, Santoni S, Vigouroux Y, Couvreur TLP (2014) Cost-effective enrichment hybridization capture of chloroplast genomes at deep multiplexing levels for population genetics and phylogeography studies. Molecular Ecology Resources 14: 1103–1113. https://doi.org/10.1111/1755-0998.12258

Meglécz E, Costedoat C, Dubut V, Gilles A, Malausa T, Pech N, Martin J-F (2010) QDD: a user-friendly program to select microsatellite markers and design primers from large sequencing projects. Bioinformatics 26: 403–404. https://doi.org/10.1093/bioinformatics/btp670

Penner GA, Bush A, Wise R, Kim W, Domier L, Kasha K, Laroche A, Scoles G, Molnar SJ, Fedak G (1993) Reproducibility of random amplified polymorphic DNA

(RAPD) analysis among laboratories. Genome Research 2: 341–345. https://doi.org/10.1101/gr.2.4.341
Raymond M, Rousset F (1995) GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. Journal of Heredity 86: 248–249. https://doi.org/10.1093/oxfordjournals.jhered.a111573
Rimlinger A, Avana M-L, Awono A, Chakocha A, Gakwavu A, Lemoine T, Marie L, Mboujda F, Vigouroux Y, Johnson V, Vinceti B, Carrière SM, Duminil J (2021) Trees and their seed networks: the social dynamics of urban fruit trees and implications for genetic diversity. PLoS ONE 16(3): e0243017. https://doi.org/10.1371/journal.pone.0243017
Russell TA (1955) The Kola of Nigeria and the Cameroons. Tropical Agriculture 32(3): 210–240.
Schuelke M (2000) An economic method for the fluorescent labeling of PCR fragments. Nature Biotechnology 18: 233–234. https://doi.org/10.1038/72708
Sonibare M, Soladoye M, Esan O, Sonibare O (2009) Phytochemical and antimicrobial studies of four species of *Cola* Schott & Endl. (Sterculiaceae). African Journal of Traditional, Complementary and Alternative Medicines 6: 518–525. https://doi.org/10.4314/ajtcam.v6i4.57182
Starin D (2013) Kola nut: so much more than just a nut. Journal of the Royal Society of Medicine 106: 510–512. https://doi.org/10.1177/0141076813507708
Sunderland TCH, Ndoye O (2004) Forest products, livelihoods and conservation: case studies of non-timber forest product systems. Volume 2 - Africa. CIFOR, Jakarta.
Tachie-Obeng E, Brown N (2001) *Cola nitida* & *Cola acuminata*. Biodiversity Support Program, Washington DC., 1–37.
Tutin CEG, Williamson EA, Rogers ME, Fernandez M (1991) A case study of a plant-animal relationship: *Cola lizae* and lowland gorillas in the Lopé Reserve, Gabon. Journal of Tropical Ecology 7: 181–199. https://doi.org/10.1017/S0266467400005320
Van Oosterhout C, Hutchinson WF, Wills DPM, Shipley P (2004) MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. Molecular Ecology Notes 4: 535–538. https://doi.org/10.1111/j.1471-8286.2004.00684.x
Weir BS, Cockerham CC (1984) Estimating F-statistics for the analysis of population structure. Evolution 38: 1358–1370. https://doi.org/10.2307/2408641
Yogom BT, Avana-Tientcheu ML, Mboujda MFM, Momo ST, Fonkou T, Tsobeng A, Barnaud A, Duminil J (2020) Ethnicity differences in use and management practices of bitter kola trees (*Garcinia kola*) in Cameroon. Economic Botany 74: 429–444. https://doi.org/10.1007/s12231-020-09508-x
Yogom B, Avana-Tientcheu ML, Monthé FK, Bissiengou P, Loumeto JJ, Zekraoui L, Hervouet C, Fonkou T, Mariac C, Duminil J (2023) Genetic diversity and structure in wild and cultivated populations of an emblematic African tree species, *Garcinia kola* (Clusiaceae). Tree Genetics & Genomes 19: 39. https://doi.org/10.1007/s11295-023-01614-w

SUPPLEMENTARY MATERIALS

Supplementary material 1

List of *Cola acuminata* and *C. nitida* individuals used in this study.
https://doi.org/10.5091/plecevo.147801.suppl1

Supplementary material 2

Number of alleles (*nA*) and allelic richness ($k = 30$) of 14 newly developed microsatellite markers in *Cola acuminata* across three Cameroonian populations.
https://doi.org/10.5091/plecevo.147801.suppl2

Supplementary material 3

Pairwise F_{ST} values between the four sites. F_{ST} were calculated using the 14 SSR markers for estimations between *Cola acuminata* populations, and using the 13 SSR markers in common for estimations between *C. acuminata* / *C. nitida* populations.
https://doi.org/10.5091/plecevo.147801.suppl3

Supplementary material 4

Genetic properties of the newly developed 13 SSRs in *Cola nitida* (N = 10). *nA* = number of alleles sampled, H_e = expected heterozygosity, H_o = observed heterozygosity, F_I = inbreeding coefficient with * p value < 0.05 and ** p value < 0.01, *AR (k = 8)* = allelic richness.
https://doi.org/10.5091/plecevo.147801.suppl4