A combined analysis of morphological traits, chloroplast and nuclear DNA sequences within *Santiria trimera* (Burseraceae) suggests several species following the Biological Species Concept

K. Guillaume Koffi[1,*], Myriam Heuertz[1,2], Charles Doumenge[3], Jean Michel Onana[4], Frederick Gavory[5] & Olivier J. Hardy[1]

[1]Université Libre de Bruxelles, Faculté des Sciences, Evolutionary Biology and Ecology, CP160/12, Av. F. D. Roosevelt 50, BE-1050 Brussels, Belgium
[2]Centre of Forest Research, Dept. Forest Genetics and Ecology, Carretera de La Coruña km 7.5, ES-28040 Madrid, Spain
[3]Cirad, Campus International de Baillarguet, TA C-105/D, FR-34398 Montpellier cedex 5, France
[4]Herbier National du Cameroun, B.P. 1601, Yaoundé, Cameroon
[5]Genoscope - Centre National de Séquençage 2 rue Gaston Crémieux CP5706, FR-91057 Evry cedex, France
*Author for correspondence: kokoffi@ulb.ac.be

Background and aims – Species delimitation remains a difficult task in many groups of organisms. Even widespread and conspicuous tropical trees make no exception. *Santiria trimera*, an African rainforest tree, displays substantial morphological variability. While classical analysis of historical herbarium samples lead to the recognition of a single species, two morphotypes of *S. trimera* are regularly collected in Gabon and Equatorial Guinea, sometimes in sympatry: one form with stilt roots (SR) and another form without stilt roots (NSR).

Methods – To assess whether these forms constitute distinct taxa and to understand evolutionary processes within African *Santiria*, we combined a morphological and a spatial analysis of both sympatric morphotypes in northern Gabon with a phylogenetic analysis of samples from West and Central Africa, including the island of São Tomé.

Key results – In Gabon, morphological traits, chloroplast DNA (cpDNA) and nuclear DNA (nuDNA) sequences congruently lead to the recognition of two well differentiated and reproductively isolated taxa corresponding to SR and NSR morphotypes. However, rare cases of the SR morphotype bearing DNA sequences typical of NSR individuals suggest that hybridization might sometimes occur. No evidence for habitat specialization of the two taxa was found. DNA sequences from São Tomé formed monophyletic clades at both marker types and the nuDNA gene tree suggests that the São Tomé population probably originated from Central Africa and subsequently underwent allopatric differentiation from continental populations.

Conclusion – In Central Africa, *S. trimera* is composed of at least two distinct sympatric species following the Biological Species Concept and calls for further morphological and phenological studies and experimental crosses between them to specify their taxonomic status. Our results also show that classical taxonomic species delimitation may not always be congruent with the Biological Species Concept, calling for more population-based in situ morphological and molecular genetic analyses.

Key words – *Santiria*, Burseraceae, phylogeny, speciation, morphotype, rainforest.

INTRODUCTION

Species delimitation remains a difficult task in many groups of organisms (Ellis et al. 2006), notably in species rich ecosystems like tropical rainforests, for at least two reasons. First, there is not a single definition of the term "species" (Lee 2003, Dayrat 2005). Taxonomists generally define species as a group of individuals characterized by distinctive morphological features, often the only information available. Sometimes, this leads to arbitrary choices and it does not ensure that species delimitation necessarily matches the Biological Species Concept (BSC), a group of individuals

able to mate freely. Second, concerning some taxonomic groups, there is insufficient research in some tropical ecosystems, such as African rainforests, to cover their remarkable diversity. For example, scientists continue to discover new species in some well-known groups such as Leguminosae, Orchidaceae and Rubiaceae (Sonké 2000, Stévart & Cribb 2004, Sonké et al. 2005, Breteler 2010, Lachenaud & Séné 2010). Even widespread and conspicuous tropical trees make no exception. However, molecular and population-based data can substantially help species delimitation approaching the BSC. For instance, new species were identified within the widespread African tropical trees traditionally described as *Carapa procera* DC. (Kenfack 2008) and *Greenwayodendron suaveolens* (Engl. & Diels) Verdc. (Dauby et al. 2010). In these examples, the fact that distinct morphotypes were occurring in sympatry raises the "species problem" (de Queiroz 2005): (1) what are the necessary properties of the species? (2) Where does the limit lie between close and sympatric taxa? Or (3) what is the necessary degree of divergence to delineate species when adult trees are clearly morphologically different? To answer these questions, Dayrat (2005) and Sei & Porter (2007) proposed the combination of various methods, including phylogeography, comparative morphology and population genetics. Some methods based on the phylogeny of chloroplast or nuclear DNA have allowed to clarify taxonomy within complexes of plant species (Murakami et al. 1998) or genera (Weeks et al. 2005, Miikeda et al. 2006).

We focused our attention on trees from the genus *Santiria* (Burseraceae), which displays substantial morphological variability in Africa. Nine African species of *Santiria* have been described in the literature although in a recent revision, Onana (2009) recognized only one species, *S. trimera* (Oliv.) H.J.Lam ex Aubrév. emend. Onana. However, two distinct morphotypes of *S. trimera* are regularly observed in sympatry in Central African forests: one form with stilt roots (hereafter denoted as SR) and another form without stilt roots (hereafter denoted as NSR; Hladik & Hallé 1973, Florence & Hladik 1980, Reitsma 1988, and personal observations). To add to the confusion, Letouzey (1985) and Senterre (2005) mentioned another morphotype in submontane forests of Cameroon and Equatorial Guinea. There is also a third form (hereafter denoted AR) on São Tomé Island with small aerial roots. A phylogeographic study of *S. trimera* based on cpDNA sequences revealed a high divergence among sympatric SR and NSR morphotypes in Gabonese and Equatoguinean forests and among allopatric AR and SR morphotypes (Koffi et al. in prep.). Therefore, to assess whether SR and NSR morphotypes occurring in sympatry should be considered as distinct taxa, we here integrate data on morphological traits of SR and NSR individuals from Gabonese populations with molecular data from nuclear DNA (nuDNA) and chloroplast DNA (cpDNA) from samples collected through West and Central African forests, including the island of São Tomé.

The objectives of this study are (1) to assess whether sympatric SR and NSR morphotypes occupy different habitats; (2) to identify the main morphological traits which differentiate SR and NSR morphotypes in Gabon; (3) to infer phylogenetic relationships between SR and NSR morphotypes from Central Africa as well as their phylogenetic relationships with populations of *S. trimera* found in West Africa and on São Tomé Island.

This work should guide further taxonomic work to delineate taxonomic groups within African *Santiria*.

MATERIALS AND METHODS

Biological model

Santiria trimera is a medium sized African tree belonging to the tribe Canarieae (Burseraceae). The shape of the tree is remarkably variable. Rarely, the base of the tree is cylindrical. Some trees have well-developed stilt roots or simply bear aerial roots on the trunk and some have buttresses (electronic appendix 1A & B, Aubréville 1948, 1962, Hladik & Hallé 1973, Florence & Hladik 1980, Reitsma 1988). The species is also variable for other vegetative traits (leaves) and reproductive traits (inflorescences, fruits). *S. trimera* is dioecious and therefore allogamous, with similar male and female inflorescences (Onana 2009). Immature drupes are brightly coloured (green, yellowish, yellow or red; electronic appendix 1A & 2). *S. trimera* is common in the undergrowth of African rainforests from Sierra Leone to the D.R.Congo and on São Tomé Island (Aubréville 1948).

According to our field observations and descriptions found in floras, the morphotype with well-developed stilt roots (SR) is common in West and Central Africa while the morphotype without stilt roots (NSR) is limited to Central Africa, more precisely Lower Guinea (Cameroon, Republic of the Congo, Equatorial Guinea, Gabon and D.R.Congo), where both morphotypes can occur in sympatry. On São Tomé Island, *Santiria* trees differ from continental morphotypes by bearing small aerial roots on their trunks (AR morphotype, electronic appendix 1B), and by containing a very fluid resin that burns like the one of *Aucoumea klaineana* Pierre (Aubréville 1948).

Morphological and spatial analyses

Morphological traits were observed during the fruiting season from October to December 2007 in four locations in northern Gabon where both SR and NSR morphotypes occur in sympatry: Makokou, Ipassa station (0°28'44"N 12°47'42"E), Belinga Mountains (1°7'12"N 13°12'22"E),Cristal Mountains (0°28'23"N 10°16'19"E) and Kongou (0°17'20"N 12°34'12"E). Morphological data and tissue samples were collected on adult trees (diameter at breast height, dbh > 15 cm) along six transects (table 1). We also collected fertile herbarium vouchers mainly at Makokou and Kongou (electronic appendix 3) and deposited them at the Herbarium of the Université Libre de Bruxelles (BRLU). Geographic coordinates of trees along transects were registered by Global Positioning System (GPS) to assess whether SR and NSR morphotypes segregate spatially. A map of the first transect in Makokou was built and the spatial autocorrelation coefficient (Moran's I statistic; Moran 1951) was calculated on the morphotype status for different distance intervals to test spatial segregation between both morphotypes.

A detailed inventory at Makokou, Ipassa station (1983–1984), of all individuals with a height ≥ 1 m, over a 2.5 ha plot (100 × 250 m), delivered additional demographic, mor-

Table 1 – Characteristics of the six transects where both morphotypes of *Santiria trimera* were studied.
Mt = Mountains; SR = trees with stilt roots; NSR = trees without stilt roots.

Transect	Geographic coordinates	Elevation range (m)	Length (m)	Samples size
Belinga Mt 1	1°7'30"N 13°12'22"E	232–932	3182	16 SR + 5 NSR
Belinga Mt 2	1°7'30"N 13°12'22"E	559–1024	2033	11 SR + 6 NSR
Makokou 1	0°28'44"N 12°47'42"E	474–608	2396	76 SR + 40 NSR
Makokou 2	0°28'44"N 12°47'42"E	464–520	2393	34 SR + 18 NSR
Cristal Mt	0°28'22"N 10°16'19"E	136–345	1711	12 SR + 9 NSR
Kongou	0°17'20"N 12°34'12"E	465–520	747	22 SR + 8 NSR

phological and ecological data. In this area, contrarily to most other places, the NSR morphotype is dominant compared to the SR one.

Quantitative trait data were collected from fertile adult trees: eleven SR and nine NSR individuals in Makokou (electronic appendix 3). Selected trees carried fruits or at least inflorescences of aborted fruits and were therefore necessarily female. Seven to ten measurements per tree (on flowering branches) were performed for each of the following traits: terminal leaflet length (TLL), terminal leaflet width (TLW), petiole length (PL), drupe length (DL), drupe width (DW), number of pairs of leaflets (NLP). Differences between morphotypes for the average value per tree of each quantitative trait were tested using a Student's *t-test* or a Wilcoxon test (non-parametric test equivalent to the *t-test*). We applied a principal component analysis (PCA) to ordinate individuals along two axes. We also recorded some qualitative traits additional to the presence/absence of stilt roots (i.e. SR or NSR morphotype), namely the position of the inflorescence on the twig (axial or terminal), the strength and colour of the inflorescence rachis, the colour of immature fruits and the intensity of the smell of the resin of 17 SR and 17 NSR individuals in three locations (Cristal Mountains, Kongou and Makokou).

Sample collection, DNA extraction and sequencing

Leaflet or cambium tissue were collected in 42 locations in Cameroon, the Central African Republic, Equatorial Guinea, Gabon, Guinea, Ivory Coast, and on São Tomé island (electronic appendix 4). This sampling was originally designed for a cpDNA based phylogeographic analysis of *S. trimera* (Koffi et al. in prep.). In Equatorial Guinea and Gabon, eighteen locations were sampled: fourteen where SR and NSR morphotypes were found in sympatry (including Belinga Mountains, Cristal Mountains, Kongou, Makokou, table 1) and four with SR trees only. To obtain outgroups for phylogenetic analyses, samples of the following species from the same tribe as *Santiria* were also collected in Gabon or Cameroon: *Dacryodes buettneri* Engl. (specimen MH0794), *D. edulis* (G. Don) H.J.Lam (specimen MH1400), *D. klaineana* (Pierre) H.J.Lam (GK487 and GK1117), *D. macrophylla* (Oliv.) H.J.Lam (specimen MH1555) and *Canarium schweinfurthii* Engl. (specimen MH1355). Tissues were dehydrated with silicagel in the field. Total DNA was extracted from c. 20 mg dry plant

material ground in liquid nitrogen using a CTAB procedure (Doyle & Doyle 1987), in which case DNA pellets were resuspended in 100 µl light TE buffer (20 mM Tris-HCl and 0.1 mM EDTA), or from c. 10 mg dry plant material using the Nucleospin Plant kit (Macherey-Nagel, Düren, Germany).

Three cpDNA regions (*trnL-F*, *rbcL* and *rpl36-infA-rps8*) were successfully amplified in 377 individuals using universal primers (Kress et al. 2005, and references therein). One nuDNA sequence (intron Phosphoenolpurivate Carboxylase Exon2-Exon3: *PepC E2-E3*) was also amplified in forty individuals using universal primers (forward: 5'TGCCAAAGA-CATCACTCCTG3' and reverse: 5'ATCACCACCCATC-CAAGAAG3') designed by J. Duminil (Université Libre de Bruxelles, Belgium, pers. comm.). Individuals for nuDNA sequencing were selected according to their identity of cpDNA haplotypes and their geographic origin. PCR reactions were performed in PTC-100 or PTC-200 thermocyclers (MJ Research) in a reaction volume of 25 µl. For cpDNA, they contained 2.5 mM of $MgCl_2$, 0.2 mM of each dNTP, 0.1 µM of each primer, 0.625 units of Taq polymerase (Qiagen, Venlo, the Netherlands) and 1 µl of DNA extract diluted 50 times. Cycling conditions were as follows: 94°C for 3 min, 35 cycles of 94°C for 20 sec, 50°C for 30 sec and 72°C for 1 min 20 sec, followed by 72°C for 7 min and cooling at 10°C. The nuDNA region was amplified using the Phusion polymerase (Finnzymes, Espoo, Finland) according to the manufacturer's protocol. Most PCR products were sequenced by the Genoscope facility (Centre National de Séquençage, Institut de génomique, France). Otherwise PCR products were purified on filter columns (QiAquick 96 kit, Qiagen, Hilden, Germany or MSB HTS PCRapace/C(96) kit, Invitek, Berlin, Germany) or using the enzymes Exonuclease I and Shrimp Alkaline Phosphatase (ExoSAP, Cleveland, Ohio USA). Purified PCR products were quantified on agarose gels (1%). Sequencing reactions were performed in both directions using Big Dye v.3.1 chemistry (Applied Biosystems, Lennik, the Netherlands), purified with an ethanol-sodium acetate protocol and analysed on an ABI 3100 or ABI 3730 sequencer (Applied Biosystems, Lennik, the Netherlands).

Note – DNA sequences obtained in this paper have been deposited with the EMBL/Genebank libraries under accession numbers: FM162235 - FM162350.

Sequences alignment

Sequence trace files were analysed and edited using the Codoncode Aligner software (available from www.codoncode.com/aligner/download.htm) including base-calling with Phred and alignment with Phrap. A site of a single nucleotide was considered polymorphic if different variants had at least a Phred quality value of 25, corresponding to an error probability of 3/1000. For nuDNA, heterozygotes (double peaks) were detected in c. 31% of sequences. At most four double peaks were detected in one heterozygous individual. It was however straightforward to infer the sequences of the two constituent haplotypes using the PHASE program implemented in the DnaSP v5 software (Librado & Rozas 2009).

Molecular data analysis

Phylogenetic analyses of nuDNA sequences and combined cpDNA sequences were performed separately. Maximum parsimony (MP) and likelihood (ML) analyses were performed using PAUP v. 4.0b10 (Swofford 2003). The best evolutionary model was determined using Modeltest 3.7 (Posada & Crandall 1998). The evolutionary model inferred by Hierarchical Likelihood Ratio Tests (hLRT) was F81 + I(0.8673) + G(0.0087) with A = 29.6%, C = 19.82%, G = 18.62%, T = 31.96% for cpDNA and HKY + G(0.2317) with A = 28.98%, C = 16.84%, G = 18.13%, T = 36.05% for nuDNA. Non-parametric bootstrap support was estimated for each clade (100

Table 2 – Comparison of various traits between the morphotype with stilt roots (SR) and the morphotype without stilt roots (NSR).
Made in an inventory (2.5 ha = 100 × 250 m) at Makokou field research station (Ivindo National Park, Gabon). [1]all individuals with a dbh ≥ 10 cm were measured (diameter, total height and height of the first fork) and mapped. [2]all individuals over 1 m total height and less than 10 cm dbh were counted by 10 × 10 m quadrats. Hf: height of first fork (m). Ht: total height (m). [3]SR and NSR means significantly different (t-test; p-value = 0.020).

Variables	Stilt Roots (SR)	No Stilt Root (NSR)
density (trees/ha)[1]	6	12
basal area (m²/ha)[1]	0.34	0.68
regeneration (stems/ha)[2]	20	58
regeneration (maximum number of stems/10 × 10m)[2]	3	11
ratio Hf/Ht: mean (standard deviation)[1,3]	0.75 (0.13)	0.60 (0.15)
trees bearing woody vines (%)[1]	53	35

iterations). Bayesian inference of phylogeny was performed using MrBayes v. 3.1.2 (Huelsenbeck & Ronquist 2001). For cpDNA, MP, ML and Bayesian approaches (StdDev = 0.002; 3,000,000 generations) converged to the same consensus gene tree. Concerning nuDNA, MP and ML converged to the
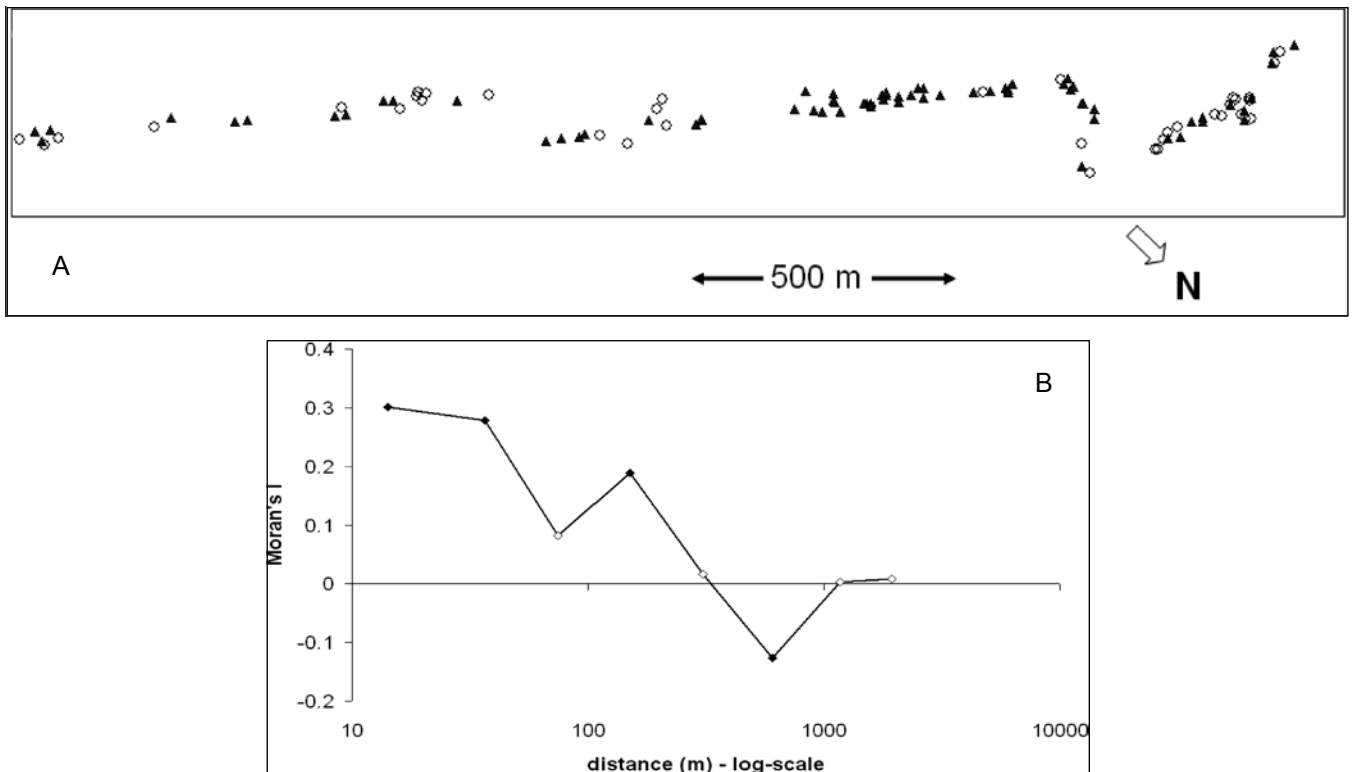

Figure 1 – A, spatial locations of both morphotypes of *Santiria trimera* along a transect in North-eastern of Gabon (Makokou). Filled triangles: morphotype with stilt roots (SR); empty circles: morphotype without stilt roots (NSR); B, spatial autocorrelation of morphotype status for the above transect using Moran's *I* statistic for a set of distance intervals. Filled symbols indicate values departing significantly from the null hypothesis that morphotypes are distributed at random. Significantly positive values at short distance demonstrate a tendency of aggregation of trees of the same morphotype.

same gene tree but the best Bayesian tree (StdDev = 0.002; 3,350,000 generations) diverged from the two first ones for a clade.

<div align="center">RESULTS</div>

Spatial distribution of morphotypes

Along the six transects, trees of the SR and NSR morphotypes were largely intermingled and could be found very close to each other (illustration of a transect in Makokou; fig. 1A). In particular, there was no spatial trend in the respective distributions of the morphotypes even in transects situated on a steep altitudinal gradient (e.g. in Belinga Mountains, results not shown). However, the spatial autocorrelation analysis revealed a tendency of aggregation of trees belonging to the same morphotype (fig. 1B), a pattern expected under limited seed dispersal.

In the detailed inventory of the Makokou 2.5 ha plot, the NSR morphotype is twice as abundant as the SR morphotype (6 vs. 12 trees/ha with dbh ≥ 10 cm), with a much bigger basal area (table 2). Regeneration (individuals over 1m height and less than 10 cm dbh) is three times more important for NSR than for SR, and sometimes highly clumped (table 2).

Morphological analysis

To test whether it is statistically possible to differentiate sympatric morphotypes based on quantitative and qualitative morphological traits, we focused our observations on morphological traits a priori distinguishing morphotypes and enabling rapid identification in the field. Vegetative traits on flowering branches (length and width of terminal leaflet, petiole length and the number of leaflets) and reproductive traits (length and width of drupes) were significantly different between both sympatric morphotypes (table 3). Drupes and terminal leaflets of SR were narrower and shorter than those of NSR. Petioles of NSR were longer than the ones of SR whereas the number of pairs of leaflets was higher for SR. The principal component analysis (fig. 2) displayed two groups without any intermediate between both morphotypes.

Table 3 – Morphological quantitative traits.
Comparison between the morphotype with stilt roots (SR, N = 11) and the morphotype without stilt roots (NSR, N = 9) of *Santiria trimera* in northern Gabon. Seven to ten measures on floriferous branches per tree were performed. All differences at quantitative traits are significant (P < 0.001; Wilcoxon tests).

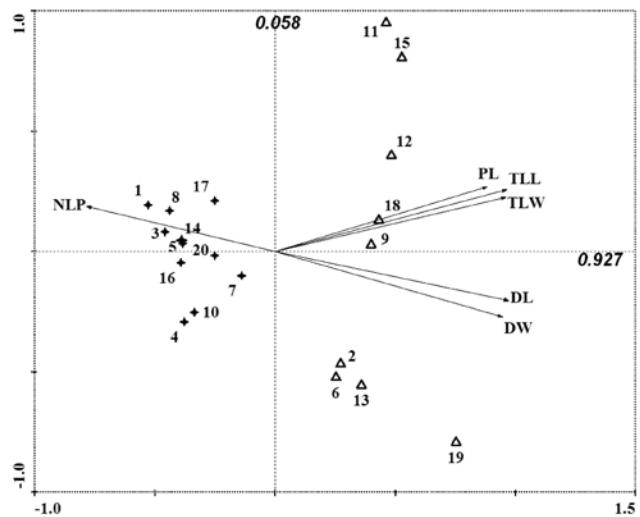| Variables	Mean (standard deviation)	
	Stilt Roots (SR)	**No Stilt Root (NSR)**
Drupe width (mm)	16.49 (1.69)	25.16 (3.83)
Drupe length (mm)	20.95 (2.38)	33.28 (3.14)
Terminal leaflet width (cm)	5.63 (2.29)	10.93 (2.34)
Terminal leaflet length (cm)	15.35 (5.55)	26.12 (5.53)
Petiole length (cm)	2.53 (0.75)	4.15 (1.01)
Number of leaflet pairs	4.21 (1.02)	3.51 (0.63)

Figure 2 – Principal Component Analysis of quantitative traits assessed on individuals of *S. trimera* of morphotypes with stilt roots (SR, + symbols) and without stilt roots (NSR, open triangles) found in sympatry in north-east Gabon. NLP = numbers of pairs of leaflets; PL = length of petiole; TLL = terminal leaflet length; TLW = terminal leaflet width; DL = drupe length; DW = drupe width. Values along each axis indicate eigenvalues. For each individual, seven to ten measures of each trait were averaged.

The six quantitative variables could be summarized in a single axis expressing 92.7% of the total variance. Hence, the six morphological traits were highly correlated. Qualitative traits also distinguished both morphotypes: the SR has a resin strongly smelling to turpentine, green fruits when immature, and a thin green inflorescence rachis, whereas the NSR has a resin with a mild smell, yellowish, yellow or red immature fruits and a brown, robust inflorescence rachis (table 4, electronic appendix 2).

Molecular analysis

The polymorphism of cpDNA sequences was low. Combining *trnL*, *rpl36-infA-rps8* and *rbcL* sequences, fourteen sites out of 1598 base pairs (bp) were polymorphic and defined twelve haplotypes among 377 individuals of *S. trimera* (electronic appendix 5). By contrast, sequences of the nuclear intron *PepC E2-E3* were much more variable: 39 sites were polymorphic among 737 bp, defining 19 haplotypes over 40 individuals of *S. trimera* (electronic appendix 6). Phylogenetic analyses supported the monophyly of *S. trimera* with respect to other African species from the same tribe.

The twelve cpDNA haplotypes of *S. trimera* formed three well supported clades (A, B and C) and clade C was composed of a further three well supported clades (C1, C2, C3; fig. 3A). Individuals of clade A came from São Tomé island (AR), being thus geographically isolated and genetically differentiated from the continental populations. All individuals of morphotype NSR found in Gabon and Equatorial Guinea were monophyletic in clade C3, bearing haplotypes H11 or H12. Interestingly, clade C3 was sister to clade C2 containing all individuals from West Africa which bear stilt roots (SR

Table 4 – Morphological qualitative traits.
Comparison between the morphotype with stilt roots (SR, N = 17) and the morphotype without stilt roots (NSR, N = 17) of *Santiria trimera* in northern Gabon.

Variables	Stilt Roots (SR)	No Stilt Root (NSR)
Resin smell (turpentine)	very strong	mild like *Dacryodes*
Fruit colour	green	yellowish, yellow or red
Inflorescence rachis	green, thin	brown, robust
Inflorescence	axial, sometimes terminal	terminal

morphotype), while the large majority of Central African SR individuals gathered in clade B. Hence, sympatric SR and NSR morphotypes in Gabon and Equatorial Guinea generally belonged to deeply divergent clades. However, seven SR individuals found in sympatric populations with NSR individuals from Gabon also bore H11 or H12 from clade C3. As a conclusion, all 304 individuals (100%) from clades B, C1 and C2 had stilt roots (SR), and 90.3% ([100/(65 + 7)]*65) of individuals belonging to clade C3 lacked stilt roots (NSR).

Sequences of the intron *PepC E2-E3* gave globally similar results as cpDNA. All sequences of the NSR morphotype (S13-S17) were monophyletic in a clade also containing six sequences from three SR individuals coming from sympatric populations, so that 76.9% of individuals with haplotypes S13-S17 corresponded to the NRS morphotype. This clade was sister to a haplotype (S12) found in the SR morphotype from south Gabon. Interestingly, S12 was observed in the same individuals that bore H07 (clade C1) and were therefore the SR closest related to NSR at cpDNA as well. Similarly as with cpDNA, the other SR individuals from Central Africa belonged to several distinct clades and were not monophyletic. Notably, one of these clades (harbouring S01 and S02) was inserted at the base of the gene tree, so that sequences found among SR individuals were much more divergent than those found among NSR individuals. Individuals from São Tomé formed a monophyletic clade, as well as individuals from West Africa. The São Tomé clade (S10 and S11) was sister to most SR clades from Central Africa (fig. 3B).

DISCUSSION

Differentiation between morphotypes in Gabon

According to observations on adult trees (dbh > 15 cm) from north Gabon, we easily distinguished two forms of *S. trimera* based on the presence or absence of stilt roots (Hladik & Hallé 1973, Florence & Hladik 1980, Reitsma 1988) and confirmed the congruence of this character with other morphological traits. Qualitative morphological traits were constant within each individual but varied to some extent among individuals within morphotypes (e.g. immature drupes of the NSR morphotype were yellowish, yellow or red). Consider-

ing quantitative traits, the SR morphotype is characterized by smaller leaflets and drupes, shorter petioles but a higher number of leaflet pairs than the NSR morphotype.

The PCA graph (fig. 2) showed two non-overlapping groups of individuals corresponding to SR and NSR morphotypes. Individuals of the SR morphotype constituted a fairly homogeneous group while individuals of the NSR morphotype were spread along the second PCA axis. Hence, quantitative and qualitative traits converged on the hypothesis that there are at least two distinct sympatric taxa within *S. trimera* in Gabonese forests. This hypothesis is also supported by both chloroplast and nuclear DNA sequences, which indicate that the haplotypes of the NSR morphotype gathered in a single clade.

Interestingly, a few individuals of the SR morphotype carried haplotypes typically found in NSR individuals, and this was observed both at the chloroplast and nuclear sequences, but not necessarily for the same individuals. This phenomenon could be due to incomplete lineage sorting (shared ancestral variation) between taxa or to rare events of hybridization (e.g. Muir & Schlötterer 2005). Shared ancestral haplotypes may be observed if effective population sizes are large and the speciation event recent (Pamilo & Nei 1988). The fact that all SR individuals bearing a haplotype typical of the NSR morphotype were found only in areas where both morphotypes occur supports the hybridization hypothesis, especially if the flowering periods of both morphotypes are overlapping. In Makokou, where phenological analyses have been undertaken, the main flowering period of SR and NSR morphotypes are respectively September-October and August-September (Hecketsweiler 1992). This fact probably fosters genetic isolation, while it might allow for hybridization events from time to time. At the nuclear locus, the three SR individuals with sequences typical of the NSR were homozygous (S15-S15) and thus could not be first generation hybrids. Even if the hybridization hypothesis is plausible, hybridization events might be ancient. We did not observe any individual of the NSR morphotype carrying a chloroplast haplotype typical of the SR morphotype. Hence, gene flow might be unidirectional, assuming maternal inheritance of the chloroplast like in most angiosperms (SR pollen to NSR ovule; e.g. Birky 1995). Moreover, the fact that morphotypes remain well differentiated for many traits despite spatial proximity suggests that hybridization should be rare and/or hybrid progenies should be counter selected. Experimental crosses and/or the use of highly polymorphic nuclear genetic markers (e.g. microsatellites) would be required to test further the extent of reproductive isolation between these distinct morphotypes.

No evidence was found for habitat specialization of the two morphotypes in the area of sympatry. Along transects covering environmental gradients, both morphotypes were intermingled, although a trend for spatial aggregation of each morphotype probably results from seed dispersal limitation. The absence of habitat differentiation can be surprising for two reasons. First, sympatric sister species are expected to show some niche differentiation, otherwise one should be excluded according to competitive exclusion principles (Hardin 1960). Second, the highly contrasted trunk architecture of the two morphotypes suggests a specialization for different
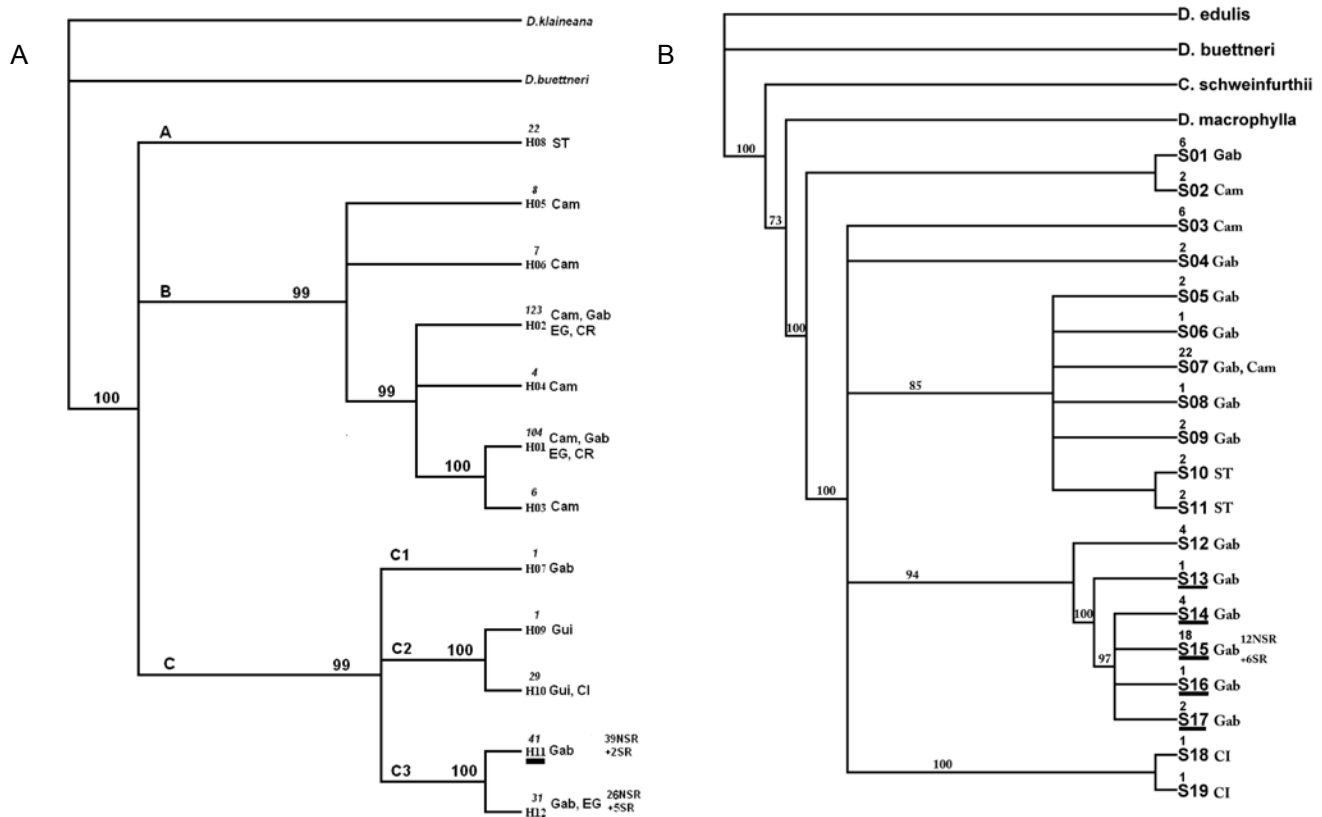
Figure 3 – A, Phylogenetic consensus tree obtained with three methods (see materials and methods) on cpDNA haplotypes (*trnLF*, *rbcL*, *rpl36-infA-rps8* combined). A, B, C, C1, C2 and C3 on branches refer to clades discussed in the text. B, Phylogenetic tree obtained with Bayesian analysis on sequences of the *PepC E2-E3* intron. Outgroups are *Dacryodes* spp. and *Canarium schweinfurthii*. Numbers on branches indicate non-parametric bootstrap values. Haplotypes underlined are typical of morphotype without stilt roots (NSR). Numbers above haplotype names indicate the number of individuals for each haplotype. The origin of haplotypes follows haplotype names: Cam = Cameroon; CR = Central African Republic; CI = Ivory Coast; EG = Equatorial Guinea; Gab = Gabon; Gui = Guinea; ST = São Tomé. For the few haplotypes found in both morphotypes with and without stilt roots (SR and NSR), the numbers of individuals of each morphotype are indicated.

edaphic conditions. In fact, the putative adaptive role of stilt roots in the widespread SR morphotype is unclear. Stilt roots can be associated with unstable grounds like marshy soils or mangroves, but *S. trimera* is found in terra firme forests and is not associated with unstable soils, even if it can occasionally be found on marshy ground. In addition, the development of stilt roots is variable: while some trees can stand on many well-developed stilt roots and lack a basal trunk (electronic appendix 1), others can present a main trunk with just one or two stilt roots sprouting at 0.5 m height. The role of phenotypic plasticity versus genetic variability in this polymorphism is unknown but deserves to be studied to gain insight into the adaptive role of stilt roots.

Despite the lack of strong habitat specialization, both morphotypes present some ecological traits suggesting more subtle ecological niche differentiation. At Makokou, where more detailed ecological research had been undertaken, the SR morphotype has a higher first fork/total height ratio than NSR (table 2), suggesting that SR may be growing under more shaded conditions. Typically, SR trees insert themselves between other canopy trees, competing for light; this strategy allows woody vines to develop more abundantly on SR morphotypes (table 2). On the other hand, the NSR mor-

photype seems to react more readily to canopy openings by higher potential growth rates than the SR morphotype when close to tree falls (Hladik 1982). In the Makokou plot, more abundant regeneration has been noted in tree fall openings at some tens of metres from big NSR trees. Moreover, seedlings and very young shoots of the NSR morphotype seem to grow more vigorously than SR when in or at the edge of a tree fall gap and leaf turnover is quicker (Hladik & Blanc 1987). So the NSR morphotype seems to be more light demanding than the SR one.

Evolutionary processes within *Santiria trimera*

The depth of a gene tree can be interpreted as a signal of long time isolation of clades. Chloroplast and nuclear DNA congruently show that sequences from São Tomé and those from West Africa form two monophyletic groups, indicating ancient isolation from Central Africa. As we observed (O.J. Hardy) that the balsam of São Tomé (common name of the particular form of *S. trimera* in São Tomé, AR) is fairly distinct from the continental SR morphotype (electronic appendix 1), having thin aerial hanging roots emerging through the trunk, a very fluid resin which burns like the one of *Au-*

coumea klaineana (Aubréville 1948) and a trunk which can grow very large (one meter in diameter), we suspect that the AR morphotype could represent a separate taxon that arose by geographic isolation (see also Aubréville 1948). Further morphological analyses including traits of reproductive organs and, ideally, experimental crosses with continental morphotypes are necessary to test this hypothesis. The nuDNA gene tree suggests that AR is a sister taxon of a Central African clade of the SR morphotype. As São Tomé is a volcanic island that emerged c. 200 km off the coast of Gabon 13 My ago (Plana et al. 2004), it is likely that it has been colonised by seeds from Central Africa.

Another interesting observation is that both cpDNA and nuDNA identified the same SR individuals as closest related to NSR (CD0777 of SE Gabon, and CD0539 of SW Gabon, although the latter had only two of the three cpDNA fragments sequenced; electronic appendix 4). Detailed morphological studies in south Gabon might reveal whether these individuals constitute a transition form between SR and NSR morphotypes.

The SR morphotype is polyphyletic and harbours more diverse and more divergent nuDNA and cpDNA lineages than the NSR. This suggests that SR is more ancient than NSR, and therefore that the presence of stilt roots is probably the ancestral state of *S. trimera*. Under this hypothesis, stilt roots have probably been lost during the differentiation process leading to the NSR morphotype. The cpDNA clade characterizing NSR (C3) is close to a West African (C2) and to a Central African clade (C1), suggesting that the differentiation event leading to the NSR morphotype originated from a widespread ancestor different from the one leading to most contemporary SR individuals from Central Africa (clade B). Combining these results, one possible evolutionary scenario is that stilt roots were an adaptive trait of *S. trimera* when the species appeared (because Asian *Santiria* species do not bear stilt roots), perhaps because it was initially adapted to marshy conditions, *S. trimera* then expanded to wider range of habitats without losing stilt roots. Another explanation might rely in the mechanical role of the stilt roots, favouring the stabilization of the tree and the development of the root system of SR trees, which are growing more in height, with a higher crown than NSR trees. The NSR morphotype appeared subsequently by allopatric differentiation and lost stilt roots by genetic drift or selection in particular conditions. Finally, according to distribution range shifts, the SR and NSR morphotypes encountered in secondary contact zones and can coexist because of sufficient niche differentiation and/or because competitive exclusion is too slow to eliminate one morphotype, as shown by neutral community models (Hubbell 2001). In addition, prezygotic barriers seem to strongly limit gene flow and the production of low viability hybrids.

Taxonomic implications

Considering the Biological Species Concept (BSC) at the light of morphological traits and the phylogenies of cpDNA and nuDNA sequences, the sympatric morphotypes SR and NSR found in Gabon should be defined as two distinct species (reproductive entities). In fact, *S. trimera* is dioecious

(and consequently allogamous), so the genetic divergence between both sympatric morphotypes with only rare occurrences of shared haplotypes is probably due to reproductive isolation. The correlation of patterns of genetic and morphological divergence in sympatric zones is evidence of a speciation process. Further work is required to assess whether populations from São Tomé or West Africa, which belong to distinct phylogenetic clades, also constitute distinct species. Species delimitation might be more difficult and arbitrary at this scale because reproductive isolation occurs de facto by geographic isolation and it is not possible to assess whether reproductive isolation would occur under sympatry without performing experiments.

In sharp contrast with our conclusions, in a revision of the genus *Santiria* in Africa based on herbarium samples collected over a large area (Guineo-Congolian region), Onana (2009) recognized a single African species, although both SR and NSR morphotypes were represented in the vouchers analysed. Herbarium samples showed a wide range of variation with complex combinations of morphological traits (O. Lachenaud, Université Libre de Bruxelles, Belgium, pers. comm.; Onana 2009), preventing the recognition of distinct species using unambiguous morphological traits from a taxonomic perspective. As herbarium samples came from the whole distribution area of African *Santiria*, the distribution of quantitative traits appeared continuous while our sample from Gabon was clearly bimodal. Herbarium samples often lack information for some key distinctive traits such as the presence/absence of stilt roots or the colour of immature fruits and the distinction between morphotypes at quantitative traits may require relatively large sample sizes. Although this may partly explain the discrepancy between our respective conclusions, intermediate or recombining forms between SR and NSR morphotypes have been occasionally observed in herbarium vouchers and in the field, and might be related to our genetic data suggesting occasional hybridization. This highlights that species concepts in herbarium taxonomy or following the BSC do not always converge because reproductive isolation needs not to be concomitant with the appearance of morphological traits. Incongruence likely originates from the limits of the notion of species. Speciation is rarely a one step process and species delimitation can remain very arbitrary during a speciation process. African *Santiria* might thus undergo an ongoing speciation. From a taxonomic perspective, as long as diagnostic characters valid over the whole distribution of *Santiria* are lacking to identify unambiguously our species following the BSC, the latter should be classified at an infra-specific level.

Our results also highlight that species delimitations based on herbarium samples alone may not be sufficient to define species in concordance with the BSC. Wrong species delimitations can lead to wrong inference of evolutionary processes (biogeography). Hence, while herbarium based analyses remain essential for taxonomy, population level analysis and/or molecular markers could often be required to delineate closely related taxa. This study allows defining distinct genetic units which agree with morphological divergences within *S. trimera*. Because of complex combinations of some morphological traits and polyphyletic clades of nuDNA and cpDNA for the SR morphotype (as reported above), further morpho-

logical (even histological), ecological and phenological studies of taxa outlined in this study are required to specify their taxonomic status. Additional genetic work using the power of highly polymorphic nuclear genetic markers, such as microsatellites, is also necessary to identify all reproductive entities that might match the BSC.

SUPPLEMENTARY DATA

Supplementary data are available at *Plant Ecology and evolution*, Supplementary Data Site (http://www.ingentaconnect.com/content/botbel/plecevo/supp-data), and consist of the following: (1) colour photographs showing characteristic traits of the several morphotypes (pdf format); (2) colour photographs of fruits of the NSR morphotype (pdf format); (3) list of fertile vouchers (pdf format); (4) list of *Santiria trimera* samples with their origin, haplotype identity for chloroplast markers and genotype at the nuclear intron Phosphoenolpurivate Carboxylase Exon2-Exon3 (PepCE2-E3) sequence (pdf format); (5) table showing the polymorphic positions of the twelve cpDNA haplotypes of *Santiria trimera* (pdf format); (6) table showing the polymorphic positions of the nineteen haplotypes on the nuclear intron Phosphoenolpurivate Carboxylase Exon2-Exon3 (pdf format).

ACKNOWLEDGEMENTS

REFERENCES

Aubréville A. (1948) Dacryodes (Pachylobus) et Santiria de l'Ouest africain. Bois et forêts des tropiques 8: 342–348.

Aubréville A. (1962) Flore du Gabon vol. 3: Irvingiacées, Simaroubacées, Burseracées. Paris, Muséum National d'Histoire Naturelle Laboratoire de Phanérogamie.

Birky C.W.J. (1995) Uniparental inheritance of mitochondrial and chloroplast genes: mechanisms and evolution. Proceedings of the National Academy of Sciences of the USA 92: 11331–11338.

Breteler F.J. (2010) Revision of the African genus Anthonotha (Leguminosae, Caesalpinioideae). Plant Ecology and Evolution 143: 70 – 99.

Dauby G., Duminil J., Heuertz M., Hardy O.J. (2010) Chloroplast DNA polymorphism and phylogeography of a central African tree species widespread in mature rainforests: Greenwayodendron suaveolens (Annonaceae). Tropical Plant Biology 3: 4–13.

Dayrat B. (2005) Towards integrative taxonomy. Biological Journal of the Linnean Society 85: 407–415.

de Queiroz K. (2005) Different species problems and their resolution. BioEssays 27: 1263–1269.

Doyle J.J., Doyle J.L. (1987) A rapid DNA isolation procedure from small quantities of fresh leaf tissues. Phytochemical Bulletin19: 11–15.

Ellis J.S., Knight M.E., Carvell C., Goulson D. (2006) Cryptic species identification: a simple diagnostic tool for discriminating between two problematic bumblebee species. Molecular Ecology Notes 6: 540–542.

Florence J., Hladik A. (1980) Catalogue des Phanérogames du N.-E. du Gabon (6e liste). Adansonia 20: 235–253.

Hardin G. (1960) The competitive exclusion principle. Science 131: 1292–1297.

Hecketsweiler P. (1992) Phénologie et saisonnalité en forêt gabonaise. L'exemple de quelques espèces ligneuses. PhD thesis, Université des Sciences et Techniques – Montpellier II, Montpellier, France.

Hladik A., Hallé N. (1973) Catalogue des Phanérogames du Nord-Est du Gabon (5e liste). Adansonia 13: 527–544.

Hladik A. (1982) Dynamique d'une forêt équatoriale Africaine: mesures en temps réel et comparaison du potentiel de croissance des différentes espèces. Acta Oecologica, Oecologica Generalis 3: 373–392.

Hladik A., Blanc P. (1987) Croissance des plantes en sous-bois de forêt dense humide (Makokou, Gabon). Revue d'Ecologie: La Terre et la Vie 42: 209–234.

Hubbell S.P. (2001) The Unified Neutral Theory of Biodiversity and Biogeography. Princeton, Princeton University Press.

Huelsenbeck J.P., Ronquist F. (2001) MRBAYES: Bayesian inference of phylogeny. Bioinformatics 17: 754–755.

Kenfack D. (2008) Systematic studies in Carapa (Meliaceae-Swietenioideae). PhD thesis, University of Missouri – St Louis, St Louis.

Kress W.J., Wurdack K.J., Zimmer E.A., Weigt L.A., Janzen D.H. (2005) Use of DNA barcodes to identify flowering plants. Proceedings of the National Academy of Science of the USA 102: 8369–8374.

Lachenaud O., Séné O. (2010) Un nouveau Psychotria (Rubiaceae) du sud Cameroun. Plant Ecology and Evolution 143: 105–108.

Lee M.S.Y. (2003) Species concepts and species reality: salvaging a Linnaean rank. Journal of Evolutionary Biology 16: 179–188.

Letouzey R. (1985) Notice de la carte phytogéographique du Cameroun au 1:500 000. IRA, Yaoundé, Cameroun. Toulouse, Institut National de la Carte de Végétation.

Librado P., Rozas J. (2009) DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. Bioinformatics 25: 1451–1452. DOI: 10.1093/bioinformatics/btp187

Miikeda O., Kita K., Handa T., Yukawa T. (2006) Phylogenetic relationships of Clematis (Ranunculaceae) based on chloroplast and nuclear DNA sequences. Botanical Journal of the Linnean Society 152: 153–168.

Moran P.A.P. (1951) Partial and multiple rank correlation. Biometrika 38: 26–32.

Muir G., Schlötterer C. (2005) Evidence for shared ancestral polymorphism rather than recurrent gene flow at microsatellite loci differentiating two hybridizing oaks (Quercus spp.). Molecular Ecology 14: 549–561.

Murakami N., Yokoyama J., Cheng X., Iwasak H., Imaichi R., Iwatsuki K. (1998) Molecular α-taxonomy of Hymenasplenium obliquissimum complex (Aspleniaceae) based on rbcL sequence comparisons. Plant Species Biology 13: 51–56.

Onana J.M. (2009) Le genre Santiria (Burseraceae) en Afrique: re-définition de Santiria trimera. Systematics and Geography of Plants 79: 215–224.

Pamilo P., Nei M. (1988) Relationships between gene trees and species trees. Molecular Biology and Evolution 5: 568–583.

Plana V., Gascoigne A., Forrest L.L., Harris D., Pennington R.T. (2004) Pleistocene and pre-pleistocene Begonia speciation in Africa. Molecular phylogenetics and evolution 31: 449–461.

Posada D., Crandall K.A. (1998) Modeltest: testing the model of DNA substitution. Bioinformatics 14: 817–818.

Reitsma J.M. (1988) Végétation forestière du Gabon/Forest vegetation of Gabon. Ede, Technical series Tropenbos, Stichting Tropenbos.

Sei M., Porter A.H. (2007) Delimiting species boundaries and the conservation genetics of the endangered maritime ringlet butterfly (Coenonympha nipisiquit McDunnough). Molecular Ecology 16: 3313–3325.

Senterre B. (2005) Recherches méthodologiques pour la typologie de la végétation et la phytogéographie des forêts denses d'Afrique tropicale. PhD thesis, Université Libre de Bruxelles, Brussels, Belgium.

Sonké B. (2000) Une nouvelle espèce de Rothmannia (Rubiaceae, Gardenieae) de Banyang Mbo (Cameroun). Systematics and Geography of Plants 70 : 149–153.

Sonké B., Esono P., Nguembou K.C., Stévart T. (2005) Une nouvelle espèce de Bertiera Aubl. (Rubiaceae) du sous-genre Bertierella découverte en Guinée Équatoriale et au Cameroun. Adansonia sér. 3, 27: 309–315.

Stévart T., Cribb P. (2004) Five new taxa of Tridactyle (Orchidaceae) from West Central Africa. Kew Bulletin 59: 195–205.

Swofford D.L. (2003) PAUP*. Phylogenetic Analysis Using Parsimony (*and other methods), v. 4. Sunderland, Massachusetts, Sinauer Associates.

Weeks A., Daly D.C., Simpson B.B. (2005) The phylogenetic history and biogeography of the frankincense and myrrh family (Burseraceae) based on nuclear and chloroplast sequence data. Molecular Phylogenetics and Evolution 35: 85–101.

Communicating Editor: Elmar Robbrecht.