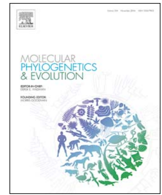
Molecular Phylogenetics and Evolution 120 (2018) 83–93

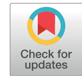
Contents lists available at ScienceDirect

Molecular Phylogenetics and Evolution

journal homepage: www.elsevier.com/locate/ympev



Evolution in the Amphi-Atlantic tropical genus *Guibourtia* (Fabaceae, Detarioideae), combining NGS phylogeny and morphology



Félicien Tosso[a,b,*], Olivier J. Hardy[b], Jean-Louis Doucet[a], Kasso Daïnou[c,d,e], Esra Kaymak[b], Jérémy Migliore[b]

[a] TERRA Research Centre, Central African Forests, Gembloux Agro-Bio Tech, University of Liège, Passage des Déportés 2, B-5030 Gembloux, Belgium
[b] Evolutionary Biology and Ecology Unit CP 160/12, Faculté des Sciences, Université Libre de Bruxelles, 50 avenue F. Roosevelt, B-1050 Brussels, Belgium
[c] Nature + asbl / TERRA Research Centre, Central African Forests, Gembloux Agro-Bio Tech, University of Liege, Passage des Déportés 2, 5030 Gembloux, Belgium
[d] BIOSE Department, Management of Forest Resources, Gembloux Agro-Bio Tech, Université de Liège, Passage des Déportés 2, B-5030 Gembloux, Belgium
[e] Université Nationale d'Agriculture, BP 43 Kétou, Benin

ARTICLE INFO

Keywords:
Plastid genomes captures
Guibourtia
Molecular phylogeny
Morphology
Next generation sequencing
Speciation

ABSTRACT

Tropical rain forests support a remarkable diversity of tree species, questioning how and when this diversity arose. The genus *Guibourtia* (Fabaceae, Detarioideae), characterized by two South American and 13 African tree species growing in various tropical biomes, is an interesting model to address the role of biogeographic processes and adaptation to contrasted environments on species diversification. Combining whole plastid genome sequencing and morphological characters analysis, we studied the timing of speciation and diversification processes in *Guibourtia* through molecular dating and ancestral habitats reconstruction. All species except *G. demeusei* and *G. copallifera* appear monophyletic. Dispersal from Africa to America across the Atlantic Ocean is the most plausible hypothesis to explain the occurrence of Neotropical *Guibourtia* species, which diverged *ca.* 11.8 Ma from their closest African relatives. The diversification of the three main clades of African *Guibourtia* is concomitant to Miocene global climate changes, highlighting pre-Quaternary speciation events. These clades differ by their reproductive characters, which validates the three subgenera previously described: *Pseudocopaiva*, *Guibourtia* and *Gorskia*. Within most monophyletic species, plastid lineages start diverging from each other during the Pliocene or early Pleistocene, suggesting that these species already arose during this period. The multiple transitions between rain forests and dry forests/savannahs inferred here through the plastid phylogeny in each *Guibourtia* subgenus address thus new questions about the role of phylogenetic relationships in shaping ecological niche and morphological similarity among taxa.

1. Introduction

Despite the remarkable species richness of the tropical biomes and conservation issues, one critical question remains how and when rain forest diversity arose (e.g. Plana, 2004; Couvreur et al., 2011a; Couvreur, 2015). Biodiversity is unequally distributed across tropical continents (Slik et al., 2015) and we are still far from having a synthetic explanation about the lower apparent species diversity in African rain forests when compared to the Neotropics or South-East Asia (Couvreur, 2015). Dated molecular phylogenies are thus needed to reconstruct the history and evolution of taxa and understand the current biogeographical dynamics of tropical ecosystems at the inter- and intra-continental scales.

At an inter-continental scale, many plant families characteristic of rain forest ecosystems show a pantropical distribution in the three major tropical regions (America, Africa and Asia) and many of their taxa (e.g. genera, sometimes even species) are found in two or three of these regions (Couvreur et al., 2011b). Several hypotheses have been proposed to explain such disjunct distributions: (i) the break-up of the Gondwanan supercontinent (e.g. Raven and Axelrod, 1974; Conti et al., 2002), (ii) the degradation of boreotropical flora and the end of a northern mid-latitude migration corridor (e.g. Davis et al., 2002; Zerega et al., 2005; Muellner et al., 2006), and (iii) long-distance dispersal (e.g. Renner et al., 2001). In many plant taxa, when molecular data succeed in rejecting the vicariance hypotheses, they tend to confirm long distance dispersal (Pennington et al., 2006; Christenhusz and Chase, 2013; Baker and Couvreur, 2013; Armstrong et al., 2014). However, new molecular-dated inter-continental phylogenies are still needed to

* Corresponding author at: TERRA Research Centre, Central African Forests, Gembloux Agro-Bio Tech, University of Liège, Passage des Déportés 2, B-5030 Gembloux, Belgium.
E-mail addresses: tfelicien@yahoo.fr, dnftosso@ulg.ac.be (F. Tosso), ohardy@ulb.ac.be (O.J. Hardy), jldoucet@ulg.ac.be (J.-L. Doucet), kdainou@ulg.ac.be (K. Daïnou), ekaymak@ulb.ac.be (E. Kaymak), jeremy.migliore@ulb.ac.be (J. Migliore).

https://doi.org/10.1016/j.ympev.2017.11.026
Received 29 November 2016; Received in revised form 22 November 2017; Accepted 30 November 2017
Available online 05 December 2017
1055-7903/ © 2017 Elsevier Inc. All rights reserved.

provide a more reliable view of the age and geographical origin of taxa showing a disjunct distribution and whether the timing and tempo of speciation on each continent coincides with geoclimatic events (Couvreur, 2011a; Richardson et al., 2015).

At the intra-continental scale, a main challenge is to validate species delineations and to infer the most putative evolutionary history of target taxa. However, there is still some limitations when applying the phylogenetic concept of species due to heterogeneous evolutionary rates between genetic markers and incomplete lineage sorting, the latter being especially important for trees with long generation time and large effective size (Daïnou et al., 2014, 2016). For Central African trees, the process of species diversification remains poorly documented, despite the emergence of some global trends. Firstly, among the few attempts to characterize the evolutionary history of rain forest trees, it appears that Miocene speciation events are not incompatible with incomplete lineage sorting and scarce morphological differences (e.g., Daïnou et al., 2014; Donkpegan et al., 2017). Secondly, studies highlighted that the most recent Pleistocene glacial/interglacial climatic oscillations promoted intraspecific diversification, but not speciation (e.g. Couvreur et al., 2011a; Duminil et al., 2015; Pineiro et al., 2017). Finally, both congruent phylogeographical patterns and idiosyncratic histories are detected among African rain forest species, especially for explaining both the role of gene flow barriers and refuge areas (e.g. Hardy et al., 2013; Dauby et al., 2014).

Here, we focus on the genus *Guibourtia* Benn. (Fabaceae, Detarioideae) because it is characterized by two South American and 13 African species (Léonard, 1949; Tosso et al., 2015) whereas the Neotropical flora is globally richer than the African one (Couvreur, 2015). In addition, the 13 African species of *Guibourtia* occupy contrasted forest habitats and appear well suited to infer the imprints left by past environmental changes on rain forest evolution. Besides, to date, botanists did not fully agree on species boundaries in this genus due to the existence of morphologically close species found in sympatry in Africa (Tosso et al., 2015). Using whole plastid genome sequencing by enrichment/hybridization/capture (McPherson et al., 2013; Mariac et al., 2014), the present study aims to better understand the evolutionary history of speciation and diversification in the genus *Guibourtia*. Modern high-throughput sequencing techniques and their applications in phylogeny have become essential to understand complex evolutionary scenarios that could not have been resolved before by sequencing a few genes (Faye et al., 2016). Indeed, they improve substantially the resolution of phylogenetic inferences even at low taxonomic levels or where recent divergence, rapid speciation or slow genome evolution occurred (Malé et al., 2014; Williams et al., 2016). This should allow a higher reliability of both phylogenetic relationships and molecular dating, benefiting systematics.

Specifically, we will address the following questions in this study:

– What are the phylogenetic relationships between Neotropical and African *Guibourtia* species? How and when did *Guibourtia* occupy an Amphi-Atlantic distribution (vicariance *versus* long-distance dispersal)? For these questions, we will analyze the relative position of American and African taxa in the phylogeny, examining the temporal window of speciation and migration between continents.
– Is the current species delimitation of African *Guibourtia* supported by the phylogeny of their plastomes? How far does phylogenetic divergence reflect morphological differentiation among *Guibourtia* species? When and in which direction(s) did habitat transition(s) occur and can the diversification of the genus in Africa be explained by the climatic history of the continent? Here, we will verify whether species represented by multiple samples appear monophylic, and study the phylogenetic relationships between African *Guibourtia* taxa in the light of their morphologic features and habitat preferences.

2. Materials and methods

2.1. Biological model and sampling strategy

The study was conducted on all the 15 species of the genus *Guibourtia* (Fabaceae, Detarioideae), all found in tropical regions. The two neotropical species studied were *G. chodatiana* (Hassl.) J. Léonard and *G. hymenaefolia* (Moric.) J. Léonard which occur in South American rain forests. The 13 African species of *Guibourtia* include seven tropical rain forest species, *G. arnoldiana* (De Wild. & T. Durand) J. Léonard, *G. demeusei* (Harms) J. Léonard, *G. dinklagei* (Harms) J. Léonard, *G. ehie* (A. Chev.) J. Léonard, *G. leonensis* J. Léonard, *G. pellegriniana* J. Léonard, *G. tessmannii* (Harms) J. Léonard, and six dry forest or savannah species, *G. carrissoana* (M.A. Exell) J. Léonard, *G. coleosperma* (Benth.) J. Léonard, *G. conjugata* (Bolle) J. Léonard, *G. copallifera* Benn., *G. schliebenii* (Harms) J. Léonard, and *G. sousae* J. Léonard.

The whole plastid genome was captured and sequenced for each species using one to five individuals per species ($n = 40$; Supplementary material Table S1). In addition, five individuals from highly divergent species (based on known morphological differentiation) were selected for genomic libraries sequencing without enrichment. Two genomic libraries were sequenced for two outgroup taxa among the closest known sister taxa of *Guibourtia*: *Crudia harmsiana* De Wild and *Daniellia pilosa* (J. Léonard) Estrella (Fougère-Danezan, 2005). Samples were collected in herbaria (Botanic Garden Meise "BR", Natural History Museum Paris "P") and from field samples from which leaf material was directly dried in silica gel.

2.2. DNA library construction and whole plastid genome capture

Total genomic DNA was extracted from dried leaves using the CTAB protocol (Doyle, 1987) coupled with QIAquick purification kits (Qiagen, Venlo, Netherlands), and followed by Qubit® 2.0 Fluorometer quantification (Life Technologies, Invitrogen, Foster City, USA) and by QIAxcel (Qiagen) DNA quality control.

Library preparation for multiplexed individuals was based on the protocol of Rohland and Reich (2012) but with the inclusion of additional steps published by Mariac et al. (2014) for the plastid enrichment procedure (biotinylated probes capture). A 5–10 µg of DNA aliquot per sample was sheared by sonication using a Bioruptor® Pico (Diagenode SA., Liège, Belgium) to a mean fragment size of *ca.* 400 bp. Sheared DNA was sized by dual fragment size selection using AMPure XP magnetic beads (Agencourt, Beckman Coulter, Brea, USA) to remove remaining larger (> 600 bp) and shorter (< 200 bp) DNA fragments. The next steps included blunting and 5′-phosphorylation using the Fast DNA End Repair Kit, ligation of tagged adapters using T4 DNA ligase, and a nick fill-in step performed with Bst DNA polymerase (New England Biolabs Inc. NEB, Beverly, USA). Tagging was made with multiplex 6-bp indexed adapters to enable pooling of libraries for sequencing, following the protocol published by Rohland and Reich (2012). A real-time PCR was then performed to extend Illumina adaptors and enrich library fragments using the KAPA HiFi RT PCR library amplification kit (KAPA Biosystems, Boston, USA), on a StepOnePlus analyzer (Applied Biosystems, Foster City, USA). Most of the steps were followed by AMPure XP bead-based sample clean-up steps and a check of DNA quality using a QIAxcel (Qiagen).

To enrich genomic libraries in plastid sequences, we applied the sequence capture using PCR probes (SCPP) method (Peñalba et al., 2014) adapted to plastid DNA by Mariac et al. (2014). Biotinylated probes were produced in the laboratory from one sample of *Guibourtia tessmannii* collected in the field, using 18 long-range universal PCR primers combinations (Uribe-Convers et al., 2014) with the LongAmp® Taq PCR Kit (NEB). These probes were sheared and sized, targeting 400 bp in length, as previously done for genomic libraries. Probes biotinylation was done with a 5′ TEG-biotinylated linker using a Phusion® High-Fidelity PCR Master Mix (Invitrogen, Carlsbad, USA). In

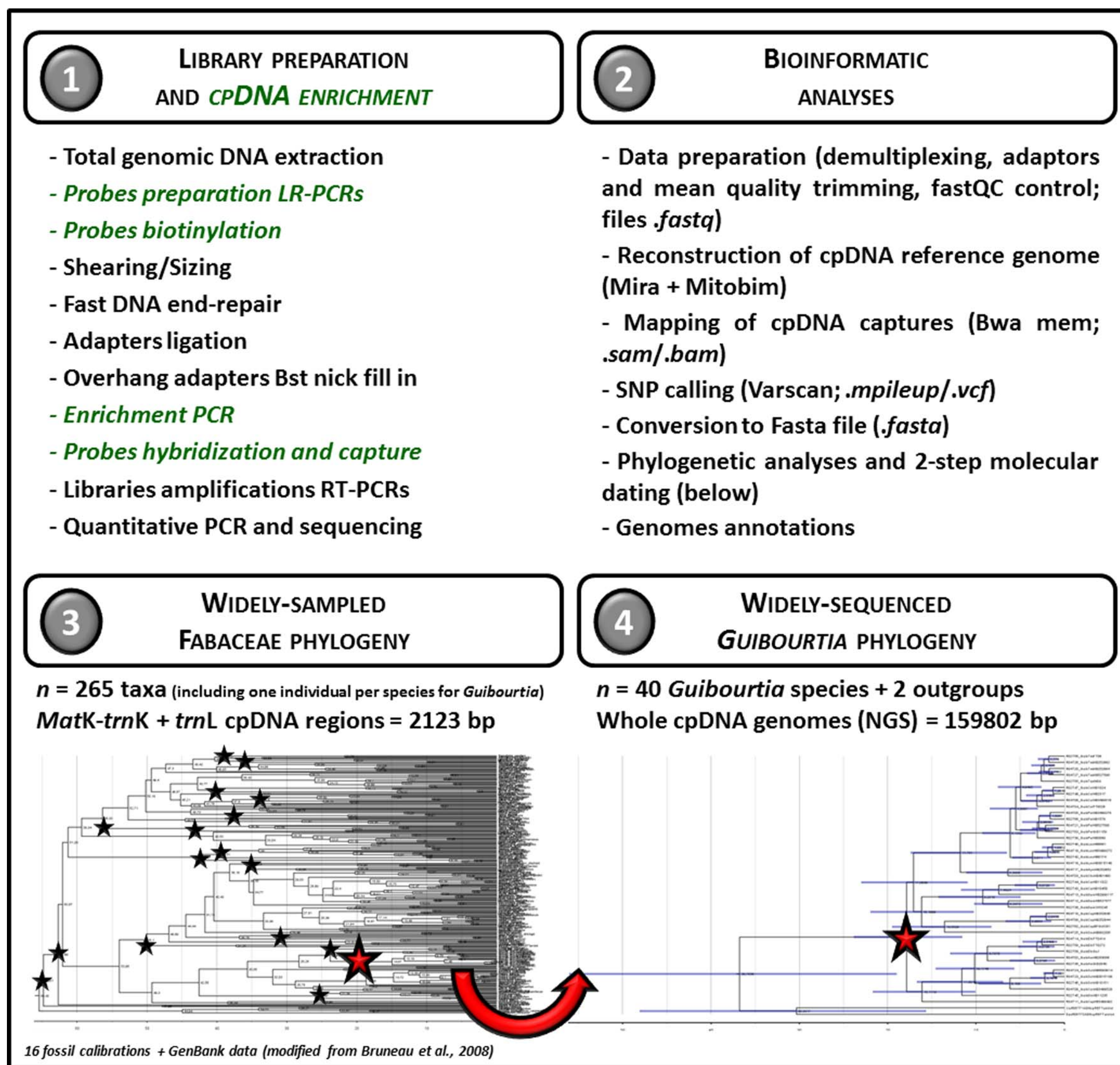


Fig. 1. Main laboratory and bioinformatic steps followed for reconstructing the phylogeny of the genus *Guibourtia*.

solution capture of plastid DNA was done by hybridization of biotinylated probes to genomic libraries during 48 h, with then immobilization of target DNA by streptavidin beads (Dynabeads C1, Invitrogen). A prehybridization has been undertaken before enrichment with Phusion Taq to increase the concentration of DNA aliquots.

The plastid genome enrichment phase does not guarantee that all parts of the plastid genome will be captured, thus the protocol of genomic libraries preparation was also applied without enrichment to reconstruct a reference plastid genome on one sample among five species tested (highest quality library conserved), using *ca.* 1/20th of the MiSeq run *versus ca.* 1/100th of the run for each plastid genome enriched sample. Independent genomic libraries were also sequenced for the outgroup taxa (without enrichment). All these steps are summarized in Fig. 1.

Libraries were titrated for sequencing using the KAPA Library Quantification Kit (Kapa Biosystems, Boston, USA) and fragment size distributions checked using the QIAxcel, before pooling libraries. Paired-end sequencing (2×150 bp) was performed on an Illumina MiSeq with reagent kit V2 at the GIGA platform (Liège, Belgium).

2.3. Plastid genomes reconstruction

Standard Illumina pipeline software (v 1.7 or later) was used for base-calling and quality filtering, with standard chastity settings used to produce fastq data files whose quality was checked with FastQC. Demultiplexing based on the 6-bp barcodes was performed using the freely available python script Demultadapt (https://github.com/Maillol/demultadapt), using a 0-mismatch threshold. Adapters and low-quality bases were trimmed using Cutadapt 1.2.1 (Martin, 2011) with the following options: quality cut-off = 20, minimum overlap = 7 and minimum length = 35. Reads with a mean quality lower than 30 were discarded afterwards using a freely available perl script (https://github.com/SouthGreenPlatform/arcad-hts/blob/master/scripts/arcad_hts_2_Filter_Fastq_On_Mean_Quality.pl).

The software Mitobim 1.5 (Hahn et al., 2013) was used to reconstruct the plastid genome of *Guibourtia*, assembling the reads from the non-enriched genomic DNA library. Mitobim mapped these reads using MIRA 3.4.1.1 (Chevreux et al., 1999) to highly conserved regions of the closely related reference genomes available in GenBank; *Tamarindus indica* (Fabaceae, Detarioideae; KJ468103) was retained after comparison with other genomes. The GenBank genome used to initiate the process of reference plastid genome reconstruction is not assumed to have the correct linear order and the resolving process specifically detects rearrangements and inversions, and corrects these.

The mapping of plastid DNA (pDNA) enriched libraries was performed using Bwa Mem 0.7.5a-r405 (Li and Durbin, 2009) with -M and -B 4 options, using the plastid reference genomes reconstructed in the previous steps with the reference genomic library of *Guibourtia*.

Samtools 1.1 (Li and Durbin, 2009) was used to generate a mpileup file with option -B, following recommendations of Scarcelli et al. (2016). Finally, we used Varscan 2.3.7 (Koboldt et al., 2012) to call SNPs (option –min-var-freq set to 50%) before converting it to a fasta multi-alignment file (bioinformatic pipeline summarized in Fig. 1).

The final complete plastid genomes were annotated using Geneious v. 7.1.3 (Kearse et al., 2012) and through comparison to published complete plastid genome sequences available in GenBank. The newly generated sequences are available in GenBank (see Table S1 for accession numbers) and bam files are available on request to authors.

2.4. Phylogenetic analyses and molecular dating

A two-step molecular dating approach was used. The first step used a large taxonomic coverage of Fabaceae and several fossil calibrated nodes to estimate the crown age of the genus *Guibourtia* using mostly published sequences for a limited number of genes. The second step used our whole plastid genomes of *Guibourtia* samples to estimate the nodes ages of the *Guibourtia* phylogeny (Fig. 1). The phylogenies were reconstructed with the Bayesian MCMC analysis implemented in BEAST 1.8.2 (Bayesian Evolutionary Analysis by Sampling Trees; Drummond and Rambaut, 2007).

For the first step, we extracted the *mat*K/*trn*K (analyzed separately) and *trn*L plastid DNA regions from one plastid genome sequenced for each *Guibourtia* species ($n = 15$), and the NGS genomic libraries of *Daniellia pilosa* and *Crudia harmsiana*. These data were combined to the same plastid DNA regions sequenced for 250 Fabaceae species (including only taxa with data available for these three plastid regions) by Bruneau et al. (2008) and available in GenBank. Sequences were aligned in MAFFT 7 (Multiple alignment program for amino acid or nucleotide sequences: http://mafft.cbrc.jp/alignment/software/). The models GTR + G, GTR + G and HKY + G were chosen using jModelTest 2.1.7 (Darriba et al., 2012) for *mat*K, *trn*K, and *trn*L alignments, respectively. Sixteen fossil calibration points were used (see Table S2) to anchor the phylogeny, following Bruneau et al. (2008), where all these fossils are described in detail. Each calibration point was parameterized as minimum age with a normal, lognormal or gamma distribution (Table S2). These nodes and those of strongly supported clades in the Fabaceae (Bruneau et al., 2008) were constrained as monophyletic.

For the second step (Fig. 1), we anchored the phylogeny of the plastid genomes of 40 *Guibourtia* samples by setting the age of the crown node of *Guibourtia* taxa as a normal distributed random variable (mean = 19.41, StDev = 3.0) with a 95% interval between 14.48 and 24.35 Ma, as obtained through the first step. The model GTR + G was applied.

At each molecular dating step, an XML file was generated using BEAUti (Bayesian Evolutionary Analysis Utility). An uncorrelated lognormal relaxed clock model and a Yule process of speciation were applied. One MCMC analysis was run for 100 million generations each, sampling trees at 10,000 step intervals. We used Tracer 1.4 (Drummond and Rambaut, 2007) to assess convergence, estimate Effective Sample Sizes (ESS), and examine the posteriors of all the parameters. Mean heights were taken in TreeAnnotator 1.4.8 (Drummond and Rambaut, 2007), and trees were plotted in FigTree 1.1.2 (http://tree.bio.ed.ac.uk/software/figtree/).

In parallel, a phylogenetic network of the 40 *Guibourtia* plastomes was established from a Neighbor-Net analysis (Bryant and Moulton, 2004) using SplitsTree 4.14.2 (Huson and Bryant, 2006). This phylogenetic Neighbor-Joining distance-based method works quickly by agglomerating clusters and tends to construct resolved haplotype networks representing both groupings in the data and evolutionary distances between taxa.

2.5. Ancestral characters reconstruction

Ancestral states reconstructions were performed for habitat types (rain forest *versus* dry forest/savannah), using a maximum likelihood approach implemented in the R package phytools (Revell, 2012). We also applied maximum parsimony principles to infer habitat transitions, assuming that species can be adapted to only one type of habitat.

2.6. Morphological data

In total, 45 descriptors were scored for the 40 individuals used for whole plastome sequencing. These descriptors were sorted into vegetative and reproductive traits (Table S3). Specimens were collected between 1889 and 2010 and kept in the National Museum of Natural History in Paris P (France) and the Botanic Garden Meise BR (Belgium). Only the specimens for which the determination has been validated by botanists specialized in Fabaceae (most often by J. Leonard, J. Wieringa, M. Fougère-Danezan or R. Letouzey) were taken into account.

The morphological characters used were those listed in the determination keys of Léonard (1949), Aubréville (1970), Fougère-Danezan et al. (2010), and Tosso et al. (2015) (Table S3). We added characters potentially displaying high interspecific variation as the presence of glands on the leaflets, the shape of the apex of leaflets, the persistence of stipules, and the type of fruit (Table S3). After being removed, floral organs to be measured were rehydrated in boiling water at 90 °C for 3 min. The flowers (1–3) were then dissected and observed. The floral pieces were measured with a micrometer incorporated into a binocular microscope (Nachet GLI 154), magnification x 10–40. Microscopy was also used to check for the presence/absence of glands and the hairiness on the leaflets, on the flowers or fruits. The other qualitative characteristics (form of leaflets, position of the median veins, and secondary veins, etc.) were directly observed.

To describe the morphological similarity between species, a matrix of morphometric distances between the 40 analysed specimens was constructed by calculating the Gower distance (Gower, 1971) for (i) all morphometric characters, (ii) reproductive characters only, and (iii) vegetative characters only, using the R package FD (Laliberté et al., 2010). Hill and Smith analysis (Hill and Smith, 1976; Kiers, 1991) allowed to ordinate both qualitative and quantitative morphometric variables. Then, we considered the scree plot (Cattell, 1966) and Kaiser criterion (Kaiser, 1960) to assess the number of axes that represent enough the ordination patterns after a Principal Coordinate Analysis (PCoA), using the R package ade4 (Chessel et al., 2012).

3. Results

3.1. Plastome sequencing data

The non-enriched genomic library of *G. leonensis* used for reconstructing the reference plastid genome (size of 159,802 bp) contained 1,454,120 R1-R2 paired reads, of which 2.29% appeared to be of plastid origin, making an average depth of 27×. The genomic libraries of *Daniellia pilosa* and *Crudia harmsiana* used for reconstructing outgroup plastid genomes (sizes of 160,314 bp and 159,664 bp, respectively) contained 1,307,116 and 1,638,114 R1-R2 paired reads, of which 4.5% and 0.9% were of plastid origin, respectively.

For the 40 pDNA enriched libraries of *Guibourtia*, on average, 332,974 R1-R2 paired reads (standard deviation, stdev 272,188 R1-R2 paired reads) were obtained for each NGS plastid genome capture, of which 96.7% (stdev 1.8%) appeared to be of plastid origin on average, illustrating a good yield for pDNA enrichment (Table S1). The average depth was 285× (mean stdev 223) and the mean coverage at ≥10× was around 84.4% (stdev 7.5%). After SNP (Single Nucleotide Polymorphic) calling, 6300 SNPs and 1412 indels were retained and used for reconstructing fasta files of complete plastid genomes.

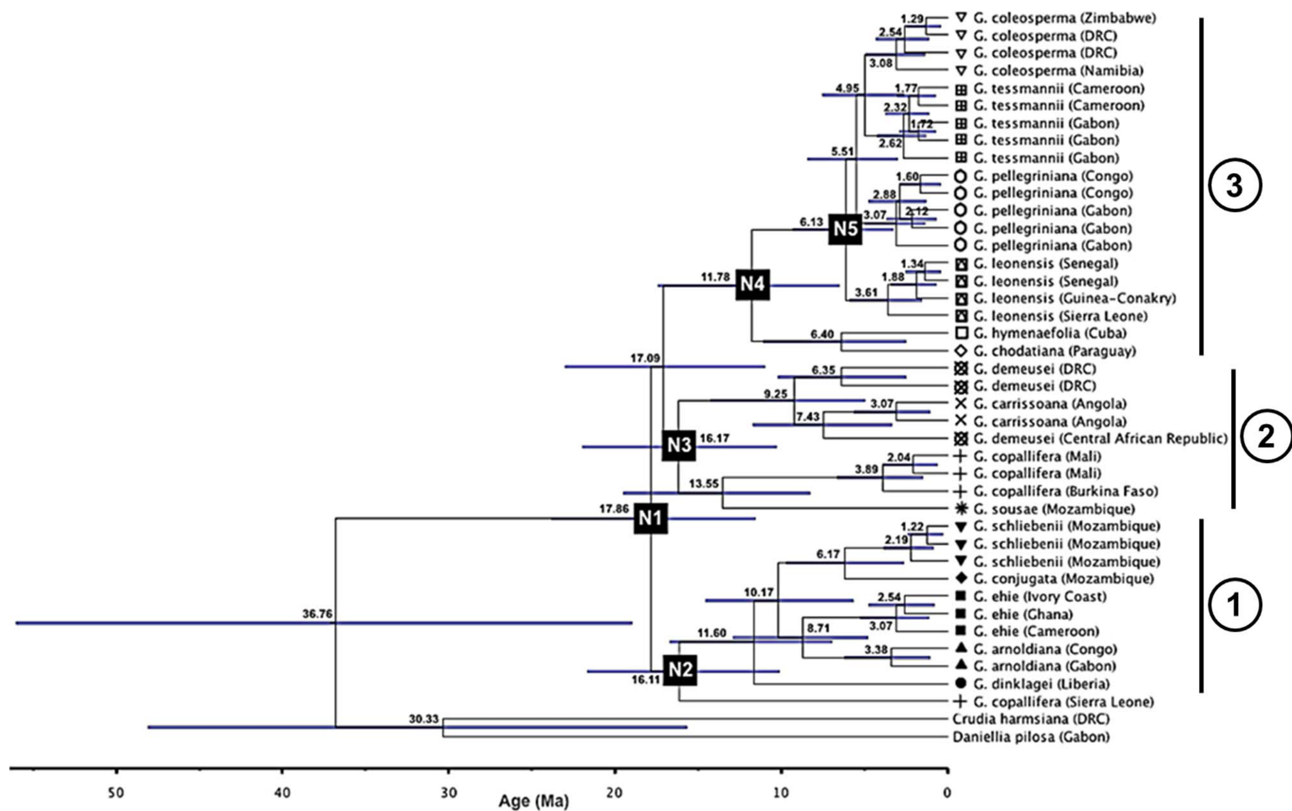


Fig. 2. Divergence time chronogram of the whole plastid genome sequenced for *Guibourtia* and two outgroup taxa, using BEAST software (Bayesian maximum clade credibility tree). Numbers on nodes indicate the mean divergence time estimates (in Ma), with bars referring to the 95% highest posterior density intervals around node ages. Nodes N1 to N5 are discussed in the text (see Table S2). All, except two nodes, are supported at posterior probabilities of at least 0.99 (see Fig. S2).

3.2. Phylogenetic relationships and molecular dating from plastid genomes

Integrating our plastid genomes for one individual per species within the Fabaceae phylogeny published by Bruneau et al. (2008) for the *mat*K-*trn*K and *trn*L plastid DNA regions, the genus *Guibourtia* was well supported as monophyletic (posterior probability of 1; Table S2). The genus *Hymenaea* is sister to *Guibourtia*, and *Peltogyne* is sister to both genera. In this Fabaceae phylogeny anchored by 16 fossil calibration points, the crown node of *Guibourtia* was dated to 19.41 Ma, with a 95% higher posterior density interval ranging between 14.50 and 24.26 Ma (Fig. S1). These divergence values were then implemented in our intra-*Guibourtia* phylogeny exclusively based on complete plastid genomes. For each molecular dating step, most of the parameters had reached stationarity with ESS values over 200, thus deemed reliable (all the parameters have reached at least 100).

The Bayesian Maximum Clade Credibility phylogeny provided by BEAST at the intra-*Guibourtia* level (40 *Guibourtia* taxa + 2 outgroups) showed strong overall support for each node (posterior probabilities mainly of 1 or 0.99 for recent divergences; Fig. S2). Three clades diverged around 17.08–17.86 Ma (95% HPD of 10.98–22.98 and 11.57–23.80 Ma, respectively between clades 1/2–3 and clades 2/3; Fig. 2). The clade 1 is composed of *G. arnoldiana*, *G. ehie*, *G. conjugata*, *G. schliebenii*, *G. dinklagei* and one of the four samples of *G. copallifera* from Sierra Leone. The clade 2 included *G. demeusei*, *G. carrissoana*, *G. sousae*, and three samples of *G. copallifera* collected in Burkina Faso, Mali, and Congo. The clade 3 is composed of *G. tessmannii*, *G. pellegriniana*, *G. leonensis*, and *G. coleosperma*. The Neotropical taxa *G. chodatiana* and *G. hymenaefolia* are included in the clade 3, but diverged from other African *Guibourtia* around 11.78 Ma (95% HPD: 6.52–17.38 Ma). When several samples were analysed per species (10/15 species), the monophyly of the species was always well supported except for *G. demeusei* and *G. copallifera*.

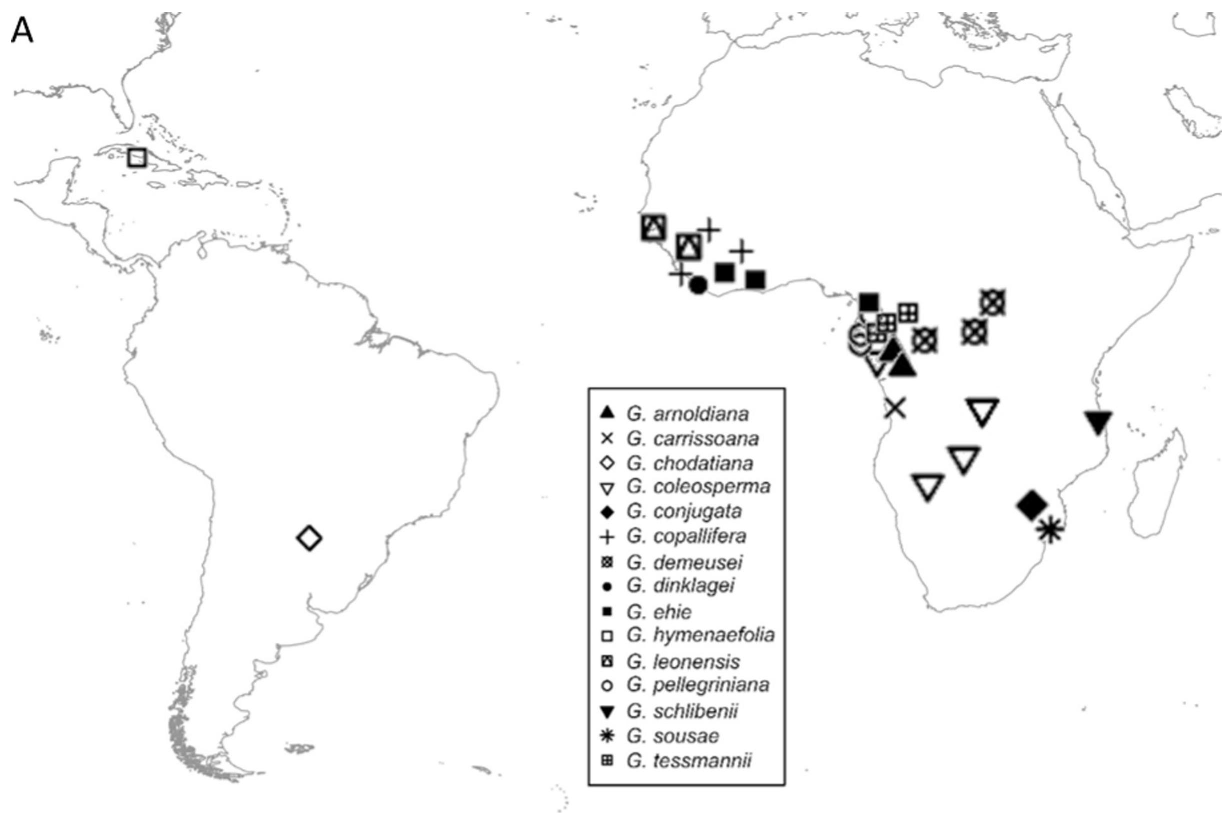
The NeighborNet analysis (Fig. 3) is congruent with the result of the BEAST analyses, with the same three clades identified. The divergence between taxa belonging to clade 1 is relatively more recent than that observed within the other clades. Especially, the NeighborNet tree revealed an obvious star-like topology for the clade 1.

3.3. Habitat transitions

Each of the three clades contains rain forest (RF) species and dry forest/savannah (DFS) species (Fig. 4). By using maximum likelihood approach for reconstructing ancestral states, the most recent common ancestor of *Guibourtia* species was difficult to assign to a specific type of habitat, the scaled likelihood at the root reaching 57% for RF and 43% for DFS habitats (Fig. 4). Applying maximum parsimony principles, if we assume a RF ancestral state, at least four habitat transitions are needed, either four RF to DFS transitions, or three RF to DFS and one DFS to RF transitions. If we assume a DFS ancestral state, at least five habitat transitions are needed, either four DFS to RF and one RF to DFS transitions, or three DFS to RF and two RF to DFS transitions. The only transition common to all these scenarios is the RF to DFS transition leading to *G. coleosperma*, which would have occurred between 3 and 5 Ma, thus during the Pliocene.

3.4. Morphometric delineation in Guibourtia species

To disentangle morphological units, the first two components of the Principal Coordinate Analysis (PCoA) explained 33.75% and 19.38% of the total morphometric variation. The PCoA analysis scatter plot showed three morpho-groups (Fig. 5A). The first group is composed of *G. copallifera*, *G. demeusei*, *G. carrissoana*, and *G. sousae*. The second group included *G. arnoldiana*, *G. ehie*, *G. conjugata*, *G. schliebenii*, and *G. dinklagei*. The third group is composed of *G. tessmannii*, *G. pellegriniana*,



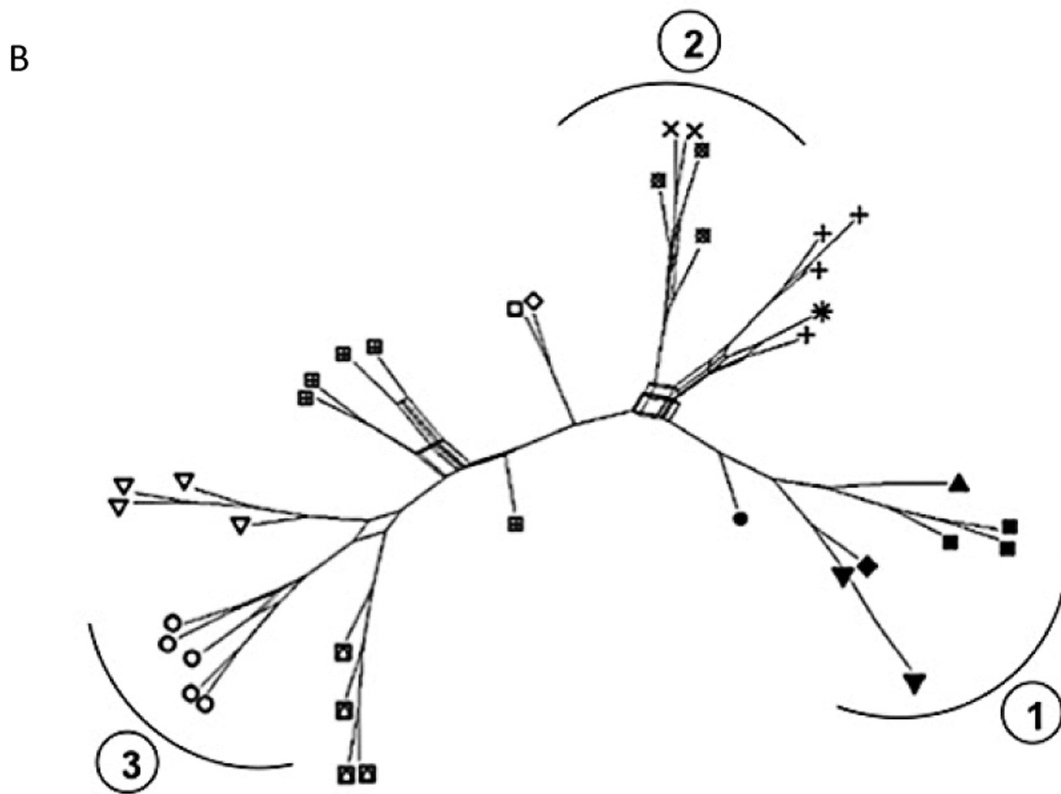
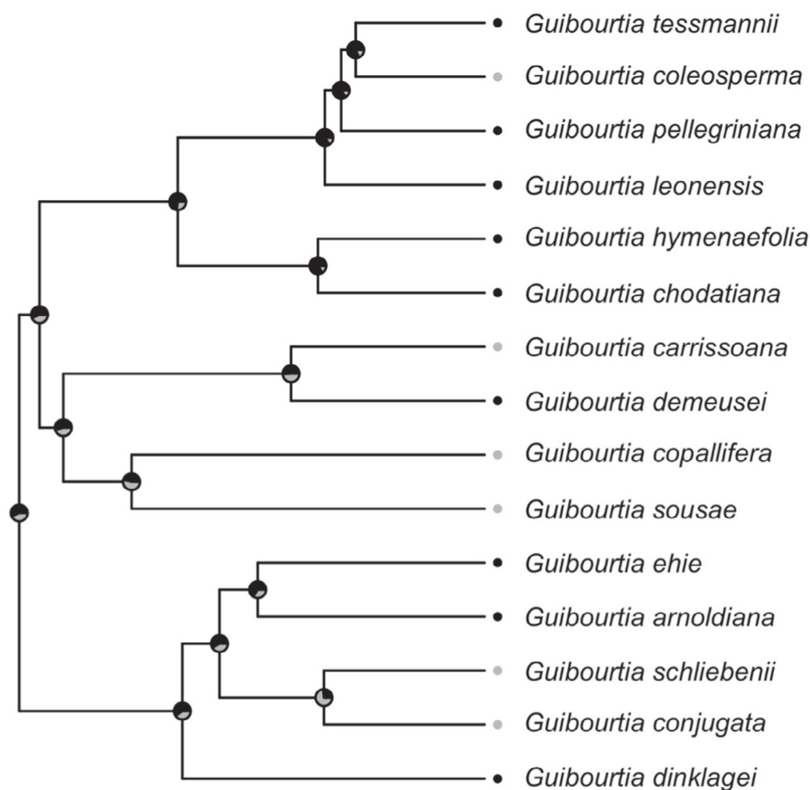


Fig. 3. Distribution map of sequenced samples of *Guibourtia* species (A) and Neighbor-Net tree of whole plastid genomes (B).

G. leonensis, *G. coleosperma*, *G. chodatiana*, and *G. hymenaefolia*. Hence, morphometric groups and phylogenetic clades are highly congruent.

The first two components of the PCoA based on reproductive morphometric characters explained 40.27% and 25.56% of the total variance. The corresponding scatter plot (Fig. 5B) displayed the same three distinct morpho-groups as described above. However, the PCoA



Black: Rain forests (scaled likelihoods at the root: 56.79%)

Grey: Dry forests/savannas (scaled likelihoods at the root: 43.20%)

Fig. 4. Reconstruction of ancestral habitat types using the ancestral state reconstruction method (maximum likelihood) implemented in the R package phytools (Revell, 2012).

scatter plot based on vegetative characters did not separate the above-mentioned morpho-groups (43.19% of the total variance explained by the first two components; Fig. 5C).

Through Hill and Smith's analysis on both all morphometric characters and then reproductive characters, it appears that the first morpho-group can be characterized by the following combination of characters: secondary leaf veins basilar, marginal venation, persistent or sessile stipules, persistent bracteoles, glabrous sepals and non arrilate seeds, no venation on the fruit and non arillate seeds. The second morpho-group was characterized by coriaceous limb, apex of leaflets obtuse, persistent stipule, deciduous bracteole, axillary and terminal inflorescence, globular bud and presence of venation on the fruit. The third group showed the following features: deciduous bracteole, axillary and terminal inflorescence, dehiscent fruit with long and stipitate stipe, no gland on the fruit and arillate seeds.

4. Discussion

4.1. Origin of the Amphi-Atlantic tropical distribution of Guibourtia

The pantropical distribution of 59 families and 334 genera of seed plants is one of the most important plant biogeographic patterns of intercontinental range disjunction (Thorne, 1973). Given that the origin of many families and genera postdates the progressive fragmentation of the Gondwana continent which started *ca.* 180 Ma (McLoughlin, 2001), transcontinental dispersal must have occurred. Three pathways can be seen as alternative hypotheses for explaining the biogeographical connections between Afrotropical and Neotropical floras (Fig. S3): (1) the Boreotropical North Atlantic Land Bridge (NALB) between Europe and North America in the early Tertiary (mainly *ca.* 35–54 Ma), (2) the Bering Land Bridge (BLB or Beringia) between Asia and North America from the early Palaeogene into the late Miocene/early Pliocene, and (3) the direct Atlantic dispersal between Africa and Neotropics. This third pathway was characterized successively by land connection (until *ca.* 96 Ma), potential insular stepping stones (short-distance overseas dispersal via island chains of the Sierra Leone Rise and/or the Rio Grande Rise–Walvis Ridge), and oceanic long-distance dispersal of diaspores by wind or oceanic currents (Renner, 2004; Erkens et al., 2009). Phylogenetic trees, as developed here for the genus *Guibourtia* can thus be used to test the likelihood of these different hypotheses, especially to infer the source and time of entry of taxa into continents (Donoghue and Moore, 2003).

Only two *Guibourtia* forest species are growing in the Neotropics, compared to the 13 African ones, suggesting that Africa is probably the center of origin of this genus. In addition, Neotropical *Guibourtia* are not early branching but included in a clade of otherwise African species. This is supported by our highly resolved plastid phylogeny: a single transatlantic migration event is necessary if the ancestor of all *Guibourtia* occurred in Africa, while three transatlantic migration events in the reverse direction should be assumed if the ancestor of *Guibourtia* occurred in America. Nevertheless, it is interesting to note that *Hymenaea* and *Peltogyne*, the two genera most related to *Guibourtia*, occur in the Neotropics (except *H. verrucosa* which is East-African), suggesting that transoceanic migration was not so exceptional. Among the Neotropical taxa, *G. hymenaefolia* is a deciduous tree mostly found in dry forests in Brazil, Bolivia and Paraguay (Jardim et al., 2003) while the deciduous *G. chodatiana* grows in gallery forests in North and South of America (Cuba, Brazil, etc.; Lorenzi, 1998). Although their taxonomic status has been debated (one species instead of two according to Barneby, 1996), the age of divergence between the two Neotropical

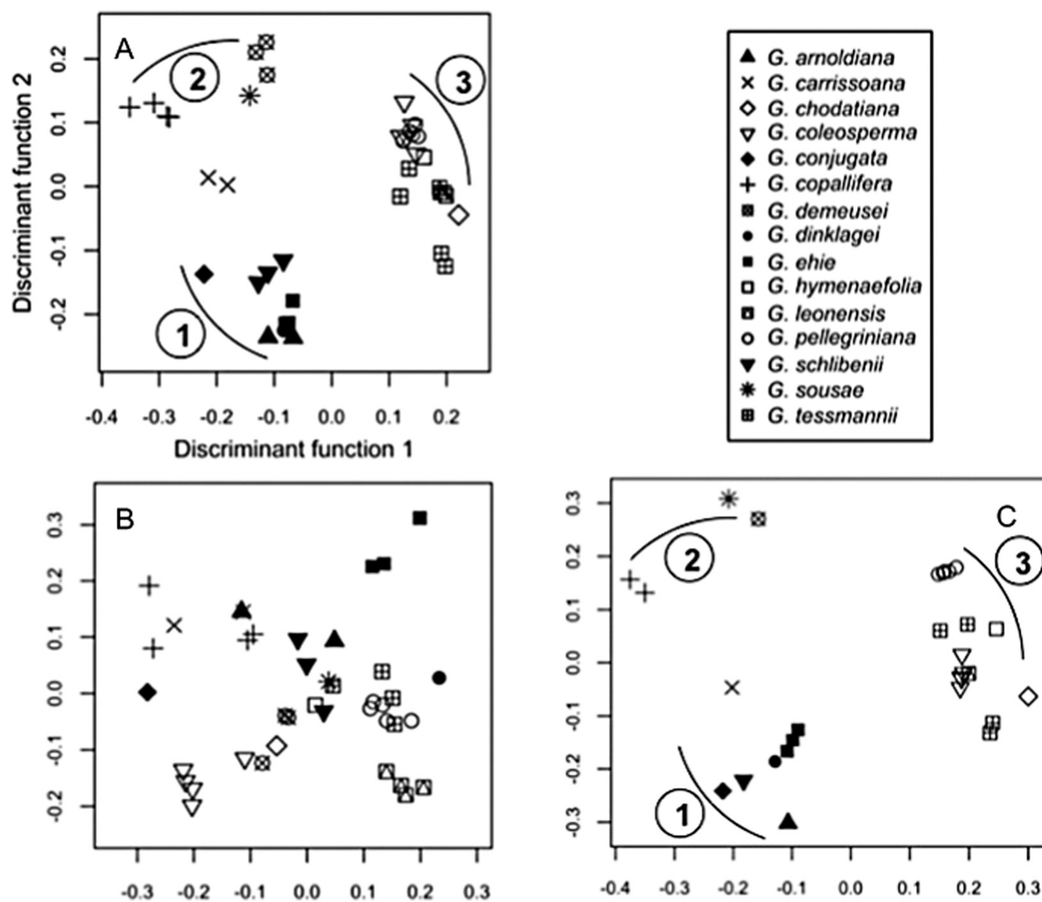


Fig. 5. Morphometric Principal Coordinate Analysis considering all *Guibourtia* species, and either all tested morphological characters (A), or only vegetative characters (B), or only reproductive characters (C). See Table S1 for the list of characters included.

samples was estimated around 6.40 Ma (95% HPD between 1.59 and 11.24 Ma), and the Neotropical clade diverged from the closest African samples around 11.78 Ma (95% HPD between 6.52 and 17.38 Ma), during the Miocene. Hence, assuming a single migration event, the latter must have occurred between these two time estimates, which are far more recent than the age of the Gondwanan breakup (*ca.* 96–105 Ma; Morley, 2003) and of the Boreotropics land bridge (*ca.* 35–54 Ma; Tiffney, 1985; Davis et al., 2002). Regarding the Bering Land Bridge, which connected Asia and North America from the early Palaeogene until its closure between 7.4 and 4.8 Ma (Tiffney and Manchester, 2001), no *Guibourtia* taxa has been described in fossil layers in Eurasia and northern America that could validate this pathway of migration. In addition, the BLB route, which was probably too far north, was less frequently used by tropical taxa than the NALB route, due to the fact that daylight might have been a limiting factor for migration of megathermal plant taxa, such as *Guibourtia* (Tiffney and Manchester, 2001; Morley, 2003).

The most plausible explanation for the Amphi-Atlantic *Guibourtia* distribution appears to be long-distance (oceanic) dispersal from Africa to South America. This is in line with an increasing number of molecular dating studies which reveal that such dispersal events occurred, especially from Africa to South America during the Miocene: Atherospermataceae (Renner et al., 2000), *Symphonia globulifera* (Clusiaceae; Dick et al., 2003), *Gossypium* (Malvaceae; Cronn and Wendel, 2004), Myrtaceae, Vochysiaceae (Sytsma et al., 2004), *Manilkara* (Armstrong et al., 2014), *Barleria oenotheroides* (Acanthaceae; Martín-Bravo and Daniel, 2016) (see the review in Fig. S3). Other cases of transatlantic dispersal were summarized by Renner (2004) and De Queiroz (2005), who also suggested possible mechanisms for these events. Since some seeds or fruits of Detarioideae have good ability to float, Fougère-Danezan (2005) assumed that dispersal by oceanic long-distance dispersal is plausible. In addition, transoceanic dispersal by floating seeds or by seeds associated to large mats of vegetation was estimated to occur in less than two weeks during the Miocene (Houle, 1998). This case exemplifies once again that oceans do not constitute an insurmountable barrier for plants over geological time, and disjunct continental floras are more connected than had been thought previously when vicariance was the main hypothesis for the occurrence of Amphi-Atlantic taxa (De Queiroz, 2005).

4.2. Diversification of Guibourtia in Africa

Our results based on whole plastid genome sequencing highlight that the genus *Guibourtia* is well supported as monophyletic (Fig. S1), as suggested by Fougère-Danezan (2005) who used only seven *Guibourtia* species. Among the 13 African species, eight of the 10 species represented by more than one individual appeared monophyletic for the plastome (Fig. S2). Our plastome phylogeny gives therefore some support to the current species delimitation. It must be noted that incongruence between species delimitation and plastome phylogeny is not uncommon and could result from incomplete lineage sorting and/or chloroplast captures, even if species are correctly delineated (e.g. Duminil et al., 2012; Daïnou et al., 2014; Cavender-Bares et al., 2011; Neophytou et al., 2011; Acosta and Premoli, 2010; Premoli et al., 2012). The globally good congruence in the case of *Guibourtia* suggests that the plastome phylogeny should be a good proxy of the species phylogeny. The two exceptions showing paraphyly (*G. copallifera* and *G. demeusei*) might thus result from incomplete lineage sorting or chloroplast captures. Nevertheless, to delineate more reliably *Guibourtia* species and identify reproductively isolated groups, nuclear markers

would be needed.

The whole-plastid-genome-based chronogram revealed that the genus *Guibourtia* diverged during the Miocene (*ca.* 14.5–24 Ma; node N1, Fig. 2) and continued until Late Miocene and Early Pliocene, as confirmed by Fougère-Danezan (2005). This period coincided with the second major environmental changes that occurred from the Early to the Mid-Miocene (15–23 Ma): the African continent moved northwards, the Equator down positioned, humid vegetation disappeared in the Sahara, and thus the rain forest belt shifted southwards (Maley, 1996; Jacobs, 2004). An increasingly number of genetic studies tends to confirm the crucial role of Miocene events in the African flora composition, especially for: *Inga* (Fabaceae, Mimosoideae; Richardson et al., 2001), *Afromomum* (Zingiberaceae; Auvrey et al., 2010), *Isolona* and *Monodora* (Annonaceae; Couvreur et al., 2011), *Erythrophleum* (Fabaceae, Detarioideae; Duminil et al., 2013), *Manilkara* (Sapotaceae; Armstrong et al., 2014), *Milicia* (Moraceae; Daïnou et al., 2014), the African rattans palms *Ancistrophyllinae* (Arecaceae; Faye et al., 2016b), *Afzelia* (Fabaceae, Detarioideae; Donkpegan et al., 2017). The three main clades identified in our phylogeny (as described below; nodes N2-4, Fig. 2) diverged around the middle Miocene climatic transition between *ca.* 14.8–16 Ma (Flower and Kennett, 1994), when continuous rain forest has split, due to renewed cooling and aridity across Africa (Davis et al., 2002; Couvreur et al., 2008). Especially, this major environmental change may have caused shift in habitats and then species divergence (Losos et al., 1997; Reznick et al., 1997), promoting speciation within many taxa (Smith et al., 1997; Schluter, 1998). Finally, if speciation largely predates the Quaternary, Pleistocene glacial/interglacial climatic oscillations (from 2.7 Ma, and mainly 1.8 Ma) might promote the recent divergence within taxa belonging to the clade 1: *G. coleosperma*, *G. tessmannii*, *G. pellegriniana* and *G. leonensis* (node N5, Fig. 2). It is important to note that these divergence events seem congruent to some major environmental changes but we cannot establish reliable causal links based on the outcomes of the present study.

Inferring ancestral states through the plastid phylogeny, the genus *Guibourtia* is characterized by multiple habitat shifts, with at least four transitions, mainly between rain forests and dry forest/savannahs and at least one occurring during the Pliocene. These multiple transitions differ from the unique transition revealed for Afzelia species, where Quaternary climatic oscillations could have favoured the habitat shift between rain and dry forests combined with the phenomenon of polyploidization (Donkpegan et al., 2017).

4.3. New phylogenetic and morphological insights for systematics of *Guibourtia*

A significant contribution of floral and fruit traits emerged for the morphological delimitation of *Guibourtia* species. This can be explained by some literature statements which assume that reproductive characters evolve faster than vegetative ones in taxa with specialist pollinators, which should be the case for *Guibourtia* species. Plants with generalist pollinators or wind pollination could develop more specific vegetative traits whereas their reproductive characters could remain very similar, as observed in the genus *Milicia* (Daïnou et al., 2014). Besides, Cardoso et al. (2012) highlighted that reproductive characters cannot be used to predict phylogenetic relationships in the early branches of Papilionoid legumes, hypothesizing that reproductive characters might be better at shallow phylogenetic time scale.

Fundamentally, we distinguish three main clades in the phylogeny, which are congruent with the three subgenera described by Léonard (1949) and confirmed by our morphometric analyses (Fig. 5). The clade 1 encompasses *G. tessmannii*, *G. pellegriniana*, *G. leonensis*, *G. coleosperma*, *G. chodatiana* and *G. hymenaefolia*, and was described as the subgenus *Pseudocopaiva* (Britton & Wilson) by Léonard (1949). The clade 2, with *G. copallifera*, *G. demeusei*, *G. carrissoana* and *G. sousae*, corresponds to the subgenus *Guibourtia*, and the clade 3, including *G. arnoldiana*, *G. ehie*, *G. conjugata*, *G. schliebenii* and *G. dinklagei*, belongs to the subgenus *Gorskia* (Bolle).

The subgenus *Pseudocopaiva* is different from the two other ones by the type of fruit and seed (dehiscent fruit and arillate seeds for one and indehiscent fruits and not arillate seeds for others). The difference between the subgenus *Guibourtia* and the subgenus *Gorskia* can be linked to the fundamental difference regarding (i) the persistence of bracteoles (the bracteoles of the first are deciduous while those of the second are persistent) and (ii) the thickness of the seeds (membranous for the first and thick for the second).

5. Conclusions

In this paper, we highlighted the important role of dispersal across the Atlantic Ocean, which promoted the diversification of *Guibourtia* into the Neotropics. The majority of species appeared monophyletic and the crown age of the plastomes sampled within each monophyletic species was usually anchored in the Pliocene or early Pleistocene (2.6–6 Ma). Besides, the divergence between the plastomes of sister monophyletic species was usually dated in the late Miocene or Pliocene (5–13 Ma). Finally, in each clade, morphologically supported by reproductive characters, we can assume at least one habitat transition, mainly from rain forest to dry forest/savannah.

The next step, under progress, will be to use the genus *Guibourtia* as a model to address questions about niche conservatism and convergent evolution. Targeting the ecological amplitude of African *Guibourtia*'s taxa, we could now rely on this new highly resolved plastid DNA phylogeny to examine how far ecological niche influences organism morphology among related species.

Acknowledgements

This work was financially supported by the Belgian “Fonds pour la Recherche Scientifique” (F.R.S.-FNRS) through the “Fonds pour la Formation à la Recherche dans l'Industrie et l'Agriculture” (FRIA PhD grant attributed to Félicien Tosso) and the FRS-FNRS PDR project T.0163.13, by the Belgian Science Policy (Belspo) through the BRAIN research project AFRIFORD (postdoctoral grant to Jérémy Migliore), and by FFEM-AFD through the DynAfFor project. We thank the National Museum of Natural History in Paris (France) and the Botanic Garden Meise (Belgium), for allowing us to sample and study their herbarium collections. We are thankful to partners that offered technical and logistical support during field collection: Wijma Cameroun S.A., Precious Woods Gabon and Nature+. We sincerely appreciate the help of Laurent Grumiau (ULB-EBE, Belgium), Cédric Mariac (IRD-DIADE, Montpellier, France), Adriana Anzil and Pierre Servais (ULB, Belgium), Latifa Karim and Wouter Coppieters (GIGA Liège, Belgium), Jérôme Duminil (IRD, Montpellier, France), Charles Doumenge and Honore Agbazahou (CIRAD, Montpellier). Finally, we thank the two anonymous reviewers for their very constructive comments which helped improving this manuscript.

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at https://doi.org/10.1016/j.ympev.2017.11.026.

References

Acosta, M.C., Premoli, A.C., 2010. Evidence of chloroplast capture in South American Nothofagus (subgenus *Nothofagus*, Nothofagaceae). Mol. Phylogenet. Evol. 54, 235–242.

Armstrong, K., Stone, G., Nicholls, J., Valderrama Escallon, E., Anderberg, A., Smedmark, J., Gautier, L., Naciri, Y., Milne, R., Richardson, J.E., 2014. Patterns of diversification amongst tropical regions compared: a case study in Sapotaceae. Front Genet. 5.

Aubréville, A., 1970. Légumineuses. Césalpinioidées. Flore du Cameroun. Museum National d'Histoire Naturelle, Paris.

Baker, W.J., Couvreur, T.L.P., 2013. Global biogeography and diversification of palms sheds light on the evolution of tropical lineages. I. Historical biogeography. J.

Biogeogr. 40, 274–285.
Barneby, R.C., 1996. Neotropical Fabales at NY: asides and oversights. Brittonia 48, 174–187.
Bruneau, A., Mercure, M., Lewis, G.P., Herendeen, P.S., 2008. Phylogenetic patterns and diversification in the caesalpinioid legumes This paper is one of a selection of papers published in the Special Issue on systematics research. Botany 86, 697–718.
Bryant, D., Moulton, V., 2004. Neighbor-net: an agglomerative method for the construction of phylogenetic networks. Mol. Biol. Evol. 21 (2), 255–265.
Cattell, R., 1966. The scree test for the number of factors. MBR 1, 245–276.
Cardoso, D., De Queiroz, L.P., Pennington, R.T., De Lima, H.C., Fonty, É., Wojciechowski, M.F., Lavin, M., 2012. Revisiting the phylogeny of papilionoid legumes: new insights from comprehensively sampled early-branching lineages. Am. J. Bot. 99 (12), 1991–2013.
Cavender-Bares, J., Gonzalez-Rodriguez, A., Pahlich, A., Koehler, K., Deacon, N., 2011. Phylogeography and climatic niche evolution in live oaks (*Quercus* series Virentes) from the tropics to the temperate zone. J. Biogeogr. 38, 962–981.
Chessel, D., Dufour, A., Dray, S., Jombart, T., Lobry, J., Ollier, S., Thioulouse, J., 2012. ade4: Analysis of Ecological Data: Exploratory and Euclidean Methods in Environmental Sciences. R Package Version 1.5-0.
Chevreux, B., Wetter, T., Suhai, S., 1999. Genome sequence assembly using trace signals and additional sequence information. In: German Conference on Bioinformatics, Heidelberg, pp. 45–56.
Christenhusz, M.J.M., Chase, M.W., 2013. Biogeographical patterns of plants in the Neotropics – dispersal rather than plate tectonics is most explanatory. Bot. J. Linn. Soc. 171, 277–286.
Conti, E., Eriksson, T., Schonenberger, J., Sytsma, K.J., Baum, D.A., Orti, G., 2002. Early Tertiary out-of-India dispersal of the Crypteroniaceae: evidence from phylogeny and molecular dating. Evolution 56, 1931–1942.
Couvreur, T.L.P., Chatrou, L.W., Sosef, M.S., Richardson, J.E., 2008. Molecular phylogenetics reveal multiple tertiary vicariance origins of the African rain forest trees. BMC Biol. 6 (1), 54.
Couvreur, T.L.P., Porter-Morgan, H., Wieringa, J.J., Chatrou, L.W., 2011a. Little ecological divergence associated with speciation in two African rain forest tree genera. BMC Evol. Biol. 11 (1), 296.
Couvreur, T.L.P., Pirie, M.D., Chatrou, L.W., Saunders, R.M.K., Su, Y.C.F., Richardson, J.E., Erkens, R.H.J., 2011b. Early evolutionary history of the flowering plant family Annonaceae: steady diversification and boreotropical geodispersal. J. Biogeogr. 38 (4), 664–680.
Couvreur, T.L.P., 2015. Odd man out: why are there fewer plant species in African rain forests? Plant Syst. Evol. 301 (5), 1299–1313.
Cronn, R., Wendel, J.F., 2004. Cryptic trysts, genomic mergers, and plant speciation. New Phytol. 161, 133–142.
Daïnou, K., Mahy, G., Duminil, J., Dick, C., Doucet, J.-L., Donkpégan, A., Pluijgers, M., Sinsin, B., Lejeune, P., Hardy, O.J., 2014. Speciation slowing down in widespread and long-living tree taxa: insights from the tropical timber tree genus *Milicia* (Moraceae). Heredity 113, 74.
Daïnou, K., Blanc-Jolivet, C., Degen, B., Kimani, P., Ndiade-Bourobou, D., Donkpegan, A.S., Tosso, F., Kaymak, E., Bourland, N., Doucet, J.-L., 2016. Revealing hidden species diversity in closely related species using nuclear SNPs, SSRs and DNA sequences–a case study in the tree genus *Milicia*. BMC Evol. Biol. 16, 259.
Darriba, D., Taboada, G.L., Doallo, R., Posada, D., 2012. jModelTest 2: more models, new heuristics and parallel computing. Nat. Methods. 9 (8) 772-772.
Dauby, G., Duminil, J., Heuertz, M., Koffi, G., Stevart, T., Hardy, O.J., 2014. Congruent phylogeographical patterns of eight tree species in Atlantic Central Africa provide insights into the past dynamics of forest cover. Mol. Ecol. 23, 2299–2312.
Davis, C.C., Bell, C.D., Mathews, S., Donoghue, M.J., 2002. Laurasian migration explains Gondwanan disjunctions: evidence from Malpighiaceae. Proc. Nat. Acad. Sci. 99, 6833–6837.
De Queiroz, A., 2005. The resurrection of oceanic dispersal in historical biogeography. Trends Ecol. Evol. 20, 68–73.
Dick, C.W., Abdul-Salim, K., Bermingham, E., 2003. Molecular systematic analysis reveals cryptic tertiary diversification of a widespread tropical rain forest tree. Am. Nat. 162, 691–703.
Donkpegan, A.S., Doucet, J.L., Migliore, J., Duminil, J., Dainou, K., Piñeiro, R., Wieringa, J.J., Champluvier, D., Hardy, O.J., 2017. Evolution in African tropical trees displaying ploidy-habitat association: the genus *Afzelia* (Leguminosae). Mol. Phylogenet. Evol. 107, 270–281.
Donoghue, M.J., Moore, B.R., 2003. Toward an integrative historical biogeography. Integr. Comp. Biol. 43, 261–270.
Doyle, J.J., 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochem. Bull. 19, 11–15.
Drummond, A.J., Rambaut, A., 2007. BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evol. Biol. 7, 1.
Duminil, J., Brown, R.P., Ewédjè, E.-E.B., Mardulyn, P., Doucet, J.-L., Hardy, O.J., 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.
Duminil, J., Kenfack, D., Viscosi, V., Grumiau, L., Hardy, O.J., 2012. Testing species delimitation in sympatric species complexes: the case of an African tropical tree, Carapa spp. (Meliaceae). Mol. Phylogenet. Evol. 62 (1), 275–285.
Duminil, J., Mona, S., Mardulyn, P., Doumenge, C., Walmacq, F., Doucet, J.L., Hardy, O.J., 2015. Late Pleistocene molecular dating of past population fragmentation and demographic changes in African rain forest tree species supports the forest refuge hypothesis. J. Biogeogr. 42, 1443–1454.
Erkens, R.H., Maas, J., Couvreur, T.L.P., 2009. From Africa via Europe to South America: migrational route of a species-rich genus of Neotropical lowland rain forest trees (*Guatteria*, Annonaceae). J. Biogeogr. 36, 2338–2352.
Faye, A., Deblauwe, V., Mariac, C., Damien, R., Sonké, B., Vigouroux, Y., Couvreur, T.L.P., 2016a. Phylogeography of the genus *Podococcus* (Palmae/Arecaceae) in Central African rain forests: climate stability predicts unique genetic diversity. Mol. Phylogenet. Evol. 105, 126–138.
Faye, A., Pintaud, J.C., Baker, W.J., Vigouroux, Y., Sonke, B., Couvreur, T.L.P., 2016b. Phylogenetics and diversification history of African rattans (Calamoideae, Ancistrophyllinae). Bot. J. Linn. Soc. 182 (2), 256–271.
Flower, B.P., Kennett, J.P., 1994. The middle Miocene climatic transition: East Antarctic ice sheet development, deep ocean circulation and global carbon cycling. Palaeogeogr. Palaeoclimatol. Palaeoecol. 108, 537–555.
Fougère-Danezan, M., 2005. *Phylogénie moléculaire et morphologique des Detarieae résinifères (Leguminosae, Caesalpinioideae): contribution à l'étude de l'histoire biogéographique des légumineuses*. Toulouse 3.
Fougère-Danezan, M., Herendeen, P.S., Maumont, S., Bruneau, A., 2010. Morphological evolution in the variable resin-producing Detarieae (Fabaceae): do morphological characters retain a phylogenetic signal? Ann. Bot. 105, 311–325.
Gower, J.C., 1971. A general coefficient of similarity and some of its properties. Biometrics 857–871.
Hahn, C., Bachmann, L., Chevreux, B., 2013. Reconstructing mitochondrial genomes directly from genomic next-generation sequencing reads—a baiting and iterative mapping approach. Nucl. Acid Res. 41 (13) e129-e129.
Hardy, O.J., Born, C., Budde, K., Daïnou, K., Dauby, G., Duminil, J., Ewédjé, E.-E.B., Gomez, C., Heuertz, M., Koffi, G.K., 2013. Comparative phylogeography of African rain forest trees: a review of genetic signatures of vegetation history in the Guineo-Congolian region. CR Geosci. 345, 284–296.
Hill, M., Smith, A., 1976. Principal component analysis of taxonomic data with multi-state discrete characters. Taxon 249–255.
Houle, A., 1998. Floating islands: a mode of long distance dispersal for small and medium-sized terrestrial vertebrates. Divers. Distributions 4, 201–216.
Huson, D.H., Bryant, D., 2006. Application of phylogenetic networks in evolutionary studies. Mol. Biol. Evol. 23, 254–267.
Jacobs, B.F., 2004. Palaeobotanical studies from tropical Africa: relevance to the evolution of forest, woodland and savannah biomes. Philos. Trans. Roy. Soc. B. 359, 1573–1583.
Jardim, A., Killeen, T.J., Fuentes, A., 2003. Guia de los arboles y arbustos del Bosque Seco Chiquitano, Bolívia. Santa Cruz de la Sierra, Bolivia: Editorial FAN ix, 324p.-illus., col. illus.. ISBN, 1400631683.
Kaiser, H.F., 1960. The application of electronic computers to factor analysis. Educ. Psychol. Meas. 20 (1), 141–151.
Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., Buxton, S., Cooper, A., Markowitz, S., Duran, C., Thierer, T., Ashton, B., Meintjes, P., Drummond, A., 2012. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics 28 (12), 1647–1649.
Kiers, H.A., 1991. Simple structure in component analysis techniques for mixtures of qualitative and quantitative variables. Psychometrika 56, 197–212.
Koboldt, D.C., Zhang, Q., Larson, D.E., Shen, D., Mclellan, M.D., Lin, L., Miller, C.A., Mardis, E.R., Ding, L., Wilson, R.K., 2012. VarScan 2: somatic mutation and copy number alteration discovery in cancer by exome sequencing. Genome Res. 22, 568–576.
Laliberté, E., Legendre, P., Shipley, B., Laliberté, M.E., 2010. Package 'FD'. Measuring functional diversity (FD) from multiple traits, and other tools for functional ecology. Ecology 91 (1), 299–305.
Léonard, J., 1949. Notulae systematicae IV (Caesalpiniaceae-Amherstieae africanae americanaeque). Bull. Jard. Bot. État Bruxelles. 19, 383–408.
Li, H., Durbin, R., 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25, 1754–1760.
Lorenzi, H., 1998. Árvores brasileiras: manual de identificação e cultivo de plantas arbóreas nativas do Brasil. Nova Odessa: Editora Plantarum 352p.-col. illus.. Por Geog, 4.
Losos, J.B., Warheit, K., Schoener, T., 1997. Adaptive differentiation following experimental island colonization inAnolis lizards. Nature 387, 70–73.
Martin, M., 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet. J. 17 (1), 1–10.
Malé, P.J.G., Bardon, L., Besnard, G., Coissac, E., Delsuc, F., Engel, J., Lhuillier, E., Scotti-Saintagne, C., Tinaut, A., Chave, J., 2014. Genome skimming by shotgun sequencing helps resolve the phylogeny of a pantropical tree family. Mol. Ecol. Resour. 14, 966–975.
Maley, J., 1996. The African rain forest–main characteristics of changes in vegetation and climate from the Upper Cretaceous to the Quaternary. Proc. Royal Soc Edinburgh. Section B. Biol. Sci. 104, 31–73.
Mariac, C., Scarcelli, N., Pouzadou, J., Barnaud, A., Billot, C., Faye, A., Kougbeadjo, A., Maillol, V., Martin, G., Sabot, F., 2014. Cost-effective enrichment hybridization capture of chloroplast genomes at deep multiplexing levels for population genetics and phylogeography studies. Mol. Ecol. Resour. 14, 1103–1113.
Martín-Bravo, S., Daniel, T.F., 2016. Molecular evidence supports ancient long-distance dispersal for the amphi-Atlantic disjunction in the giant yellow shrimp plant (*Barleria oenotheroides*). Am. J. Bot. 103, 1103–1116.
Mcloughlin, S., 2001. The breakup history of Gondwana and its impact on pre-Cenozoic floristic provincialism. Aust. J. Bot. 49, 271–300.
Mcpherson, H., Van Der Merwe, M., Delaney, S.K., Edwards, M.A., Henry, R.J., Mcintosh, E., Rymer, P.D., Milner, M.L., Siow, J., Rossetto, M., 2013. Capturing chloroplast variation for molecular ecology studies: a simple next generation sequencing approach applied to a rainforest tree. BMC Ecol. 13, 8.
Morley, R.J., 2003. Interplate dispersal paths for megathermal angiosperms. Perspect. Plant Ecol. Evol. Syst. 6, 5–20.

Muellner, A.N., Savolainen, V., Samuel, R., Chase, M.W., 2006. The mahogany family "out-of-Africa": divergence time estimation, global biogeographic patterns inferred from plastid rbcL DNA sequences, extant, and fossil distribution of diversity. Mol. Phylogenet. Evol. 40, 236–250.
Neophytou, C., Dounavi, A., Fink, S., Aravanopoulos, F.A., 2011. Interfertile oaks in an island environment: I. High nuclear genetic differentiation and high degree of chloroplast DNA sharing between *Q. alnifolia* and *Q. coccifera* in Cyprus. A multi-population study. Eur. J. For. Res. 130, 543–555.
Peñalba, J.V., Smith, L.L., Tonione, M.A., Sass, C., Hykin, S.M., Skipwith, P.L., McGuire, J.A., Bowie, R.C.K., Moritz, C., 2014. Sequence capture using PCR - generated probes: a cost - effective method of targeted high - throughput sequencing for nonmodel organisms. Mol. Ecol. Resour. 14 (5), 1000–1010.
Pennington, R.T., Richardson, J.E., Lavin, M., 2006. Insights into the historical construction of species-rich biomes from dated plant phylogenies, phylogenetic community structure and neutral ecological theory. New Phytol. 172, 605–616.
Piñeiro, R., Dauby, G., Kaymak, E., Hardy, O.J., 2017. Pleistocene population expansions of shade-tolerant trees indicate fragmentation of the African rainforest during the Ice Ages. In: Proc. R. Soc. B., vol. 284, 1866. doi:10.1098/rspb.2017.1800.
Plana, V., 2004. Mechanisms and tempo of evolution in the African Guineo-Congolian rainforest. Philos Trans. Roy. Soc. B-Biol. Sci. 359, 1585–1594.
Premoli, A.C., Mathiasen, P., Cristina Acosta, M., Ramos, V.A., 2012. Phylogeographically concordant chloroplast DNA divergence in sympatric *Nothofagus ss* How deep can it be? New Phytol. 193, 261–275.
Raven, H.P., Axelrod, D.I., 1974. Angiosperm biodiversity and past continental movements. Ann. Mo. Bot. Gard. 61, 539–673.
Renner, S.S., Clausing, G., Meyer, K., 2001. Historical biogeography of Melastomataceae: the roles of Tertiary migration and long-distance dispersal. Am. J. Bot. 88 (7), 1290–1300.
Renner, S., 2004. Plant dispersal across the tropical Atlantic by wind and sea currents. Int. J. Plant Sci. 165, S23–S33.
Renner, S., Foreman, D., Murray, D., 2000. Timing transantarctic disjunctions in the *Atherospermataceae* (Laurales): evidence from coding and noncoding chloroplast sequences. Syst. Biol. 49, 579–591.
Revell, L.J., 2012. phytools: an R package for phylogenetic comparative biology (and other things). Methods Ecol. Evol. 3 (2), 217–223.
Reznick, D.N., Shaw, F.H., Rodd, F.H., Shaw, R.G., 1997. Evaluation of the rate of evolution in natural populations of guppies (*Poecilia reticulata*). Science 275, 1934–1937.
Richardson, J.E., Weitz, F.M., Fay, M.F., Cronk, Q.C., Linder, H.P., Reeves, G., Chase, M.W., 2001. Rapid and recent origin of species richness in the Cape flora of South Africa. Nature 412 (6843), 181–183.
Richardson, J.E., Whitlock, B.A., Meerow, A.W., Madriñán, S., 2015. The age of chocolate: a diversification history of *Theobroma* and Malvaceae. Front. Ecol. Evol. 3, 120.
Rohland, N., Reich, D., 2012. Cost-effective, high-throughput DNA sequencing libraries for multiplexed target capture. Genome Res. 22, 939–946.
Scarcelli, N., Mariac, C., Couvreur, T.L.P., Faye, A., Richard, D., Sabot, F., Berthouly-Salazar, C., Vigouroux, Y., 2016. Intra - individual polymorphism in chloroplasts from NGS data: where does it come from and how to handle it? Mol. Ecol. Resour. 16 (2), 434–445.
Schluter, D., 1998. Ecological causes of speciation. Endless forms: species and speciation, 114-129.
Slik, J.F., Arroyo-Rodríguez, V., Aiba, S.-I., Alvarez-Loayza, P., Alves, L.F., Ashton, P., Balvanera, P., Bastian, M.L., Bellingham, P.J., Van Den Berg, E., 2015. An estimate of the number of tropical tree species. Proc. Natl. Acad. Sci. 112, 7472–7477.
Smith, T.B., Wayne, R.K., Girman, D.J., Bruford, M.W., 1997. A role for ecotones in generating rainforest biodiversity. Science 276, 1855–1857.
Sytsma, K.J., Litt, A., Zjhra, M.L., Pires, J.C., Nepokroeff, M., Conti, E., Walker, J., Wilson, P.G., 2004. Clades, clocks, and continents: historical and biogeographical analysis of Myrtaceae, Vochysiaceae, and relatives in the Southern Hemisphere. Int. J. Plant Sci. 165, S85–S105.
Thorne, R.F., 1973. Floristic relationships between tropical Africa and tropical America. In: Meggers, B.J., Ayensu, E.S., Duckworth, W.D. (Eds.), Tropical Forest Ecosystems in Africa and South America: A Comparative Review. Smithsonian Instn. Press, Washington, pp. 27–47.
Tiffney, B.H., 1985. The Eocene North Atlantic land bridge: its importance in Tertiary and modern phytogeography of the Northern Hemisphere. J. Arnold Arbor. 66 (2), 243–273.
Tiffney, B.H., Manchester, S.R., 2001. The use of geological and paleontological evidence in evaluating plant phylogeographic hypotheses in the Northern Hemisphere Tertiary. Int. J. Plant Sci. 162, S3–S17.
Tosso, F., Daïnou, K., Hardy, O.J., Sinsin, B., Doucet, J.-L., 2015. Le genre *Guibourtia* Benn., un taxon à haute valeur commerciale et sociétale (synthèse bibliographique). Biotechnol. Agron. Soc. Environ. 19, 71–88.
Uribe-Convers, S., Duke, J.R., Moore, M.J., Tank, D.C., 2014. A long PCR-based approach for DNA enrichment prior to next-generation sequencing for systematic studies. Appl. Plant Sci. 2, 1300063.
Williams, A.V., Miller, J.T., Small, I., Nevill, P.G., Boykin, L.M., 2016. Integration of complete chloroplast genome sequences with small amplicon datasets improves phylogenetic resolution in *Acacia*. Mol. Phylogenet. Evol. 96, 1–8.
Zerega, N.J.C., Clement, W.L., Datwyler, S.H., Weiblen, G.D., 2005. Biogeography and divergence times in the mulberry family (Moraceae). Mol. Phylogenet. Evol. 37, 402–416.