Journal of Wood Science

ORIGINAL ARTICLE Open Access

Wood spectral variability in *Entandrophragma cylindricum* across northern Congo: limited association with genetic structure despite high genetic diversity

Boniface T. Yogom[1*], Bouka U. D. Gaël[1], Makouanzi Ekomono C. G.[8], Gilles Chaix[4,5,6*], Olivier J. Hardy[9], Patrick Langbour[3,4], Grace Jopaul Loubota Panzou[1,2], Yannick Bocko[1], Styven Mahoungou[1], Eléngué Jospin[1], Reine Niémé[1], Raina Mbou[1], Ray Dzalamou[1], Jean-Joël Loumeto[1] and Charles Doumenge[6,7]

Abstract

Wood plays a crucial role in the economies of Central African countries, yet knowledge of the ecological and genetic determinants of intraspecific wood quality variation remains limited. In this study, we combined microsatellite-based population genetic analyses with short-wave near-infrared (NIR) hyperspectral imaging (HSI) of transverse wood slices to investigate the genetic diversity and wood spectral variability of three *Entandrophragma cylindricum* populations across distinct forest management units in northern Republic of the Congo. Genotyping of adult trees using 10 nuclear microsatellite loci revealed high levels of genetic diversity and low inbreeding across all populations. No significant genetic differentiation was detected among sites, although a weak but significant spatial genetic structure was observed, consistent with an isolation-by-distance (IBD) pattern, suggesting ongoing gene flow over the landscape. Hyperspectral imaging of transverse wood samples highlighted substantial inter-individual and site-related variability in NIR spectral signatures, with the highest spectral heterogeneity observed in the site characterized by the long-term logging history. However, this variability was not correlated with genetic distance, as confirmed by Mantel tests, indicating that site-specific context and logging intensity are the primary drivers of wood spectral variation. The results suggest that *E. cylindricum* populations retain considerable genetic potential for conservation and improvement. Moreover, the observed site-related variation in wood spectral traits highlights the importance of considering local growing context when interpreting wood spectral variability in tropical timber species. Further research combining near-infrared spectroscopy with detailed wood anatomical analyses and direct measurements of environmental variables would be necessary to disentangle the mechanisms underlying site-associated spectral variation.

Keywords Wood spectral variation, NIR hyperspectral imaging, Genetic diversity, Mantel test, *Entandrophragma cylindricum*

*Correspondence:
Boniface T. Yogom
yogomboni@yahoo.com
Gilles Chaix
gilles.chaix@cirad.fr
Full list of author information is available at the end of the article

© The Author(s) 2026. **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/.

Introduction

Tropical forests of Central Africa play a crucial role in global biodiversity conservation and regional economies, particularly through timber production, which represents a major source of revenue and livelihoods for local populations [1]. Despite the implementation of forest management plans over the past decades, sustainable timber exploitation still relies largely on conventional forest inventories that primarily consider tree abundance, diameter distribution, and species composition [1–3]. These approaches often overlook fine-scale intraspecific variability, especially variability expressed in wood-related characteristics, which may reflect differences among individuals and sampling sites [4–6].

Intraspecific variability in wood-related characteristics is increasingly recognized as an important component of tree functional diversity [7–9]. Such variability arises from complex interactions between genetic background, local growing context, and tree growth history, and can be expressed through phenotypic plasticity across heterogeneous forest landscapes [10–12]. While genetic variation can contribute to phenotypic differentiation among populations, weak genetic structuring is often reported in widespread tropical tree species characterized by extensive gene flow [13–18]. Disentangling the respective influence of genetic structure and site-related context therefore remains a major challenge in tropical forest ecology.

Near-infrared (NIR) hyperspectral imaging has emerged as a powerful approach for characterizing subtle variation in wood spectral signatures at high spatial resolution [19–23]. NIR spectra integrate information related to molecular bonds, particularly O–H, C–H, and N–H functional groups, and thus capture integrative signals associated with moisture-related features and unmeasured physico-chemical variability of wood tissues [24–26]. Unlike classical wood trait measurements, hyperspectral imaging enables high-throughput exploration of wood spectral variability within and among individuals. However, NIR spectral variation should be interpreted as an integrative phenotypic signal rather than a direct proxy for specific technological or mechanical wood properties, unless explicitly calibrated against independent measurements.

To explore the potential drivers of wood spectral variability, combining hyperspectral data with population genetic approaches offers a promising framework. Nuclear microsatellite markers provide robust insights into genetic diversity, population structure, and gene flow [27–32], allowing assessment of whether spectral variation among individuals is spatially structured by genetic differentiation or not.

The main aim of this study was to characterize hyperspectral (NIR) wood variation across three contrasting sites and to assess whether this variation is associated with genetic differentiation or consistent with site-related patterns. For this study, we selected *Entandrophragma cylindricum* (Sprague), commonly known as sapelli, a widespread and economically important timber species in Central African forests [1, 33, 34]. *E. cylindricum* is a key species of the Guineo-Congolian rainforest, belonging to the Meliaceae family, with a distribution extending from Sierra Leone to Uganda, excluding Togo and Benin [35–37]. Previous studies have documented substantial variability in wood-related characteristics across its distribution range [33], but the relative contribution of genetic structure to this variability remains poorly resolved [17, 38, 39]. This study focuses on three forest management units (FMUs) in northern Republic of the Congo, providing a suitable context to examine spatial patterns of wood spectral variability.

Specifically, the objectives of this study were threefold: (1) to characterize the genetic diversity and population structure of *E. cylindricum* across three sampling sites using nuclear microsatellite markers; (2) to quantify intraspecific and site-related variability in wood NIR spectral signatures using hyperspectral imaging; and (3) to test whether spectral dissimilarities among individuals are associated with genetic distances. We hypothesized that, given the expected weak genetic differentiation among populations, spatial variation in wood NIR spectral signatures would show limited association with genetic structure. By integrating population genetics and hyperspectral imaging, this study aims to clarify the extent to which observed wood spectral variability is genetically structured in a key tropical timber species.

Materials and methods

Study species and study sites

Entandrophragma cylindricum can grow up to 60 m tall with diameters exceeding 2 m [40]. It typically occurs at elevations below 1,000 m, on well-drained soils, in regions with annual rainfall between 1,600 and 2,000 mm and average temperatures ranging from 24 to 26 °C [41]. Despite high reproductive potential, natural regeneration is limited due to low seedling survival rates [36, 42]. Wood and genetic samples were collected from three FMUs located in northern Republic of the Congo: Kabo, Loundoungou, and Mimbeli. These FMUs differ in geomorphology, soil characteristics, hydrological context, and logging history [37, 43, 44], providing contrasted site contexts for assessing spatial patterns of genetic and wood spectral variation, without incorporating quantitative environmental parameter measurements, in accordance with the scope and objectives of the study (Table 1).

Table 1 Description of the study sites according to CIB–Olam Agri management plans

Site	Key ecological and environmental features
Kabo (1.8475875; 16.1770067)	Geomorphology: located in the Sangha basin with mostly Precambrian schist-quartz and lateritic substrates Soils: are ferralitic, acidic, and hydromorphic, with low fertility. The terrain is mostly flat with slight undulations and numerous swampy areas Human impact: the site has a long history of selective logging (since 1949), contributing to increased canopy openness and microenvironmental heterogeneity Vegetation: heterogeneous, including semi-evergreen terra firme forests and extensive swamp forests
Loundoungou (1.95498; 17.18932, at 123 km of Kabo and at 136 km from Mimbeli)	Geomorphology: dominated by Quaternary alluvial deposits with some Tertiary sandstones Soils: are deep, very acidic, ferralitic, and vary from sandy to clayey textures Human impact: logging began around 1992 Vegetation: the landscape is flat to gently dissected, with seasonal flooding and floodplain systems, particularly around the Motaba river. Forest types include diverse mixed-species terra firme stands, monodominant *Gilbertiodendron dewevrei* forests, and marshy depressions
Mimbeli (2.89321667; 17.9766433, at 240 km from Kabo)	Geomorphology: situated on Quaternary alluvium and Jurassic–Cretaceous sandstones (Carnot series). The relief is mostly flat to gently sloped (≤ 10%) Soils: are ferralitic and hydromorphic, very acidic, prone to erosion, and of low fertility Human impact: logging began since 1986, but less intensive Vegetation: forest cover is largely dense humid forest, with extensive swamp forests in the eastern zones, possibly linked to post-glacial forest refugia. Climate is a humid equatorial and seasonal waterlogging in low-lying areas

Sampling design

A total of 177 adult trees (with diameter at breast height (DBH) ≥ 90 cm) of *E. cylindricum* were sampled for genetic analyses across the three FMUs (Kabo = 51, Loundoungou = 80 and Mimbeli = 46). Due to logistical and material constraints, wood samples suitable for hyperspectral analysis were collected from a subset of 21 trees distributed among the sites (Kabo = 7, Loundoungou = 6, and Mimbeli = 8). For each tree, a radial wood block was extracted using a chainsaw along the longest radius of the stump to capture growth-related variability. Each block measured approximately 10 cm in the radial direction and 5 cm in the tangential direction, extending continuously from the pith to the bark. Trees were selected to cover the spatial extent of each FMU and to minimize spatial clustering. Genetic and wood samples were collected independently, and the reduced number of trees analyzed spectrally implies that wood spectral results should be interpreted as exploratory and indicative of site-related trends rather than exhaustive representations of within-site variability.

Genetic sampling, DNA extraction and nuclear microsatellites genotyping

For each sampled tree, leaf or cambial tissue was collected in the field and preserved in silica gel until laboratory processing. Genomic deoxyribonucleic acid (DNA) was extracted from 15 to 25 mg of dried leaf or cambial tissue using the NucleoSpin plant kit (Macherey-Nagel) following the manufacturer's instructions [45]. Genetic diversity and population structure were assessed using 10 nuclear microsatellite loci specifically developed for *E. cylindricum* and organized into two multiplex polymerase chain reaction (PCR) sets [32, 39]. Microsatellite loci were amplified in two developed multiplexes following the protocol of Micheneau et al.[46]. PCR amplifications were performed using the Type it Microsatellite PCR Kit (QIAGEN, Valencia, CA) as follows: 1 µl DNA, 7.5 µl of Multiplex Master Mix, 3.3 µl or 2.2 µl, respectively, for Mix 1 and 2. The PCR volume was adjusted to a final volume of 15 µl with ultrapure water. PCR was run on a TProfessional Basic Thermocycler (Biometra, Göttingen, Germany) using the following cycling conditions: initial denaturation at 95 °C for 5 min, then 30 cycles at 95 °C for 30 s, 57 °C for 90 s and 72 °C for 30 s, followed by 10 cycles at 95 °C for 30 s, 57 °C for 90 s and 72 °C for 30 s, and a final extension at 60 °C for 30 min. For genotyping, a mix of 1µl of PCR product, 12 µl of Hi-Di formamide (Life Technologies) and 0.3 µl of GeneScan *500 LIZ* size standard (Applied Biosystems) were used. Samples were analyzed on a 48-capillary *3730xl DNA Analyzer (Applied Biosystems)*. Genotype datasets were screened for scoring errors, null alleles, and deviations from Hardy–Weinberg equilibrium (HWE) prior to downstream analyses.

Wood spectral data acquisition (NIR hyperspectral imaging)

Radial wood samples were collected from freshly cut logs using a chainsaw. Samples were prepared as rectangular blocks exposing a transverse surface perpendicular to the stem axis. Each sample was assigned a unique identification code and transported to the Centre de Coopération

Internationale en Recherche Agronomique pour le Développement (CIRAD, France). Prior to spectral acquisition, samples were stabilized at a theoretical moisture content of 12% in a climatic chamber maintained at 20 °C and 65% relative humidity. To obtain a smooth and homogeneous surface suitable for hyperspectral imaging, one transverse face of each block was successively sanded using 180, 240, and 360 grit sandpaper.

Near-infrared hyperspectral images were acquired using a stationary short-wave infrared (SWIR) chemical imaging camera (Specim®, Finland) available at the PHIV platform (https://phiv.cirad.fr/). The system covers a spectral range from 1000 to 2500 nm, with a spectral resolution of 10 nm and a spatial resolution of 500 µm. Prior to each acquisition session, the system was calibrated using dark-current and white-reference measurements following the manufacturer's guidelines. Imaging was conducted at a constant scanning speed and a fixed sensor-sample distance to ensure consistency among samples. Each radial block was scanned across its entire exposed transverse surface, producing hyperspectral image cubes with high spatial and spectral resolution. Data acquisition was performed using LumoScanner software (Specim, Finland). Data quality control was conducted using Insight software (version 1.3.2.72, Specim, Spectral Imaging Ltd, Finland), and spectral data were subsequently exported for further processing and analysis using MATLAB.

This sampling design aimed to capture radial wood spectral variation by integrating both heartwood and sapwood tissues within each sample. However, heartwood and sapwood were not systematically distinguished during spectral acquisition. Given the known sensitivity of OH-related absorption bands to moisture content and tissue composition, potential effects of tissue heterogeneity on spectral signatures cannot be fully excluded.

Although the number of trees per site was limited, the exhaustive scanning of each sample generated a large volume of hyperspectral data, substantially increasing the effective number of observations and providing a robust representation of within-sample variability [23, 47]. To reduce potential ontogenetic effects, trees were selected within a comparable diameter (DBH ≥ 90 cm) class. Nevertheless, the large size of the hyperspectral datasets imposed practical constraints on data storage, processing, and multivariate analysis. The adopted sampling strategy therefore represents a methodological compromise that prioritizes spatial depth and data quality per tree while maintaining analytical feasibility.

Spectral preprocessing

Each hyperspectral imaging comprises 100.000 of 500.000 of pixel spectra. To reduce computational burden while maintaining representative variability, we selected one spectrum every 100 pixels from the raw image using a regular sampling interval. This yielded approximately 1000 to 5000 representative spectra per individual sample, depending on image size.

After testing several preprocessing methods (including combinations of Standard Normal Variate, polynomial detrending, and Savitzky–Golay derivatives filter with varying window sizes), we selected the Savitzky–Golay first-derivative filter (23-point window and 2 polynomial order as it enhanced spectral features while minimizing noise). The spectra were further restricted to the wavelength range of 1050–2450 nm, corresponding to the noise-free region of the SWIR spectrum. Preprocessing and analysis were performed with R v4.4.0 [48] and RStudio [49] using the *rchemo* package [50], with support from *readxl*, *ggplot2*, and *rgl* for data handling and visualization.

Genetic diversity and population structure analyses

Rare alleles were identified within each population by calculating allele frequencies across loci. Alleles with a frequency lower than 5% were classified as rare [51]. For each site, the number of rare alleles, the total number of alleles, and the proportion of rare alleles relative to the total were computed. Statistical comparisons among populations were performed using a Chi-square test of independence to evaluate differences in rare allele frequencies.

The effective number of alleles per locus (N_A [52]), the rarefied allelic richness for a subsample size (A_R), observed (H_O) and expected (H_E) heterozygosity (gene diversity [53]), Weir and Cockerham's fixation index (F_{IS}) and the deviation from Hardy–Weinberg equilibrium (HWE) were calculated with SPAGeDi 1.5d [54]. Because F can be biased by the presence of null alleles, we tested for their presence and estimated Wright's inbreeding coefficient Fc using a method accounting for null alleles and implemented in INEst 2.3 [55]. Differences in genetic diversity parameters between populations were tested using an ANOVA procedure in R. The parameters were compared among populations, accounting for the locus effect.

To further investigate patterns of spatial genetic structure, principal components analyses (PCA) were performed using the adegenet package [27] in R 4.4.0 software. The PCA scores were used to calculate the genetic distance between individuals. To explore the spatial organization of genetic variation, we used SPAGeDi 1.5d to characterize the spatial genetic structure (SGS) by computing pairwise kinship coefficient (F_{ij}) between individuals using the estimator of Loiselle et al. [56]. F_{ij} values were averaged over a set of non-overlapping

spatial distance intervals, and regressed on the logarithm of the spatial distance between individuals.

Analysis of wood spectral variation

Intraspecific wood spectral variation was explored using principal component analysis (PCA) applied to mean spectral signatures. Unlike classical PCA, which decomposes the covariance matrix, PCA on singular value decomposition (SVD) decomposes the centered data matrix. PCA is particularly suited for datasets in which there is collinearity and when the number of spectral variables far exceeds the number of observations, a common case in hyperspectral imaging, and offers improved numerical stability. Before applying PCA, we reduced the data to a representative spectrum per individual by retaining only one spectrum per sample (following the 1000-pixel sampling). This approach ensured a balanced representation across individuals and avoided over-representation of samples with larger surface areas. The first 10 principal components (PCs), explaining the majority of spectral variance, were retained. These PCs were used as input for calculating Mahalanobis distances between individual spectra. The score distance (SD) and orthogonal distance (OD) were also computed for each sample to characterize its position in and residual from the PCA model, respectively. Together, these distances allow detection of sample outliers and better understanding of intraspecific variability.

To move from pixel-level to individual-level analysis, all spectra associated with a given sample (i.e., a tree) were averaged. This step yielded a mean spectrum per individual, representative of the wood's spectral signature while smoothing pixel-scale heterogeneity. The resulting matrix of individual mean spectra was used for subsequent multivariate analyses, such as dimensionality reduction and inter-individual distance computation.

Cross-validation model (PLS-DA)

Partial least squares discriminant analysis (PLS-DA) is a supervised multivariate method that extends partial least squares (PLS) regression to categorical response variables [57]. Instead of predicting continuous values, PLS-DA models the relationship between spectral data (X) and predefined classes (Y), here corresponding to the different sites of origin of *E. cylindricum*. For each sample, a binary (dummy) response matrix is created where each column represents a class, with values of 1 if the sample belongs to that class and 0 otherwise. The algorithm extracts latent variables (LVs) also called discriminant variables (VD) which are orthogonal components maximizing the covariance between the spectral predictors and the class membership. These components summarize the most discriminative spectral information separating the groups.

For model calibration, the spectral dataset was divided into two subsets: (1) a calibration set, used to build and internally cross-validate the model, and (2) an independent validation (test) set, used for external evaluation of predictive performance. The same segmentation indices were kept constant across replicates to ensure reproducibility. During calibration, a repeat twofold cross-validation ($k=2$, repeated 20 times) was applied to the calibration set to optimize model complexity. At each iteration, the number of latent variables was varied (from 1 to 30), and model error was computed using the rchemo::err function. The optimal number of discriminant variables (opt = 13) corresponded to the minimum mean classification error across folds, as visualized on the cross-validation curve (error vs. number of LV).

The final PLS-DA model was then rebuilt using the optimal preprocessing (Savitzky–Golay first derivative, polynomial order = 2, window = 11) and the selected number of LVs (13). Model validation was performed using the independent test set composed of spectra from *E. cylindricum* cubes, while calibration was done on diametral bar spectra. The predicted classes were compared to the actual one, and a confusion matrix was generated to quantify classification accuracy for each site. From this matrix, global accuracy, sensitivity, and specificity were computed to evaluate the model's discriminatory power. To account for unbalanced sample sizes between sites, weighted PLS-DA models were also tested using class-proportional weights.

Relationship between genetic and spectral variation

To test for concordance between neutral genetic structure and wood spectral variation, Mantel tests were conducted between genetic distance matrices and spectral Mahalanobis distance matrices [58]. Statistical significance was evaluated using permutation procedures. All analyses were performed using reproducible workflows, allowing full replication of the analytical pipeline given access to the raw genetic and hyperspectral data.

Results

Genetic diversity of *Entandrophragma cylindricum* populations

All analyzed loci were polymorphic across the three studied populations (Kabo, Loundoungou, and Mimbeli), with the number of alleles per locus ranging from 7 to 33, totalling 174 unique alleles. The analysis of private alleles revealed a notable contrast in the distribution of unique genetic variants among sites. A total of 10 private alleles were identified across the nine loci examined, with five in Loundoungou, four in Mimbeli, and one in Kabo

Table 2 Rare alleles, total alleles and rare rate between studied populations

Population	Rare alleles	Total alleles	Rare rate
Kabo	1	63	1.59
Loundoungou	5	70	7.14
Mimbeli	4	79	5.06

(Table 2). Although Loundoungou has the highest proportion (~7%) and Kabo the lowest (~1.6%), these differences are not significant between sites ($\chi^2 = 2.31$; ddl = 2; $p = 0.315$).

The summary of genetic diversity parameters across the 177 sampled individuals is presented in Table 3. The effective number of alleles (N_A), observed heterozygosity (H_O), allelic richness (A_R), expected gene diversity (H_E), and inbreeding coefficient (F) did not show significant differences among populations (ANOVA, $p > 0.05$), suggesting genetic homogeneity across sites. A consistent heterozygote deficiency ($H_E > H_O$) was observed in all three populations, resulting in positive inbreeding coefficient values (F). Nonetheless, after correcting for genotyping errors and the potential presence of null alleles using the INEst model, the adjusted inbreeding coefficients ($F_{(null)}$) were close to zero, with 95% credible intervals overlapping with zero. This indicates no evidence of true inbreeding once methodological artifacts were accounted for.

Fine-scale SGS and inference of gene dispersal distances

The SGS analysis revealed a gradual decline in genetic relatedness with increasing geographic distance between individuals. Mean kinship values decreased from 0.011 at short distances (≤400 m) to -0.003 at distances greater than 20 km (Fig. 1). This pattern was further supported by a significant negative regression of kinship on the natural logarithm of geographic distance (slope $b = -0.0015$; $r^2 = 0.0011$), indicating a weak but significant spatial genetic structure consistent with an isolation-by-distance (IBD) model.

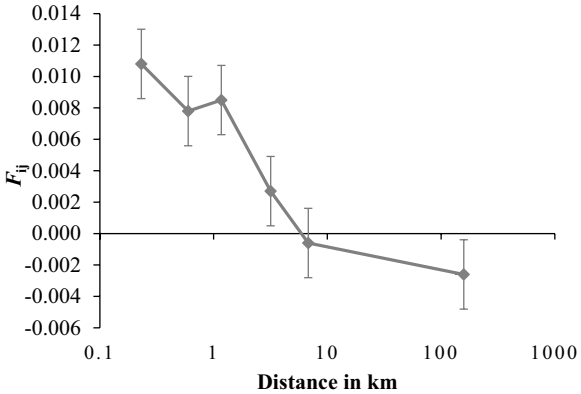


Fig. 1 Spatial genetic structure of *Entandrophragma cylindricum*. Average pairwise kinship coefficients (F_{ij}) plotted against geographic distance classes (log-transformed), illustrating the decay of genetic relatedness with increasing spatial distance

Table 4 Pairwise F_{ST} value (Nei, 1973) quantifying genetic differentiation between *E. cylindricum* populations

Sites		F_{ST}	*P*-test
Kabo	Loundoungou	0.007 ± 0.002	0.011
Kabo	Mimbeli	0.005 ± 0.003	0.101
Loundoungou	Mimbeli	0.006 ± 0.002	0.025

Although *E. cylindricum* populations in northern Congo show a high level of genetic diversity within populations, genetic differentiation between sites is limited (Table 4). This spatial genetic continuity provides a foundation for further examining the link between genetic variation and wood quality traits.

Spectral characterization of wood quality across ecological sites

Hyperspectral imaging (HSI) of *E. cylindricum* wood samples revealed subtle but consistent differences in NIR spectral signatures across the three study sites. The raw spectra recorded between 1000 and 2500 nm exhibited

Table 3 Parameters of genetic diversity and inbreeding statistics of the different population

Population	[1]N	[2]N_A	[3]A_R	[4]H_E	[5]H_O	[6]F_{IS}	[7]$F_{(null)}$	[8]HPDi(95%)	[9]HPDh(95%)
Kabo	51	14.75	13.92	0.806	0.675	0.165	0.015	0.000	0.046
Loundoungou	80	16.63	14.55	0.829	0.666	0.198	0.035	0.000	0.085
Mimbeli	46	16.38	15.97	0.800	0.642	0.200	0.015	0.000	0.049
All populations	177	21.75	15.76	0.819	0.663	0.192	0.021		

[1] Sample size, [2] effective number of alleles, [3] allelic richness, [4] expected heterozygosity (gene diversity corrected for sample size), [5]observed heterozygosity, [6] fixation index estimating the inbreeding coefficient without accounting for null allele. [7] Inbreeding coefficient corrected for null alleles under a population inbreeding model, [8] 95% highest posterior density interval, [9] 95% highest posterior density interval (for heterozygosity)

two distinct absorption peaks at approximately 1400 and 1900 nm (Fig. 2a), corresponding to overtone bands of O–H bonds [19, 59, 60]. These peaks became more pronounced after preprocessing (SAVGOL1, Fig. 2b), and the averaged spectra by site revealed distinguishable patterns, particularly between Loundoungou (blue) and Mimbeli (red), while Kabo (green) showed more overlap, suggesting a more heterogeneous profile at that site.

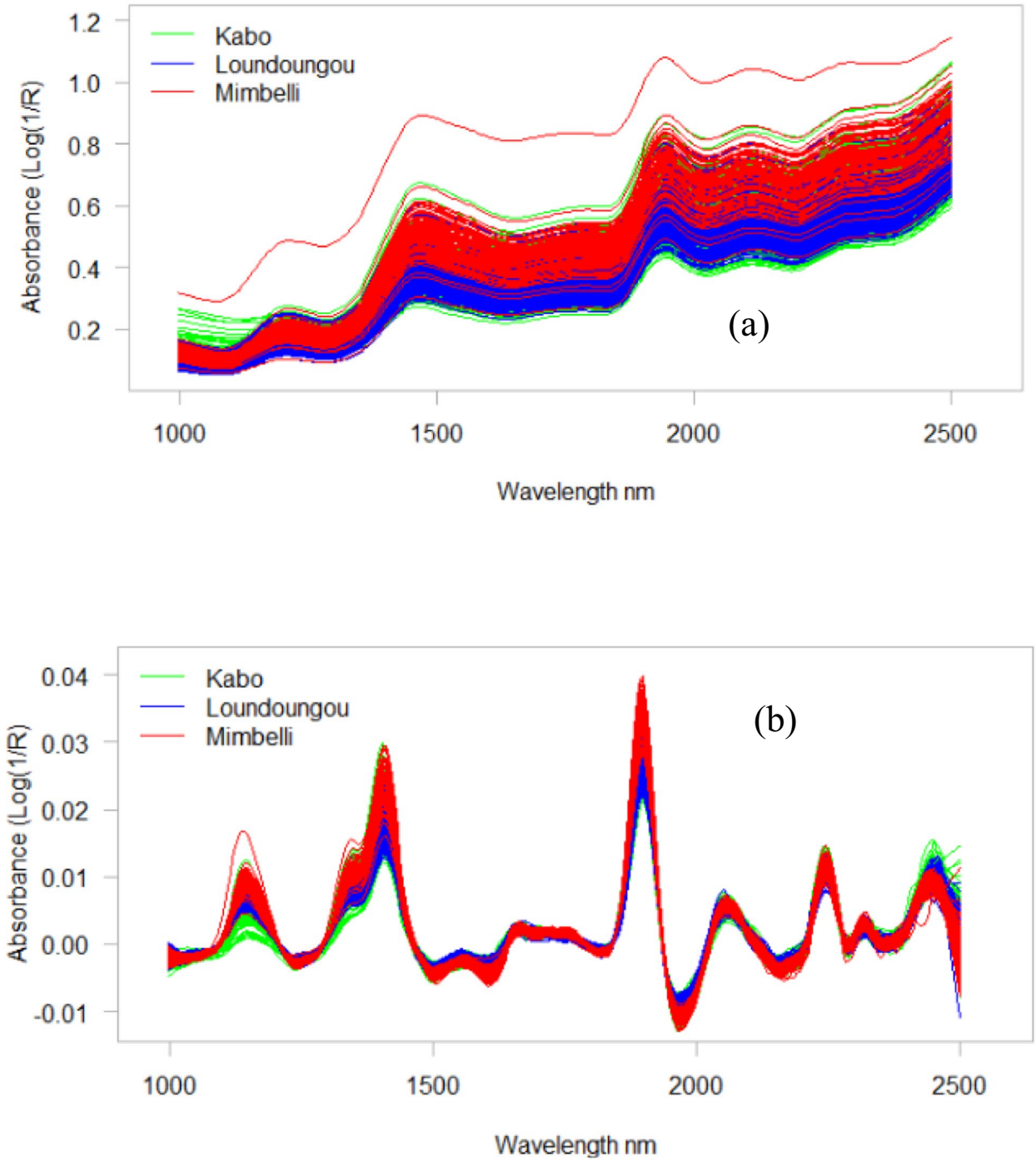


Fig. 2 Near-infrared (NIR) spectral signatures of *Entandrophragma cylindricum* wood from three sampling sites. **a** Raw NIR spectra; **b** NIR spectra after Savitzky–Golay first derivative preprocessing (SAVGOL1; window size = 11, polynomial order = 2). Colors indicate sampling sites: Kabo (green), Loundoungou (blue), and Mimbeli (red)

Principal component analysis (PCA) on raw spectral data

The PCA score plot (Fig. 3a) computed on SAVGOL1 spectra revealed moderate but discernible clustering by site in the *PC1–PC2* space, which together accounted for 90.6% of the total spectral variance. *PC1* slightly separated individuals from Loundoungou and Mimbelli, while *PC2* more clearly distinguished samples from Kabo. Notably, individuals from Kabo exhibited the widest dispersion in the PCA space, suggesting higher within-population variability in wood spectral properties. In contrast, samples from Loundoungou and Mimbelli formed more compact and homogeneous clusters (Fig. 3a). The PCA loading plots (Fig. 3b) indicated that *PC1* was primarily driven by negative peaks near 1400 nm and 1900 nm. *PC2* was influenced by positive contributions around 1150 nm and 1400 nm, along with a strong negative contribution at 1900 nm.

Spectral outlier detection and individual-level averaging

Spectral aggregation at the individual level is observed in Fig. 4. Scatter plot of Mahalanobis distance (x-axis) versus orthogonal distance (y-axis) computed for individual pixel-level spectra after dimensionality reduction via PCA computed on around 36,000 spectra. Each point corresponds to a single spectrum, colored by site of origin: Kabo (green), Loundoungou (blue), and Mimbelli (red). The dashed horizontal and vertical lines represent the thresholds used for outlier detection, beyond which spectra were flagged as statistical outliers. A total of 14 spectra were identified as outliers, most likely due to instrumental artifacts or contamination, and excluded from subsequent analyses.

Discriminant analysis by PLS-DA

Based on global accuracy determined by cross-validation, the optimal pre-treatment selected is SAVGOL1 with 23 smoothing points and polynom 2. The number of discriminant variables determined by cross-validation was 11, with a global accuracy of 69.1% based on 14,279, 14,384 and 7309 spectra, respectively, for Kabo, Loundoungou, and Mimbeli. The global accuracy of PLS-DA model validation is 67%. Samples from Mimbeli exhibited the highest accuracy (99.9%), followed by those from Kabo (70.2%) and Loundoungou (48.6%; Table 5). The three sites and individuals can be better discriminated from their spectra data (Fig. 5a). Averaged over individuals, the sites stand out more clearly (Fig. 5b).

Relationship between genetic variation and wood spectral profiles

To investigate whether the intraspecific variation in wood quality assessed via NIR hyperspectral data was genetically driven, we conducted a Mantel test comparing pairwise genetic distances (derived from PCA on microsatellite data) and Mahalanobis distances calculated from spectral data. The analysis revealed a non-significant correlation between genetic and spectral distances ($r = -0.0889$, $p = 0.803$), indicating no detectable genetic signal underlying the spectral variation in wood quality observed across the sampled individuals of *E. cylindricum*.

Discussion

Genetic diversity and population structure of *Entandrophragma cylindricum*

Our first objective was to assess the genetic diversity and population structure of *Entandrophragma cylindricum* across the three sampling sites using nuclear microsatellite markers, in order to determine whether genetic differentiation could contribute to spatial patterns of wood spectral variation. The results revealed relatively high levels of genetic diversity within populations, consistent with previous studies on *E. cylindricum* and other Central African timber species such as *Milicia excelsa* and *Distemonanthus benthamianus* [16, 17, 32, 39, 61].

Genetic differentiation among sites was weak, and inbreeding coefficients were low ($F_{null} \approx 0.02$), indicating predominantly outcrossed mating systems and substantial gene flow. Spatial genetic analyses revealed a weak but significant isolation-by-distance (IBD) pattern, a common feature in tropical tree species with extensive dispersal capacities [16, 62, 63]. These results are consistent with the reproductive ecology of *E. cylindricum*, which combines insect-mediated pollination primarily by bees with efficient wind-dispersed seeds. Mean dispersal distances of at least 500 m for pollen and 400 m for seeds have been reported [38], exceeding those of many other African wind-dispersed tree species [16, 64–66].

Such dispersal efficiency promotes gene flow at the landscape scale and can maintain genetic connectivity even in forests affected by selective logging [67–69]. In this context, the weak genetic structuring observed among the study sites suggests that spatial differentiation in wood spectral traits, if present, is unlikely to be primarily driven by population-level genetic divergence.

Interpretation of NIR spectral variability in relation to wood chemistry and maturation

Near-infrared (NIR) spectral signatures are dominated by absorption features associated with O–H, C–H, and N–H molecular bonds, which are abundant in wood cell wall polymers and water [24, 70–73]. In our study, the most influential spectral regions were centered around 1400 nm and 1900 nm, wavelengths commonly associated with O–H stretching overtones [24, 73]. These features are known to integrate signals related to moisture dynamics,

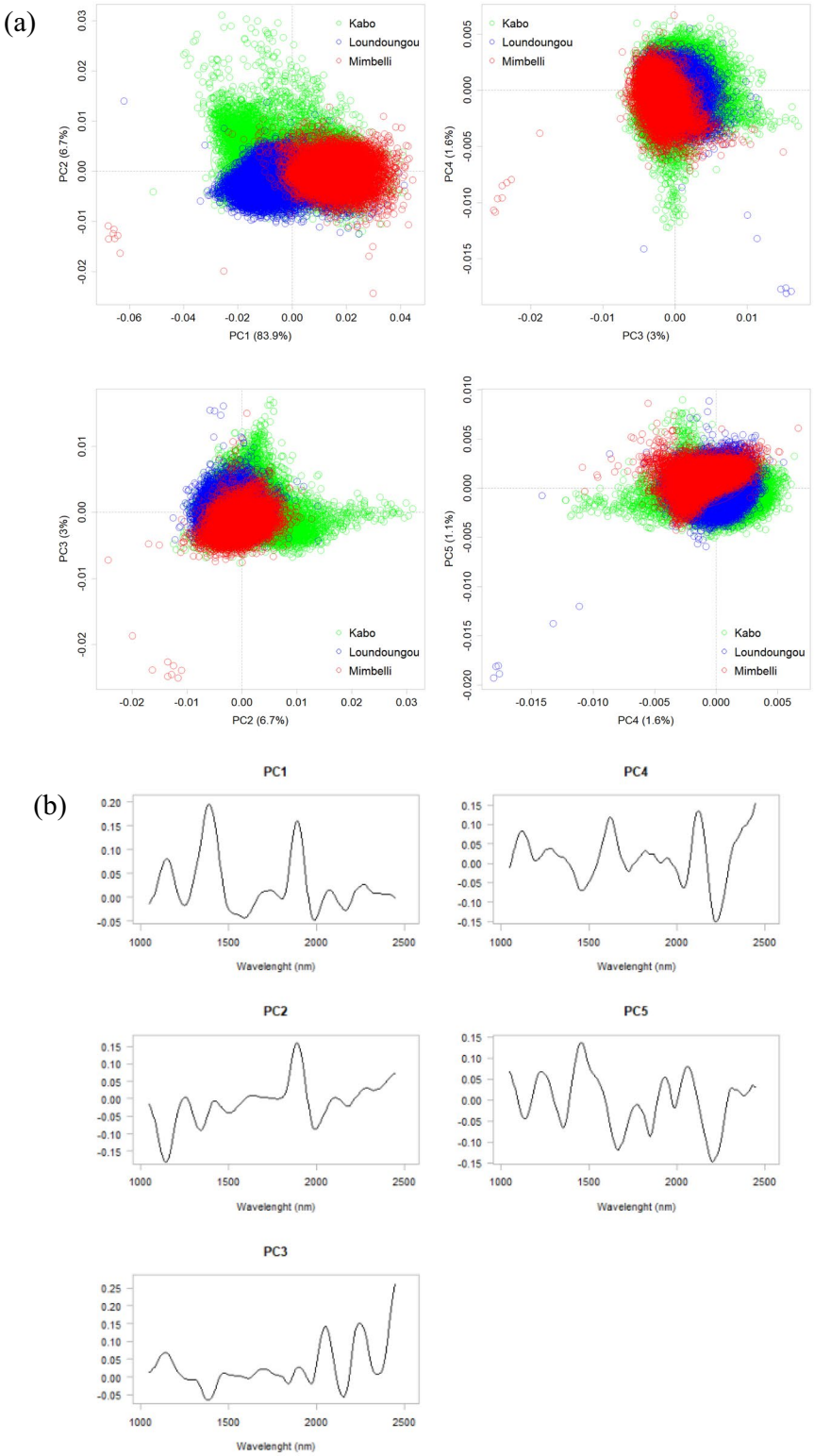


Fig. 3 Principal component analysis (PCA) of wood NIR hyperspectral data of *Entandrophragma cylindricum*. **a** PCA score plot based on preprocessed spectra (SAVGOL1), showing the distribution of individual samples according to sampling site; **b** Loading plots of the first five principal components, indicating the spectral regions contributing most to variance

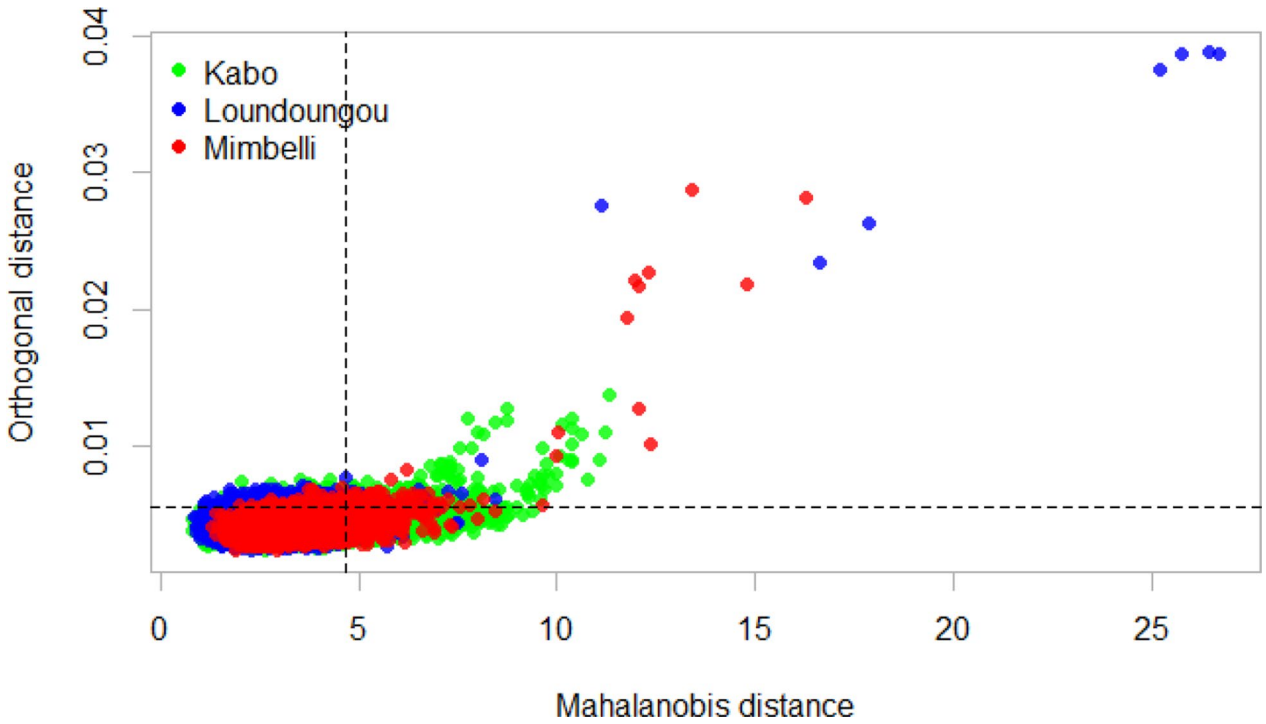


Fig. 4 Identification of spectral outliers based on multivariate distance metrics. Scatter plot of Mahalanobis distance versus orthogonal distance calculated from individual-averaged spectra ($n = 21$) in the reduced PCA space (10 components) after spectral preprocessing. Points are colored by site: Kabo (black), Loundoungou (blue), and Mimbeli (red)

hydroxyl groups in cellulose and hemicelluloses, and hydrogen bonding within the cell wall matrix [24, 26, 74]. Although no direct chemical or physical measurements were performed, the observed variation in these spectral regions likely reflects differences in the hydration state of wood tissues and in the relative organization of cell wall components. Such spectral variation may also be linked to wood maturation processes, including changes in cell wall thickness, lignification, and extractive accumulation, which modify the availability and binding state of hydroxyl groups [25, 74]. Importantly, these processes occur along radial gradients from pith to bark and can vary among individuals depending on growth conditions and tree age [71, 75, 76]. Consequently, the NIR spectral variability observed among individuals should be interpreted as an integrative signal of unmeasured physicochemical variation associated with wood formation and maturation, rather than as a direct indicator of specific wood technological properties.

Table 5 Confusion matrix of PLS-DA model validation

	Real value		
	Kabo	**Loundoungou**	**Mimbeli**
Predicted value			
Kabo	9105	2696	2
Loundoungou	2419	3322	1
Mimbeli	1440	812	2554
N sample	12,964	6830	2557
Global accuracy		67.0	
Sensitivity	70.2	48.6	99.9
Specificity	71.3	84.4	88.6

Interpreting site-associated spectral variability: cautious perspectives

Although environmental variables were not directly measured in this study, the site-associated patterns observed in wood spectral traits suggest that local growing contexts and stand histories may contribute to the observed spectral differentiation. The higher spectral dispersion observed at Kabo may reflect greater heterogeneity in individual growth conditions or disturbance legacies, whereas the more clustered spectral profiles at Loundoungou and Mimbeli could indicate more homogeneous site contexts at the scale sampled. Importantly, these interpretations remain hypothetical and should not be construed as evidence of specific environmental drivers. Rather, the results highlight that wood spectral variation captures integrated phenotypic signals shaped by multiple interacting processes, including

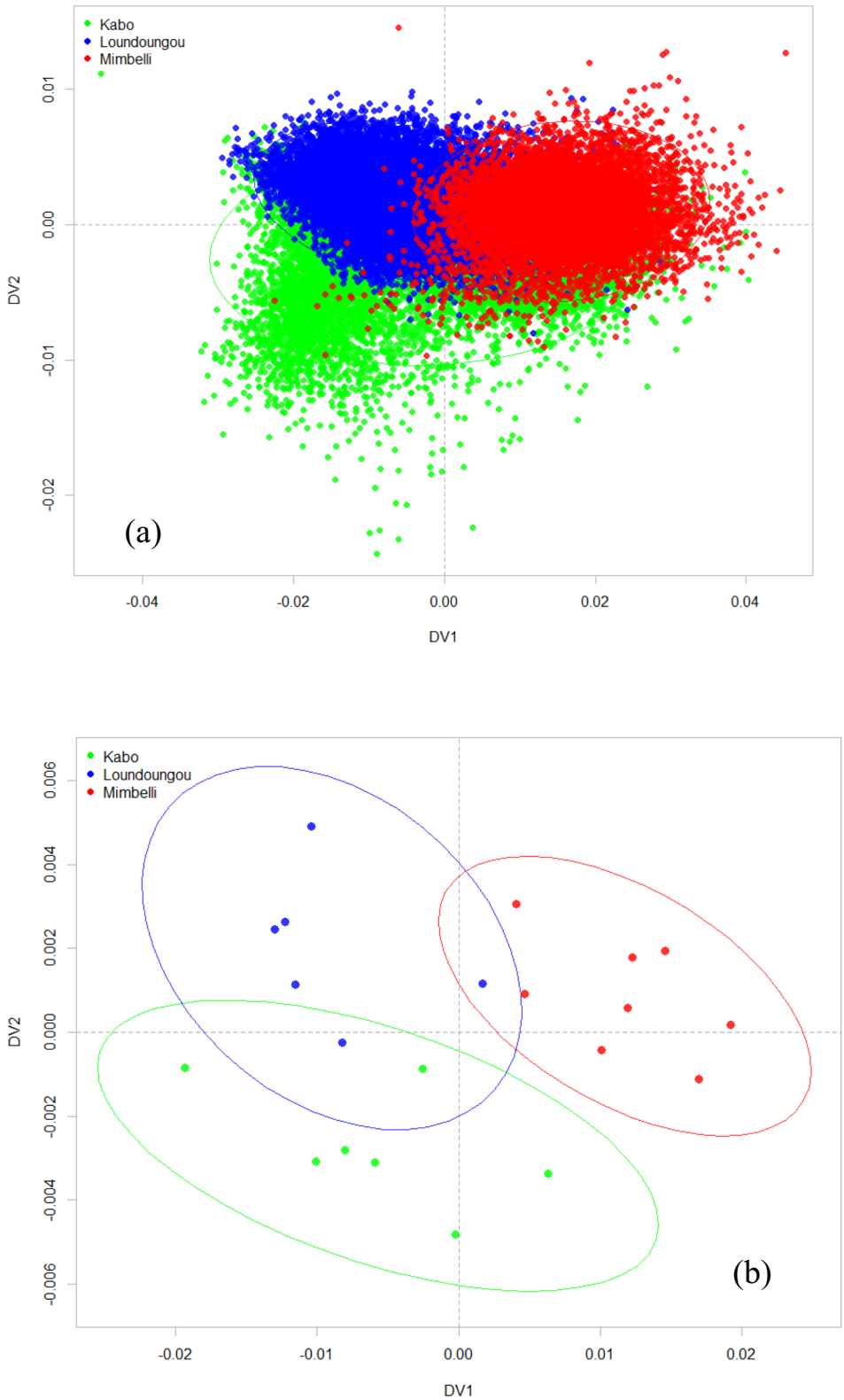


Fig. 5 Partial least squares discriminant analysis (PLS-DA) of wood NIR spectral data. **a** Score plot of the first two discriminant components based on calibration spectra after SAVGOL1 preprocessing (11 latent variables); **b** Score plot of the first two discriminant components based on individual-averaged calibration spectra

growth conditions, tree ontogeny, and possibly microscale environmental variability. Disentangling the relative contributions of these factors will require future studies explicitly combining NIR spectroscopy with direct measurements of environmental variables and detailed wood anatomical analyses. Given the limited number of trees analyzed spectrally, results should be interpreted as exploratory and cannot fully capture site-level variability.

Stand disturbance, logging history, and spectral heterogeneity

Selective logging alters forest structure by creating canopy openings, modifying light regimes, and changing competitive interactions among trees. These disturbances can have lasting effects on tree growth trajectories and wood formation processes. Increased light availability following logging can enhance radial growth rates, potentially affecting the balance between earlywood and latewood production, cell size, and wall thickness [3, 77, 78]. In the Kabo site, where logging has occurred since the mid-twentieth century, repeated disturbance events may have generated a mosaic of growth conditions within the stand. Such heterogeneity can lead to asynchronous growth patterns and variable wood maturation processes among individuals, contributing to the higher spectral dispersion observed in this population. In contrast, less intensively logged sites may exhibit more synchronized growth environments, resulting in more homogeneous spectral signatures. These findings suggest that NIR hyperspectral imaging is sensitive to stand-level disturbance history, capturing integrated signals of wood formation responses to long-term management practices.

Lack of genetic control on spectral variability

The absence of a significant correlation between genetic distances and spectral dissimilarities indicates that wood spectral variability is not structured by genetic relatedness at the spatial scale considered. This result aligns with expectations for a species characterized by high gene flow and weak population differentiation. While genetic factors undoubtedly influence wood formation at the individual level, their effects may be obscured by strong environmental modulation under heterogeneous field conditions. Our results therefore highlight the predominance of environmental filtering and phenotypic plasticity in shaping wood spectral signatures in *E. cylindricum*. Detecting a genetic basis for spectral traits would likely require approaches explicitly designed to partition genetic and environmental variance, such as common garden experiments, quantitative genetic analyses, or genome-wide association studies.

Implications for sustainable forest management

Across the three study sites, genetic diversity indices (H_E, A_R) were high and relatively homogeneous, suggesting that long-term selective logging has not resulted in a marked erosion of overall genetic diversity. However, the uneven distribution of private alleles among populations provides a more nuanced perspective. Loundoungou harbored the highest number of private alleles, followed by Mimbeli, whereas Kabo subject to the longest logging history retained only one private allele. This pattern suggests that prolonged exploitation may preferentially erode rare alleles, thereby reducing local genetic uniqueness without strongly affecting global diversity indices. Such findings are consistent with previous studies indicating that selective logging often leads to subtle genetic changes, primarily through the loss of rare alleles rather than a collapse of overall diversity [79–81]. For *E. cylindricum*, regeneration dynamics further complicate this picture. Although germination rates can be high [36, 82], seedling survival under closed canopies is low [83, 84], and correlated paternity within fruits may limit effective genetic mixing at early life stages [38]. From a management perspective, these results indicate that genetic indicators alone are insufficient to guide fine-scale operational decisions. The substantial intraspecific variation observed in wood spectral traits largely decoupled from neutral genetic structure underscores the complexity of the links between genotype, phenotype, and site context. Bridging this gap will require integrative approaches combining population genetics, wood anatomy, direct measurements of environmental variables, and forest management history to better understand how exploitation practices influence both genetic resources and phenotypic variability in tropical timber species. Because environmental variables and tree-level characteristics were not directly measured, the interpretation of site-related spectral differences in terms of environmental influence remains inferential and should be viewed as a working hypothesis rather than a demonstrated causal relationship.

Conclusion

This study provides an integrated assessment of population genetic structure and wood spectral variation in *Entandrophragma cylindricum* across the three Forest Management Units in northern Republic of the Congo. Despite prolonged selective logging in two of the sites, populations exhibited relatively high levels of genetic diversity, weak spatial genetic structuring, and low inbreeding coefficients, indicating effective gene flow likely maintained by the species' dispersal biology and landscape connectivity. In parallel, near-infrared

(NIR) spectroscopy revealed substantial intraspecific variation in wood spectral traits among individuals and sampling sites. This spectral variation showed no significant association with neutral genetic distances, suggesting that wood spectral patterns are not strongly genetically structured at the spatial scale considered. Instead, the observed site-related differentiation likely reflects integrated phenotypic responses influenced by local growing contexts, individual growth histories, and stand-level heterogeneity, although these factors were not directly quantified in this study. Together, these results highlight a decoupling between neutral genetic structure and wood spectral variation in *E. cylindricum*, underscoring the complexity of the relationships between genotype, phenotype, and site context in tropical timber species. While genetic diversity and population structure were robustly characterized, the genetic basis of wood spectral traits was not addressed here. Future studies combining NIR spectroscopy with wood anatomical analyses, direct environmental measurements, and genomic approaches such as genome-wide association studies or quantitative trait locus mapping will be necessary to disentangle the respective contributions of genetic and non-genetic factors. From a forest management perspective, our findings emphasize that maintaining overall genetic diversity may not be sufficient to capture the full extent of intraspecific phenotypic variability. Integrative and long-term approaches that jointly consider genetic resources, phenotypic heterogeneity, and site context will be essential to support the sustainable management and conservation of *E. cylindricum* populations under ongoing anthropogenic pressures and environmental change.

Abbreviations

DBH	Diameter at breast height
DNA	Deoxyribonucleic acid
FMU	Forest Management Unit
HIS	Hyperspectral imaging
HWE	Hardy–Weinberg equilibrium
IBD	Isolation-by-distance
LV	Latent variable
NIR	Near-infrared
OD	Orthogonal distance
PC	Principal component
PCA	Principal component analysis
PCR	Polymerase chain reaction
PLS-DA	Partial least square discriminant analysis
SAVGOL1	Savitzky–Golay filter-1st derivative
SD	Score distance
SGS	Spatial genetic structure
SVD	Single value decomposition
SWIR	Stationary short-wave Infrared
VD	Discriminant variable

Acknowledgements
We are grateful to CIB Olam Agri for granting access to their concessions and for their technical support throughout the study. We would like to thank the field teams, for their dedication and assistance in data collection. Special thanks go to Baptiste Leborgne for his invaluable help in retrieving information on the study sites. We would like to extend our heartfelt thanks to Netih Demba for his work and data production during her Master work. We thank an anonymous reviewer who provided useful comments that helped to improve the manuscript.
During the preparation of this work the author used ChatGPT in order to translate the text. After using this tool/service, the author reviewed and edited the content as needed and takes full responsibility for the content of the published article.

Authors contributions
GBD conceived the ideas. BY, SM, EJ, RN, RM and RD collected the data. PL received the wood samples and performed the pre-treatment prior to spectral image acquisition. BY, GC and OJH analyzed the data (spectral and genetic data, respectively). BY wrote the manuscript. JJL and CD supervised the study. CGM, GC, OJH, PL, GJL, YB, JJL and CD critically revised the manuscript.

Funding
This research was funded by the European Union through the Applied Research in Ecology and Social Science in Support of Sustainable Management of Forest Ecosystems in Central Africa (RESSAC) programme, coordinated by the Centre for International Forestry Research (CIFOR) and the International Centre for Research in Agroforestry (ICRAF).

Data availability
The datasets generated and analyzed during the current study are not publicly available due to data ownership restrictions, but are available from the corresponding author on reasonable request.

Declarations

Competing interests
The authors declare no competing interests.

Author details
[1]Laboratoire de Biodiversité, de Gestion des Ecosystèmes et de L'Environnement (LBGE), Faculté des Sciences et Techniques Brazzaville, Université Marien Ngouabi, Brazzaville, Republic of the Congo. [2]Institut Supérieur des Sciences Géographiques, Environnementales et de L'Aménagement (ISSGEA), Université Denis Sassou Nguesso, Kintélé, Republic of the Congo. [3]Unité mixte de recherche Biomasse–Bois–Energie–Bioproduits (BioWooEB), CIRAD, Montpellier, France. [4]CIRAD-UMR AGAP Institut, Montpellier, France. [5]UMR AGAP Institut, Univ Montpellier, CIRAD, INRAE, Institut Agro, Montpellier, France. [6]ChemHouse Research Group, Montpellier, France. [7]Unité Propre de Recherche Forêts & Sociétés, CIRAD, Montpellier, France. [8]Institut National de Recherche Forestière (IRF), Brazzaville, Republic of the Congo. [9]Evolutionary Biology and Ecology Unit, Faculté Des Sciences, Université Libre de Bruxelles, Brussels, Belgique.

Received: 6 November 2025 Accepted: 20 April 2026
Published online: 13 June 2026

References

1. Eba'a Atyi R, Hiol Hiol F, Lescuyer G et al (2021) Les forêts du bassin du Congo: État des Forêts 2021. CIF, Bogor
2. Gourlet-Fleury S, Mortier F, Fayolle A et al (2013) Tropical forest recovery from logging: a 24 year silvicultural experiment from Central Africa. Philos Trans R Soc B Biol Sci. https://doi.org/10.1098/rstb.2012.0302
3. Karsenty A, Gourlet-Fleury S (2006) Assessing sustainability of logging practices in the Congo Basin's managed forests: the issue of commercial species recovery. Ecol Soc. https://doi.org/10.5751/ES-01668-110126
4. Siefert A, Violle C, Chalmandrier L et al (2015) A global meta-analysis of the relative extent of intraspecific trait variation in plant communities. Ecol Lett 18:1406–1419. https://doi.org/10.1111/ele.12508
5. Whitlock R (2014) Relationships between adaptive and neutral genetic diversity and ecological structure and functioning: a meta-analysis. J Ecol 102:857–872. https://doi.org/10.1111/1365-2745.12240
6. Sungpalee W, Itoh A, Kanzaki M et al (2009) Intra- and interspecific variation in wood density and fine-scale spatial distribution of stand-level

wood density in a northern Thai tropical montane forest. J Trop Ecol 25:359–370. https://doi.org/10.1017/S0266467409006191
7. Maherali H, DeLucia EH (2000) Xylem conductivity and vulnerability to cavitation of ponderosa pine growing in contrasting climates. Tree Physiol 20:859–867. https://doi.org/10.1093/treephys/20.13.859
8. Baribault TW, Kobe RK, Finley AO (2012) Tropical tree growth is correlated with soil phosphorus, potassium, and calcium, though not for legumes. Ecol Monogr 82:189–203. https://doi.org/10.1890/11-1013.1
9. Zobel BJ, van Buijtenen JP (1989) Wood variation. Its causes and control. Springer-Verlag, Heidelberg
10. Bergès L, Nepveu G, Franc A (2008) Effects of ecological factors on radial growth and wood density components of sessile oak (*Quercus petraea* Liebl.) in Northern France. For Ecol Manage 255:567–579. https://doi.org/10.1016/j.foreco.2007.09.027
11. Zhang SB, Slik JWF, Zhang JL, Cao KF (2011) Spatial patterns of wood traits in China are controlled by phylogeny and the environment. Glob Ecol Biogeogr 20:241–250. https://doi.org/10.1111/j.1466-8238.2010.00582.x
12. Zobel B, Talbert J (1984) Applied forest tree improvement. https://archive.org/details/appliedforesttre0000zobe_i6o1
13. Duminil J, Mendene Abessolo DT, Ndiade Bourobou D et al (2016) High selfing rate, limited pollen dispersal and inbreeding depression in the emblematic African rain forest tree *Baillonella toxisperma*—management implications. For Ecol Manage 379:20–29. https://doi.org/10.1016/j.foreco.2016.08.003
14. Ndiade-Bourobou D, Hardy OJ, Favreau B et al (2010) Long-distance seed and pollen dispersal inferred from spatial genetic structure in the very low-density rainforest tree, *Baillonella toxisperma* Pierre, in Central Africa. Mol Ecol 19:4949–4962. https://doi.org/10.1111/j.1365-294X.2010.04864.x
15. Yogom BT, Avana-Tientcheu ML, Monthé FK et al (2023) Genetic diversity and structure in wild and cultivated populations of an emblematic African tree species, *Garcinia kola* (Clusiaceae). Tree Genet Genomes. https://doi.org/10.1007/s11295-023-01614-w
16. Angbonda DMA, Monthe FK, Bourland N et al (2021) Seed and pollen dispersal and fine-scale spatial genetic structure of a threatened tree species: *Pericopsis elata* (HARMS) Meeuwen (Fabaceae). Tree Genet Genomes. https://doi.org/10.1007/s11295-021-01509-8
17. Lourmas M, Kjellberg F, Dessard H et al (2007) Reduced density due to logging and its consequences on mating system and pollen flow in the African mahogany *Entandrophragma cylindricum*. Heredity (Edinb) 99:151–160. https://doi.org/10.1038/sj.hdy.6800976
18. Bizoux JP, Daïnou K, Bourland N et al (2009) Spatial genetic structure in *Milicia excelsa* (Moraceae) indicates extensive gene dispersal in a low-density wind-pollinated tropical tree. Mol Ecol 18:4398–4408. https://doi.org/10.1111/j.1365-294X.2009.04365.x
19. Chambi-Legoas R, Tomazello-Filho M, Vidal C, Chaix G (2023) Wood density prediction using near-infrared hyperspectral imaging for early selection of *Eucalyptus grandis* trees. Trees 37:981–991. https://doi.org/10.1007/s00468-023-02397-2
20. Richard N, Deborah H, Ben-Ahmed O (2018) Hyperspectral image processing: full-band, dimensionality reduction, or band selection approach? https://www.researchgate.net/publication/326293927
21. Baston G, Qian J (2015) Review of advancements in hyperspectral imaging for low-light environments. J latex Cl files 14:1–4 https://osf.io/preprints/osf/jv3p7_v1
22. Viet DD, Ma T, Inagaki T et al (2021) Near-infrared spectroscopy and hyperspectral imaging can aid in the prediction and mapping of polyploid acacia hybrid wood properties in tree improvement programs. Holzforschung 75:1067–1080. https://doi.org/10.1515/hf-2021-0024
23. Schimleck L, Ma T, Inagaki T, Tsuchikawa S (2023) Review of near infrared hyperspectral imaging applications related to wood and wood products. Appl Spectrosc Rev 58:585–609. https://doi.org/10.1080/05704928.2022.2098759
24. Schwanninger M, Rodrigues JC, Fackler K (2011) A review of band assignments in near infrared spectra of wood and wood components. J Near Infrared Spectrosc 19:287–308. https://doi.org/10.1255/jnirs.955
25. Kelley SS, Rials TG, Snell R et al (2004) Use of near infrared spectroscopy to measure the chemical and mechanical properties of solid wood. Wood Sci Technol 38:257–276. https://doi.org/10.1007/s00226-003-0213-5
26. Tsuchikawa S, Siesler HW (2003) Near infrared spectroscopic monitoring of the diffusion process of deuterium-labeled molecules in wood. Soc Appl Spectrosc 57:123–137. https://doi.org/10.1002/9780470999714.ch10
27. Jombart T, Devillard S, Dufour AB, Pontier D (2008) Revealing cryptic spatial patterns in genetic variability by a new multivariate method. Heredity (Edinb) 101:92–103. https://doi.org/10.1038/hdy.2008.34
28. Ouinsavi C, Sokpon N, Khasa DP (2009) Genetic diversity and population structure of a threatened African tree species, *Milicia excelsa*, using nuclear microsatellites DNA markers. Int J For Res 2009:1–8. https://doi.org/10.1155/2009/210179
29. Aguillon S, Din Dipita A, Lecompte E et al (2020) Development and characterization of 20 polymorphic microsatellite markers for the white-bellied pangolin *Phataginus tricuspis* (Mammalia, Pholidota). Mol Biol Rep. https://doi.org/10.1007/s11033-020-05511-6
30. Born C, Hardy OJ, Chevallier MH et al (2008) Small-scale spatial genetic structure in the Central African rainforest tree species *Aucoumea klaineana*: a stepwise approach to infer the impact of limited gene dispersal, population history and habitat fragmentation. Mol Ecol 17:2041–2050. https://doi.org/10.1111/j.1365-294X.2007.03685.x
31. Hardy OJ, Maggia L, Bandou E et al (2006) Fine-scale genetic structure and gene dispersal inferences in 10 Neotropical tree species. Mol Ecol 15:559–571. https://doi.org/10.1111/j.1365-294X.2005.02785.x
32. Monthe FK, Duminil J, Tosso F et al (2017) Characterization of microsatellite markers in two exploited African trees, *Entandrophragma candollei* and *E. utile* (Meliaceae). Appl Plant Sci 5:1600130. https://doi.org/10.3732/apps.1600130
33. CIRAD-Tropix (2023) Sapelli Propriétés physiques et mécaniques Séchage. https://tropix.cirad.fr/FichiersComplementaires/FR/Afrique/SAPELLI%202024.pdf
34. Tabi Eckebil PP, Verheggen F, Doucet JL et al (2017) *Entandrophragma cylindricum* (Sprague) sprague (Meliaceae), une espèce ligneuse concurrentielle en afrique centrale (synthèse bibliographique). Biotechnol Agron Soc Environ 21:80–97. https://doi.org/10.25518/1780-4507.13508
35. Hall JS, Mckenna JJ, Ashton PMS, Timothy GG (2004) Habitat characterizations underestimate the role of adaphic factors controlling the distribution of *Entandrophragma*. Ecology 85:2171–2183
36. Palla F, Louppe D, Forni É (2002) Sapelli. CIRAD-Forêt, France
37. Vivien J, Faure JJ (1985) Arbres des forêts denses d'Afrique centrale. https://www.abebooks.fr/Arbres-forets-denses-dAfrique-Centrale-648%20J.Vivien/18140999501/bd
38. Monthe FK, Hardy OJ, Doucet JL et al (2017) Extensive seed and pollen dispersal and assortative mating in the rain forest tree *Entandrophragma cylindricum* (Meliaceae) inferred from indirect and direct analyses. Mol Ecol 26:5279–5291. https://doi.org/10.1111/mec.14241
39. Garcia F, Noyer JL, Risterucci AM, Chevallier MH (2004) Genotyping of mature trees of *Entandrophragma cylindricum* with microsatellites. J Hered 95:454–457. https://doi.org/10.1093/jhered/esh071
40. Détienne P, Mariaux A (1977) Nature et périodicité des cernes dans les bois rouges de Méliacées Africaines. Bois For Trop 175:52–61
41. Wright SJ (1996) Phenological responses to seasonality in tropical forest plants. Springer, Boston, MA
42. Dupuy B (1998) Bases pour une sylviculture en forêt dense tropicale humide africaine. Ser FORAFRI, CIRAD 387. https://forafri.cirad.fr/ressources/forafri/09.pdf
43. Vennetier P (1963) Les hommes et leurs activités dans le Nord du Congo-Brazzaville. Université de Bordeaux. Paris, Cahiers C.R.S.T.O.M., 296 p. https://www.persee.fr/doc/rga_0035-1121_1967_num_55_1_3307_t1_0228_0000_1
44. Schwartz D, Namri M (2002) Mapping the total organic carbon in the soils of the Congo. Glob Planet Change 33:77–93. https://doi.org/10.1016/S0921-8181(02)00063-2
45. Doyle JJ, Doyle JL (1987) CTAB DNA extraction in plants. Phytochem Bull 19:11–15
46. Micheneau C, Dauby G, Bourland N et al (2011) Development and characterization of microsatellite loci in *Pericopsis elata* (Fabaceae) using a cost-efficient approach. Am J Bot. https://doi.org/10.3732/ajb.1100070
47. Ma T, Schimleck L, Dahlen J et al (2022) Comparative performance of NIR-hyperspectral imaging systems. Foundations Basel 2:523–540. https://doi.org/10.3390/foundations2030035
48. Team. RC (2024) R: A language and environment for statistical computing (4.4.0). R Found Stat Comput. https://www.r-project.org/

49. Posit team (2025) RStudio: integrated development environment for R. Posit Software, PBC, Boston, MA
50. Brandolini-Bunlon M, Jaillais B, Roger JM, Lesnoff M (2024) rchemo: Dimension reduction, regression and discrimination for chemometrics. R Packag. https://github.com/ChemHouse-group/rchemo
51. Kalinowski ST (2004) Genetic polymorphism and mixed-stock fisheries analysis. Can J Fish Aquat Sci 61:1075–1082. https://doi.org/10.1139/F04-060
52. Nielsen R, Tarpy DR, Kern Reeve H (2003) Estimating effective paternity number in social insects and the effective number of alleles in a population. Mol Ecol 12:3157–3164
53. Nei M (1972) Genetic distance between populations. Am Nat 106:283–292. https://doi.org/10.1086/282771
54. Hardy OJ, Vekemans X (2002) SPAGeDI: a versatile computer program to analyse spatial genetic structure at the individusal or population levels. Mol Ecol 2:618–620. https://doi.org/10.1046/j.1471-8278
55. Chybicki IJ, Burczyk J (2009) Simultaneous estimation of null alleles and inbreeding coefficients. J Hered 100:106–113. https://doi.org/10.1093/jhered/esn088
56. Loiselle B, Sork V, Nason J, Graham C (1995) Spatial genetic structure of a tropical understory shrub, *Psychotria officinalis* (Rubiaceae). Am JBot 1420–1425. https://www.jstor.org/stable/2445869?seq=1
57. Wold S, Sjöström M, Eriksson L (2001) PLS-regression: a basic tool of chemometrics. Chemom Intell Lab Syst 58:109–130
58. Mantel N (1967) The detection of disease clustering and a generalized regression approach. Cancer Res 27:209–220
59. Fajardo A, Piper FI, García-Cervigón AI (2022) The intraspecific relationship between wood density, vessel diameter and other traits across environmental gradients. Funct Ecol 36:1585–1598. https://doi.org/10.1111/1365-2435.14069
60. Zhang Z, Zhong H, Avramidis S et al (2025) Transfer learning for predicting wood density of different tree species: calibration transfer from portable NIR spectrometer to hyperspectral imaging. Wood Sci Technol. https://doi.org/10.1007/s00226-024-01615-5
61. Debout GDG, Doucet JL, Hardy OJ (2011) Population history and gene dispersal inferred from spatial genetic structure of a Central African timber tree, *Distemonanthus benthamianus* (Caesalpinioideae). Heredity (Edinb) 106:88–99. https://doi.org/10.1038/hdy.2010.35
62. Berry NJ, Phillips OL, Ong RC, Hamer KC (2008) Impacts of selective logging on tree diversity across a rainforest landscape: the importance of spatial scale. Landsc Ecol 23:915–929. https://doi.org/10.1007/s10980-008-9248-1
63. Dick CW, Hardy OJ, Jones FA, Petit RJ (2008) Spatial scales of pollen and seed-mediated gene flow in Tropical Rain Forest trees. Trop Plant Biol 1:20–33. https://doi.org/10.1007/s12042-007-9006-6
64. Bhasin O, Doucet J-L, Ndonda Makemba R et al (2024) Contrasted spatial, demographic and genetic structures of a light-demanding African timber species, *Cylicodiscus gabunensis* Harms—implications for a sustainable management of its populations. For Ecol Manage. https://doi.org/10.1016/j.foreco.2023.121527. (**ISSN 0378-1127,**)
65. Born C, Kjellberg F, Chevallier MH et al (2008) Colonization processes and the maintenance of genetic diversity: insights from a pioneer rainforest tree, *Aucoumea klaineana*. Proc R Soc B Biol Sci 275:2171–2179. https://doi.org/10.1098/rspb.2008.0446
66. Hardy OJ, Delaide B, Hainaut H et al (2019) Seed and pollen dispersal distances in two African legume timber trees and their reproductive potential under selective logging. Mol Ecol 28:3119–3134. https://doi.org/10.1111/mec.15138
67. Arruda CCB, Silva MB, Sebbenn AM et al (2015) Mating system and genetic diversity of progenies before and after logging: a case study of *Bagassa guianensis* (Moraceae), a low-density dioecious tree of the Amazonian forest. Tree Genet Genomes. https://doi.org/10.1007/s11295-015-0837-2
68. Duminil J, Brown RP, Ewédjè EEB et al (2013) Large-scale pattern of genetic differentiation within African rainforest trees: insights on the roles of ecological gradients and past climate changes on the evolution of *Erythrophleum* spp (Fabaceae). BMC Evol Biol 13:1. https://doi.org/10.1186/1471-2148-13-195
69. Hardy OJ, Born C, Budde K et al (2013) Comparative phylogeography of African rain forest trees: a review of genetic signatures of vegetation history in the Guineo-Congolian region. C R Geosci 345:284–296. https://doi.org/10.1016/j.crte.2013.05.001
70. Loureiro BA, Arriel TG, Ramalho FMG et al (2022) NIR-based models for estimating selected physical and chemical wood properties from fast-growing plantations. IForest 15:372–380. https://doi.org/10.3832/ifor4030-015
71. Cooper PA, Jeremic D, Radivojevic S et al (2011) Potential of near-infrared spectroscopy to characterize wood products 1. Can J For Res 41:2150–2157. https://doi.org/10.1139/x11-088
72. Silveira L, Sathaiah S, Zângaro RA et al (2003) Near-infrared Raman spectroscopy of human coronary arteries: histopathological classification based on Mahalanobis distance. J Clin Laser Med Surg 21:203–208. https://doi.org/10.1089/104454703768247774
73. Baillères H, Davrieux F, Ham-Pichavant F (2002) Near infrared analysis as a tool for rapid screening of some major wood characteristics in a *Eucalyptus* breeding program. Ann For Sci 59:479–490. https://doi.org/10.1051/forest:2002032
74. Tsuchikawa S, Kobori H (2015) A review of recent application of near infrared spectroscopy to wood science and technology. J Wood Sci 61:213–220. https://doi.org/10.1007/s10086-015-1467-x
75. Hein PRG, Chaix G, Clair B et al (2016) Spatial variation of wood density, stiffness and microfibril angle along *Eucalyptus* trunks grown under contrasting growth conditions. Trees Struct Funct 30:871–882. https://doi.org/10.1007/s00468-015-1327-8
76. Bradbury GJ, Potts BM, Beadle CL (2011) Genetic and environmental variation in wood properties of *Acacia melanoxylon*. Ann For Sci 68:1363–1373. https://doi.org/10.1007/s13595-011-0115-x
77. Zemunik G, Davies SJ, Turner BL (2018) Soil drivers of local-scale tree growth in a lowland tropical forest. Ecology 99:2844–2852. https://doi.org/10.1002/ecy.2532
78. Hacke U (2015) Functional and ecological xylem anatomy
79. Degen B, Blanc L, Caron H et al (2006) Impact of selective logging on genetic composition and demographic structure of four tropical tree species. Biol Conserv 131:386–401. https://doi.org/10.1016/j.biocon.2006.02.014
80. Sebbenn AM, Degen B, Azevedo VCR et al (2008) Modelling the long-term impacts of selective logging on genetic diversity and demographic structure of four tropical tree species in the Amazon forest. For Ecol Manage 254:335–349. https://doi.org/10.1016/j.foreco.2007.08.009
81. Wernsdörfer H, Caron H, Gerber S et al (2011) Relationships between demography and gene flow and their importance for the conservation of tree populations in tropical forests under selective felling regimes. Conserv Genet 12:15–29. https://doi.org/10.1007/s10592-009-9983-0
82. De la Mensbruge G (1966) La germination et les plantules des essences arborées de la forêt dense humide de la Côte d'Ivoire. Centre Technique Forestier Tropical, Nogent-sur-Marne, France
83. Dupuy B (1998) Bases pour une sylviculture en forêt dense tropicale humide africaine. https://forafri.cirad.fr/ressources/forafri/09.pdf
84. Hall JS (2008) Seed and seedling survival of African mahogany (*Entandrophragma* spp.) in the Central African Republic: implications for forest management. For Ecol Manage 255:292–299. https://doi.org/10.1016/j.foreco.2007.09.050

Publisher's Note

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