American Journal of Botany 94(9): 1491–1505. 2007.

Boreotropical migration explains hybridization between geographically distant lineages in the pantropical clade Sideroxyleae (Sapotaceae)[1]

Jenny E. E. Smedmark[2] and Arne A. Anderberg

Department of Phanerogamic Botany, Swedish Museum of Natural History, Box 50007, SE-104 05 Stockholm, Sweden

To determine whether the fragmented pantropical distribution of present day Sideroxyleae primarily is the result of long-distance dispersals or represents the remnants of a once continuous distribution in the northern hemisphere, the boreotropical flora, we used phylogenetic analyses of chloroplast and nuclear ribosomal DNA data, Bayesian molecular dating, and Bayesian estimation of ancestral areas. Incongruence between the two data sets was examined with a nuclear low copy gene phylogeny to discover any occurrences of reticulate evolution. The Pacific clade *Nesoluma* was shown to have two distinct copies of the nuclear low copy gene *AAT*, one from an African and one from an American ancestral lineage, indicating that it is of allopolyploid origin. We conclude that Sideroxyleae, including the ancestral lineages of *Nesoluma*, were part of the boreotropical flora and entered the New World via the north Atlantic land bridge. We also suggest that the distribution of extant species resulted from the cooling climate at the end of the Eocene. *Sideroxylon oxyacanthum* is shown not to belong in the group, but in Chrysophylloideae. A classification reflecting phylogenetic relationships, as well as new combinations for the species in *Nesoluma* under *Sideroxylon*, is presented.

Key words: allopolyploidy; biogeography; boreotropical flora; molecular dating; *Nesoluma*; phylogeny; Sapotaceae.

To understand why the most species-rich biomes, found in tropical and subtropical areas of the world, harbor such an enormous diversity of organisms, we need to understand their history. Recently, dated phylogenies have made it possible to make inferences about when the groups that constitute important components of these biomes have diversified, as well as their past distribution ranges, despite an often scarce fossil record. Such studies of different organismal groups will improve our understanding of how the biodiversity of, for example, rainforests has accumulated over time, and what factors have influenced variations in geographical ranges over time.

Sideroxyleae is a clade of about 80 species of trees and shrubs in the Sapodilla family (Sapotaceae), including the three genera *Sideroxylon* L., *Nesoluma* Baill. (three spp.) and *Argania* Roem. & Schult. (one sp.). It is found in arid or semi-arid areas as well as in rainforests, particularly in America and the islands of the Indian Ocean, including Madagascar. Considering the relatively low number of species, Sideroxyleae has a wide geographic distribution spanning mainland Africa (four spp.), Madagascar (six spp.), the Mascarenes (eight spp.), other Indian Ocean islands (three spp.), Macaronesia (two spp.), southeast Asia (three spp.), southern U.S. (nine spp.), Central America and Mexico (28 spp.), the Caribbean (24 spp.), South America (three spp.), and the Pacific islands (three spp.).

[1] Manuscript received 6 November 2006; revision accepted 29 June 2007.

The authors thank the directors of the herbaria B, P, BM, and K for kindly sending herbarium material on loan and providing samples for DNA study, J.-Y. Meyer for material of *Nesoluma nadeaudii*, G. Hao for material of *Eberhardtia aurata* and *Sideroxylon wightianum*, I. Bartish for the use of three unpublished DNA sequences, T. Eriksson for helpful discussion on fossil calibration, and two anonymous reviewers. This study was supported by a grant from the Swedish Research Council for Ericales phylogeny (to A. A.).

[2] Author for correspondence (e-mail: jenny@bergianska.se); present address: Bergius Foundation, Royal Swedish Academy of Sciences, Box 50017, SE-104 05 Stockholm, Sweden

Similar tropical or subtropical amphiatlantic disjunct distributions are seen in several other eudicot groups. By incorporating a dimension of absolute time, dated phylogenies have been able to provide different explanations for this type of fragmented distribution depending on when lineages occurring in different continents diverged. For older groups (e.g., Lauraceae, Chanderbali et al., 2001; Annonaceae, Richardson et al., 2004), vicariance caused by the breakup of Gondwana has at least partly explained extant distribution patterns. Other groups have been shown to have relict distributions caused by the disintegration of the boreotropical flora (Wolfe, 1975) that covered a large part of the northern hemisphere in the early Paleogene (e.g., Melastomataceae, Renner et al., 2001; Malpighiaceae, Davis et al., 2004; Annonaceae, Richardson et al., 2004; Burseraceae, Weeks et al., 2005; Moraceae, Zerega et al., 2005; Meliaceae, Muellner et al., 2006). During this time, the global climate was very warm (Zachos et al., 2001), and tropical vegetation was present at much higher latitudes than today (Wolfe, 1975; Tiffney, 1985a). The existence of a connection between the Eurasian and North American continental plates allowed boreotropical taxa to expand their distributions over both continents (Tiffney, 1985b; Morley, 2003). At the end of Eocene, temperatures declined markedly (Zachos et al., 2001), which caused many tropical and subtropical taxa to recede toward the equator (Wolfe, 1978, 1981; Collinson et al., 1999). Today, representatives from such lineages are found primarily in refugia in China, southeast Asia, Central America, Mexico, and Macaronesia (Morley, 2000). Less commonly, long-distance dispersal has been shown to explain transatlantic disjunctions in angiosperm groups (e.g., *Symphonia*, Dick et al., 2003; Annonaceae, Richardson et al., 2004).

In the case of Sapotaceae, no known fossils from the Cretaceous support a Gondwanan origin, and the Sapotaceae lineage as a whole has been inferred to be about 100 million years (My) old (Bremer et al., 2004) and to have probably originated in southeast Asia (Smedmark et al., 2006). Molecular phylogenetic analyses have shown that the Side-

roxyleae clade is nested within Sapotaceae (Anderberg and Swenson, 2003; Swenson and Anderberg, 2005; Smedmark et al., 2006). Sideroxyleae should consequently be younger than the 100 My estimated for the family, which refutes the idea that the breakup of Western Gondwana explains its current distribution pattern.

In the fossil record, pollen from Sapotaceae first appear in the Palaeocene of Europe (Kedves, 1967), and by the early Eocene (ca. 50 Mya) the group appears to have become widespread. From the Eocene, pollen types of the main lineages of extant Sapotaceae have been found in most continents (Harley, 1991), and the presence of fossil Sapotaceae pollen both in Europe and in North America indicates that Sapotaceae was part of the boreotropical flora (Morley, 2000). The fossil record does, however, not reveal whether the same is true for Sideroxyleae, but pollen from this group has been described from the Eocene of southern England (Gruas-Cavagnetto, 1976). If Sideroxyleae were part of the boreotropical flora and gradually migrated across Eurasia and North America during this time, we would expect to see divergences between groups occurring in the Old and New Worlds in the period from 60 to ca. 34 Mya. During this time, from the late Paleocene onwards (Morley, 2003) to the late Eocene or early Oligocene (Wolfe, 1975; Tiffney, 1985b), migration between Eurasia and North America is thought to have been made possible by the north Atlantic land bridge and/or narrow water gaps between continents. The drastic cooling of the climate in the transition between the Eocene and Oligocene (Zachos et al., 2001) made it impossible, or at least very unlikely, that tropical or subtropical taxa expanded their distributions across the north Atlantic land bridge from Eurasia to North America, or vice versa, later than 34 Mya.

Fruits in Sideroxyleae are fleshy and primarily eaten by primates and birds (Pennington, 1991). If bird dispersal was primarily responsible for distributing the group, we would expect to find that lineages have moved between geographical regions at different times, perhaps more at random without any specific pattern. We might also expect to see dispersal events between continents that coincide with the existence of land bridges or island arcs that could facilitate dispersal.

In the present study, we use Bayesian reconstruction of ancestral areas (Huelsenbeck and Bollback, 2001; Ronquist, 2004), accounting for the uncertainty both in mapping and in phylogenetic reconstruction, and Bayesian molecular dating (Thorne et al., 1998; Kishino et al., 2001) to discriminate between the two hypotheses, long-distance dispersal and boreotropical relict distribution, that may have caused the distribution patterns in extant Sideroxyleae.

Sometimes nuclear ribosomal DNA (nrDNA) and organellar DNA data support different topologies of phylogenetic trees (e.g., Smedmark and Eriksson, 2002; Kyndt et al., 2005; Zaldivar-Riveron, 2006). Therefore, it may not be enough to rely on one type of data when making inferences about organismal relationships. Previous phylogenetic studies, including a smaller taxon sample from Sideroxyleae, have been based on chloroplast DNA (cpDNA) data only (Anderberg and Swenson, 2003; Swenson and Anderberg, 2005; Smedmark et al., 2006). In this expanded study, we use sequences from the nuclear ribosomal internal transcribed spacer region (nrITS) to test the phylogenetic hypothesis obtained from analyses of cpDNA sequence data.

Reticulate evolution is common among angiosperms (Stebbins, 1947; Grant, 1971; Soltis and Soltis, 1993) and may be one explanation for incongruence between organellar and nuclear phylogenies. Hybrids are usually sterile, but a new fertile species may form through doubling of the chromosomes, allopolyploidy (Stebbins, 1947; Grant, 1971). Polyploids have been claimed to become more common with increasing geographical latitude (Löve and Löve, 1957) or altitude (Johnson and Packer, 1965). If polyploidy is rare in tropical or subtropical plants, we would expect reticulate evolution to be uncommon as well. Reports of hybrid speciation in nondomesticated tropical species are scarce, but some allopolyploid species of tropical trees in *Leucaena* Benth. are known (Hughes et al., 2002). So far there are no indications of allopolyploidization in Sapotaceae. Perhaps reticulate evolution, which is such an important mechanism for the generation of species diversity in many northern temperate plant groups, has not been sufficiently studied and does occur more frequently than currently recognized in tropical angiosperms. In the case of Sapotaceae, phylogenetic studies have been based either on cpDNA or nrITS data alone (Anderberg and Swenson, 2003; Bartish et al., 2005; Swenson and Anderberg, 2005; Smedmark et al., 2006), which will not reveal incongruence and possible reticulate evolution. In this study we aim to determine whether hybrid speciation has been a part of the evolutionary history of Sideroxyleae by comparing phylogenies based on nrDNA and cpDNA. We also use phylogenetic analysis of the nuclear low copy Aspartate aminotransaminase (*AAT*) gene, a gene that has not previously been used in any phylogenetic study. Nuclear low copy genes have the advantage of potentially providing information about both parental lineages, whereas cpDNA usually is uniparentally inherited, and the numerous copies of nrITS are most often identical because of homogenization (Arnheim et al., 1980). Many processes, however, affect low copy genes, such as lineage sorting, deletion, and duplication; these processes may make interpretation of phylogenies difficult (Doyle and Davis, 1998).

Although previous phylogenetic analyses have found support for several clades within Sideroxyleae, relationships among these groups have not been resolved with convincing support (Smedmark et al., 2006). It is, however, clear that the current classification (Pennington, 1991) is not supported by phylogeny. Previous phylogenetic studies have shown that some genera classified in this group do not belong here, namely *Sarcosperma* Hook.f., *Diploön* Cronquist (Anderberg and Swenson, 2003), and *Neohemsleya* T.D. Penn. (Smedmark et al., 2006). In Pennington's classification (Pennington, 1991), *Sideroxylon* has a broader circumscription compared to previous treatments (e.g., Dubard, 1912; Lam, 1939; Aubréville, 1964; Baehni, 1965), but phylogenetic analyses of chloroplast data have clearly indicated that both *Argania* and *Nesoluma* are nested within *Sideroxylon* s.l. (Anderberg and Swenson, 2003; Swenson and Anderberg, 2005; Smedmark et al., 2006). There is some support for recognition of smaller clades corresponding to previously recognized genera (Smedmark et al., 2006). For example, two clades consisting entirely of American species were strongly supported. One of these corresponded to the previously recognized genus *Bumelia* Sw., while the other included species that have been placed in *Mastichodendron* Cronquist and *Dipholis* A. DC. (Cronquist, 1945; Aubréville, 1964). *Bumelia* and *Dipholis* differ from the rest of Sideroxyleae in having segmented corolla lobes and are distinguished from each other by seeds with (*Dipholis*) or without (*Bumelia*) endosperm. Another well-supported clade is

the African *Sideroxylon* s.s. (Aubréville, 1964), which comprises species with horizontal embryos, in contrast to the vertical embryos of other members of Sideroxyleae. Nested within this group was also the Pacific *Nesoluma polynesicum*. Furthermore, there was support for another mainly African lineage consisting of *Sideroxylon mascatense* and *Argania spinosa*, thorny shrubs or small trees. In addition to these, a third thorny African species in Sideroxyleae, i.e., *Sideroxylon oxyacanthum*, is sometimes placed in a separate genus as *Spiniluma oxyacantha* (Aubréville, 1963). This species has not previously been included in phylogenetic analysis but was expected, based on its morphology and geographical distribution, to belong to this African lineage (Smedmark et al., 2006). In our present study we have included *S. oxyacanthum* to determine its closest relatives. By adding ITS data to the already quite large cpDNA data set (ca. 5 kb) from a previous study (Smedmark et al., 2006) and an additional 28 Sideroxyleae species, we also aimed to improve the understanding of the evolutionary history of Sideroxyleae in order to propose a classification that better reflects phylogenetic relationships.

MATERIALS AND METHODS

Plant material and taxon sampling—The analysis included 58 Sapotaceae species, 50 of which belong in Sideroxyleae. This sample represents 63% of the total species diversity in the group (Govaerts et al., 2001). Since Pennington's monograph of Sapotaceae (1991), phylogenetic studies have shown that three of the genera that he classified in Sideroxyleae do not belong in this group. *Sarcosperma* has been shown to be the sister of the remainder of Sapotaceae (Swenson and Anderberg, 2005) and was therefore used to root the trees in the present analysis, while *Diploön* Cronquist (Swenson and Anderberg, 2005) and *Neohemsleya* (Smedmark et al., 2006) were excluded from the analysis because they have been shown to belong in Chrysophylloideae. A few additional species were included to represent the remainder of Sapotoideae (*Capurodendron androyense* and *Palaquium formosanum*), as well as Chrysophylloideae (*Xantolis siamensis*, *Englerophytum natalense*, and *Omphalocarpum pachysteloides*). Leaf material sampled from the 58 Sapotaceae species was either silica gel dried or, more often, from herbarium specimens. Voucher specimens used for study are listed in Appendix 1.

Extraction, amplification, cloning, and sequencing—DNA was extracted with DNeasy Plant Mini Kit (QIAGEN, Valencia, California, USA) following the manufacturer's instructions. Leaf material was ground in AP1 buffer, using a Mini-beadbeater with 2.5-mm zirconia/silica beads (Biospec Products, Bartlesville, Oklahoma, USA). Polymerase chain reaction (PCR) was carried out using puReTaq Ready-To-Go beads (Amersham Biosciences, Buckinghamshire, UK) in a DNA Thermal Cycler 480 (Perkin Elmer, Wellesley, Massachusetts, USA). The *trnH-psbA* spacer was amplified using the primers described by Hamilton (1999). The *trnC-trnD* region (consisting of the *trnC-petN* spacer, the *petN* gene, the *petN-psbM* spacer, the *psbM* gene, and the *psbM-trnD* spacer) was amplified in two segments; the *trnC-psbM* region with the trnC (Demesure et al., 1995) and psbM2R (Lee and Wen, 2004) primers, and the *psbM-trnD* spacer with the psbM1 (Lee and Wen, 2004) and trnD (Demesure et al., 1995) primers. The 3′ end of *ndhF* was amplified with primers 972 and 2110R (Olmstead and Sweere, 1994), which are often called 5 and 10. The ITS region (ITS1, 5.8S, and ITS2) was amplified using the primers ITS18SF and ITS 26SR (Rydin et al., 2004). The Aspartate aminotransaminase gene (*AAT*) was amplified using the primers AATX5F and AATX7R (Strand et al., 1997). All PCR products were cleaned using QIAquick PCR Purification Kit (QIAGEN). In addition to the primers mentioned previously, petN1, petN2, psbM2, and petN2R (Lee and Wen, 2004) were used for sequencing the *trnC-trnD* region, 1260 (Eldenäs et al., 1999) and 1750R (Anderberg and Swenson, 2003) for *ndhF*, and ITS5.8F and ITS5.8R (Suh et al., 1993) for ITS. The data from ITS and cpDNA supported incongruent relationships of a few species (see Results section). To find any potential multiple copies of the nuclear markers in these species, PCR products were cloned using TOPO-TA Cloning Kit for Sequencing (Invitrogen, Carlsbad, California, USA). For ITS, six clones each were sequenced from *Nesoluma nadeaudii* and *N. polynesicum*. For *AAT*, five clones each from *Sideroxylon persimile*, *S. salicifolium*, *S. rotundifolium*, and *S. tepicense*, and six from *S. saxorum*, *N. nadeaudii*, and *N. polynesicum* were sequenced. If a species contains two gene copies that amplify and clone equally well, sampling five colonies results in a probability of 0.94 of finding both copies. Sampling six colonies under the same conditions would give a probability of 0.97 of finding both. A few additional species (two clones from *S. reclinatum*, three from *S. betsimisarakum* and *S. foetidissimum*, and five from *S. obtusifolium* and *S. occidentale*) were sequenced to clarify the issue of paralogy for the *AAT* gene within Sideroxyleae. Cloning procedures followed the manufacturer's instructions, except that ligation and transformation reaction volumes were halved and the ligation reactions were allowed to incubate for 15 min. Three to six colonies from each species were selected and used directly as templates in PCR reactions with the same primers that had been used to amplify the region. Sequencing reactions were performed with Big Dye Terminator Cycle Sequencing Kit v. 3.1 (Applied Biosystems, Foster City, California, USA) in a Gene Amp PCR System 9700 (PE Applied Biosystems, Warrington) and cleaned using DyeEx 96 plates (QIAGEN, Valencia, California, USA). Sequencing reactions were read in a 3100 Genetic Analyzer (Applied Biosystems).

Sequence assembly and alignment—DNA sequences were assembled and edited using the phred (Green and Ewing, 2002) and phrap (Green, 1999) modules in Pregap4 and Gap4 (Staden et al., 1998). All new sequences have been submitted to EMBL, and accession numbers are presented in Appendix 1. Sequences were aligned by eye in the sequence alignment editor Se-Al (Rambaut, 1996). Exons of the *trnC-trnD* region were removed from the data set because primers were located in these regions; therefore, data were missing for some taxa. These exons were also removed to make modeling of sequence evolution more rational. The two *Nesoluma* species were excluded from the combined data set because separate analyses indicated that their cpDNA and nrITS are of widely different origin (see Results section) and that they also had *AAT* alleles from different clades. Matrices and trees are available at TreeBASE (www.treebase.org).

Model selection—The best performing evolutionary models for the chloroplast, ITS, and *AAT* data sets were identified under two different model selection criteria, the Bayesian information criterion (BIC; Schwartz, 1978) and the Akaike information criterion corrected for small sample size in relation to the number of parameters of the model (AICc; Akaike, 1973) using MrAIC (ver. 1.4, Nylander, 2004). This program tests the substitution models that can be applied in MrBayes (ver. 3.1.1, Huelsenbeck and Ronquist, 2001; Ronquist and Huelsenbeck, 2003), in combination with different ways of modelling rate variation among sites, namely equal substitution rates, gamma distributed rate variation among sites (+Γ), and a proportion of sites with the rate zero (+I). The models evaluated by MrAIC, however, do not take rate variation over time into account. A previous study (Smedmark et al., 2006) showed that a cpDNA data set consisting of the same markers but with a partly different taxon sample within Sapotoideae was better explained by a model that allows the rate of evolution of a site to vary through time (Fitch and Markowitz, 1970; Fitch, 1971; Tuffley and Steel, 1998), viz. covarion models. For comparing the relative performance of models favored by different criteria and for comparing a covarion model to one that only accounts for rate variation across sites, Bayes factors were employed. Bayes factors measure the relative performance of two models as the ratio of their marginal likelihoods, that is, the likelihood of the data under the model (Kass and Raftery, 1995; Wasserman, 2000; Nylander et al., 2004). The marginal likelihood may be approximated by the harmonic mean of likelihood values of evolutionary hypotheses sampled from the posterior distribution (Nylander et al., 2004), which is calculated by the sump command in MrBayes (ver. 3.1.1, Ronquist and Huelsenbeck, 2003).

For the combined data set, a mixed model incorporating the models selected by Bayes factors for each of the two partitions, chloroplast and ITS data, was used.

Bayesian phylogenetic analysis—We used MrBayes (ver. 3.1.1, Huelsenbeck and Ronquist, 2001; Ronquist and Huelsenbeck, 2003) to estimate posterior probability distributions of trees and marginal likelihoods under each model. Two analyses, each with four chains, were run in parallel to better determine when convergence had been reached. The Markov chain was run for three million generations, sampling phylogenetic hypotheses every 1000 generations. Because mixing among heated chains was poor, the temperature was decreased to values between 0.07 and 0.12. For the simple models used for the ITS and cpDNA data sets, the default priors in MrBayes were used. For the

mixed model used for the combined data set, the rate prior was set to variable to allow rates to vary across partitions, and parameters applying to more than one partition, like the shape of the gamma distribution of rate variation among sites, substitution rates, and state frequencies, were unlinked to allow values to differ among partitions. Convergence of the Markov chain was assumed to be reached when plots of the overall likelihood, as well as individual parameters of the model, were fluctuating around stable values. At this point, average standard deviation of split frequencies for the two parallel runs had decreased to below 0.01 for all analyses, and the potential scale reduction factor was 1.00 for all parameters. The first 1 000 000 generations were discarded as "burnin," and the last 2000 trees were used to construct a majority rule consensus tree and calculate posterior probabilities of clades (PPs). Three independent analyses, each starting from a random tree, were performed to ensure that the Markov chain really had been sampling from the posterior distribution. This was considered to be the case if the topologies of the resulting majority rule consensus trees contained the same nodes with posterior probabilities above 0.95 and were free from supported incongruence.

Parsimony bootstrap analyses—Bootstrap analyses with 100 000 pseudoreplicates were performed on the *AAT*, chloroplast, ITS, and combined data sets with PAUP* (ver. 4.0b10, Swofford, 2002). Each pseudoreplicate was analyzed with heuristic search, creating a single start tree by random addition and improving this tree by tree-bisection-reconnection (TBR) branch swapping. One tree was saved in each pseudoreplicate and used to construct a majority rule consensus tree and calculate bootstrap proportions (BPs).

Bayesian estimation of ancestral areas—Ancestral areas were reconstructed for some nodes in the combined tree that are of interest for the biogeography of the group and also for the nodes in each of the cpDNA and nrITS trees that were incongruent regarding the ancestry of the two *Nesoluma* species. Each of these nodes was constrained to be monophyletic in a Bayesian analysis with the same settings as before, using the evolutionary model with the highest Bayes factor. The geographical distribution data were added as a separate partition that was analyzed along with the molecular data in Bayesian analyses with the same settings as before. Posterior probabilities of the different states at the node of interest were calculated based on the Markov chain Monte Carlo (MCMC) sample. This type of analysis accounts both for the uncertainty in topology and the uncertainty in mapping (Huelsenbeck and Bollback, 2001; Ronquist, 2004). The circumscription of areas was based on the biogeographical hypotheses being tested, as well as on the observed distribution patterns of species in the group and a wish to keep the number of regions as low as possible but still to have informative regions. The Macaronesian islands were included in Africa, but the Indian Ocean islands were coded as a separate area. Mexico was included in Central America, and any species occurring in Central America and one or more of the regions Caribbean islands, South America, or tropical/subtropical parts of North America were coded as Tropical America.

Bayesian molecular dating analysis—Bayesian dating (Thorne et al., 1998; Kishino et al., 2001) was performed with the multidistribute package (Thorne and Kishino, 2002) following the instructions by Rutschmann (2005). First, a Bayesian phylogenetic analysis of the cpDNA data set was performed with MrBayes ver. 3.1.1 (Huelsenbeck and Ronquist, 2001; Ronquist and Huelsenbeck, 2003) to obtain a majority rule consensus tree topology to use in the dating analysis. The same model and MCMC settings were used as in the earlier analysis of cpDNA data. Estimation of branch lengths and divergence times was based on *ndhF* data only. This was done because a few species outside Sapotaceae had to be included in the data set because the dating analysis requires a prior to be specified for the age of the root node, and the only available estimate that we could use is a published age of the Sapotaceae lineage. Based on molecular dating analyses, Bremer at al. (2004) estimated the divergence time between Sapotaceae and the group consisting of *Theophrasta*, *Myrsine*, and *Primula* to be 102 My. A potential problem with their analysis is that the r8s program (Sanderson, 2002) requires the age of at least one node to be fixed, but Bremer et al. only used fossils for calibration, which can only provide minimum ages of clades. They ran six separate r8s analyses, each fixing the age of one of the six selected fossils. Ages for individual nodes were then calculated as the mean of estimates from all six analyses. Using their estimate of the age of the Sapotaceae lineage as a prior is a potential source of error in the present analysis. *NdhF* sequences (which is the only region with sequences available in GenBank and alignable with Sapotaceae sequences) of *Barringtonia asiatica* (AF421044), *Diospyros digyna* (AF213731), *Maesa tenera* (AF213750), *Myrsine africana* (AF213751), and *Theophrasta americana* (AF213762) were included in the dating analyses. Trees were rooted with *Barringtonia* and *Diospyros.* The program Baseml in the PAML package (ver. 3.13d, Yang, 1997) was used to estimate model parameters under the F84 model (Felsenstein, 1984), and the program estbranches (ver. 8/5/03, Thorne and Kishino, 2002) was used to estimate branch lengths. In the multidivetime (ver. 9/25/03, Thorne and Kishino, 2002) analysis that estimates divergence times, the mean of the prior distribution for the rate of molecular evolution at the ingroup root node (rtrate) was set to 0.011 with a standard deviation of 0.011. The mean of the prior distribution for nu (brownmean) was set to 0.039 with a standard deviation of 0.039. The value of "bigtime," which should be substantially higher than the expected number of time units between tip and root, was set to 150 My, which is older than the age of the oldest unambiguous angiosperm fossil. For other parameter settings, the default values were used.

Three fossils were used to calibrate the dating analysis. A fossil described from a specific geologic time period and attributed to a certain group was used to set a minimum age of the crown group of that clade. We chose to assign the fossil to the crown group rather than the stem because the fossil has some characters that distinguish one group from its sister, and these characters were not yet present in their common ancestor at the stem node. Compared to assigning the fossil to the stem, this will give older age estimates. It has been argued that the only logically consistent way of assigning fossil calibration points to phylogenies is to use them as minimum ages of stem nodes (Renner, 2005). This more conservative approach will, however, lead to a consistent underestimation of divergence times. The underestimation is caused first, by fossils always being younger than the taxa they represent, and second, by fossils being assigned to nodes that are too old. The placement of fossil calibration points is difficult because there is no way of knowing where in the phylogeny the fossil belongs. It may even represent a branch inside the crown group. Recent studies commonly constrain the age of the stem node of a clade sharing some characters present in the fossil (e.g., Davis et al., 2004; Richardson et al., 2004; Pirie et al., 2006), although in some studies constraints are instead placed on the age of the crown group (e.g., Zerega et al., 2005; Muellner et al., 2006). The first calibration point that we used was based on fossil pollen of Sideroxyleae described from the early Eocene (Ypresian) of England (Gruas-Cavagnetto, 1976). This was used to set a minimum age (49 My) of node P (Figs. 1 and 2). As a second calibration point, fossil leaves of *Sideroxylon aequale* described from the middle of the Oligocene of Puerto Rico (Hollick, 1928) were used. Hollick (1928) noted that these leaves are similar to those of the extant species *Sideroxylon foetidissimum*. In our analyses, *S. foetidissimum* is found in a clade together with three Central American species (Figs. 1 and 2, node E1) that all have leaves with similar outline and venation pattern, but which differ from those of the remaining species in clade J. Therefore, we used this fossil to set a minimum age (28.4 My) of node E1 (Figs. 1 and 2). The final calibration point was fossil leaves of *Bumelia retusafolia* described from the middle of Miocene of Cuba (Berry, 1939). According to Berry (1939), these leaves resemble those of *Sideroxylon americanum* and *S. horridum*, which are found in clade K1 (Figs. 1 and 2) together with three other Caribbean species that all share the same type of leaves. Therefore, this fossil was used to assign a minimum age (14.2 My) to node K1 (Figs. 1 and 2).

The Markov chain in the program multidivtime was run for three million generations, discarding the first one million as burnin and then sampling every 100 generations. Three analyses with different initial seed numbers were run to ascertain that the MCMC was sampling from the posterior distribution and that different analyses produced the same results.

RESULTS

Molecular data—DNA sequences ranged from 400 to 827 base pairs (bp) for the ITS region, from 869 to 1064 bp for the partial *ndhF* gene, from 249 to 572 bp for the *trnH-psbA* spacer, from 683 to 789 bp for the *trnC-petN* spacer, from 927 to 1271 bp for the *petN-psbM* spacer, and from 950 to 1152 bp for the *psbM-trnD* spacer. Of the *AAT* gene, a fragment ranging from 585 to 793 bp containing two partial and one entire exon and two introns was used. One of the *AAT* copies from *Nesoluma nadeaudii* (clone 1) and two from *N. polynesicum* (clones 2 and 6) are probably pseudogenes because they have a single base-pair deletion in the exon sequence and contain

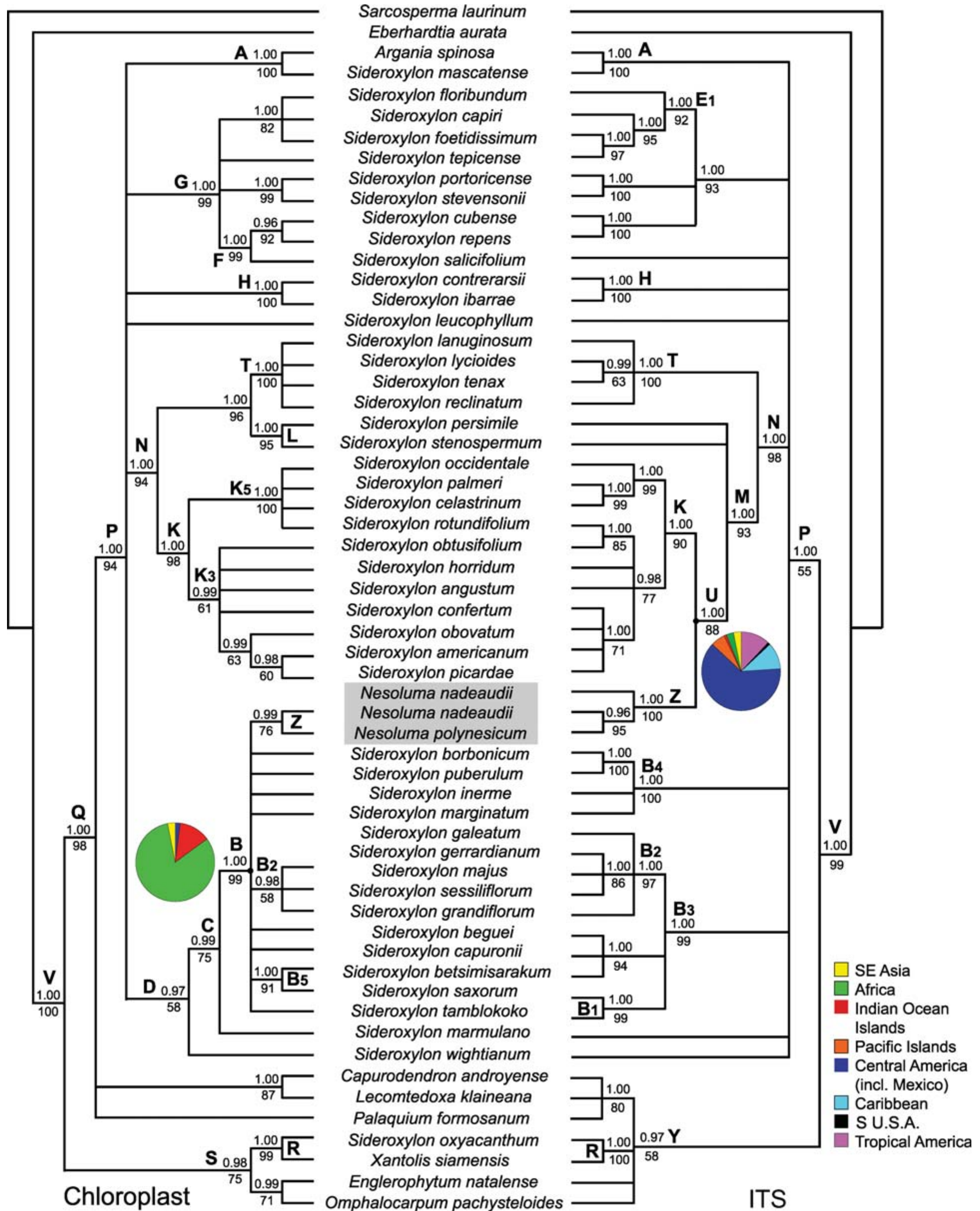

Fig. 1. The 95% majority rule consensus trees from the Bayesian phylogenetic analyses of the cpDNA and ITS data. Posterior probabilities of clades are marked above branches and parsimony bootstrap proportions are marked below. Letters refer to clades mentioned in the text. Posterior probabilities of ancestral areas for clades B and U are shown as pie charts. The shaded box marks a strongly supported incongruency between the two trees.

Species	Distribution
Sarcosperma laurinum	SE Asia
Eberhardtia aurata	SE Asia
Argania spinosa	N Africa
Sideroxylon mascatense	NE Afr + SW Asia
Sideroxylon capiri	C America
Sideroxylon foetidissimum	S U.S.A. + C Am + Caribbean
Sideroxylon tepicense	C America
Sideroxylon floribundum	C America
Sideroxylon portoricense	C America
Sideroxylon stevensonii	C America
Sideroxylon cubense	Caribbean
Sideroxylon repens	Caribbean
Sideroxylon salicifolium	S U.S.A. + C Am + Caribbean
Sideroxylon contrerarsii	C America
Sideroxylon ibarrae	C America
Sideroxylon americanum	C Am + Caribbean
Sideroxylon obovatum	Caribbean
Sideroxylon picardae	Caribbean
Sideroxylon confertum	Caribbean
Sideroxylon horridum	Caribbean
Sideroxylon angustum	Caribbean
Sideroxylon obtusifolium	Tropical Am
Sideroxylon celastrinum	Tropical Am
Sideroxylon palmeri	C America
Sideroxylon occidentale	C America
Sideroxylon rotundifolium	Caribbean
Sideroxylon persimile	Tropical Am
Sideroxylon stenospermum	C America
Sideroxylon lanuginosum	S U.S.A.
Sideroxylon lycioides	S U.S.A.
Sideroxylon tenax	S U.S.A.
Sideroxylon reclinatum	S U.S.A.
Sideroxylon leucophyllum	C America
Sideroxylon beguei	Indian Oc. Islands
Sideroxylon capuronii	Indian Oc. Islands
Sideroxylon betsimisarakum	Indian Oc. Islands
Sideroxylon saxorum	Indian Oc. Islands
Sideroxylon tamblokoko	Indian Oc. Islands
Sideroxylon grandiflorum	Indian Oc. Islands
Sideroxylon majus	Indian Oc. Islands
Sideroxylon sessiliflorum	Indian Oc. Islands
Sideroxylon borbonicum	Indian Oc. Islands
Sideroxylon puberulum	Indian Oc. Islands
Sideroxylon inerme	E Africa
Sideroxylon marginatum	NW Africa
Sideroxylon marmulano	NW Africa
Sideroxylon wightianum	SE Asia
Capurodendron androyense	Indian Oc. Islands
Lecomtedoxa klaineana	Africa
Palaquium formosanum	SE Asia
Sideroxylon oxyacanthum	Africa
Xantolis siamensis	SE Asia
Englerophytum natalense	Africa
Omphalocarpum natalense	Africa

Fig. 2. The 70% majority rule consensus tree from Bayesian phylogenetic analysis of the combined cpDNA and ITS data set. Posterior probabilities of clades are marked above branches and parsimony bootstrap proportions are below. Letters refer to clades mentioned in the text. Geographical distribution is given behind each species name.

TABLE 1. Evaluated evolutionary models with number of free parameters, model likelihood, and Bayes factors (BF). A Bayes factor of 3 to 20 indicates positive evidence against a less parameter rich model, >20 to 150 strong evidence, and above 150 very strong evidence (Kass and Raftery, 1995).

Model		Free parameters	Model likelihood	BF > 1	BF > 2
A) ITS region					
1	HKY+Γ	5	−8021	—	
2	GTR+I+Γ	10	−8006	15	
B) cpDNA (*ndhF*, *trnC-trnD* region, and *trnH-psbA*)					
1	GTR+Γ	9	−13 587	—	
2	GTR+Γ+COV	11	−13 535	52	
C) cpDNA and ITS combined (excluding *Nesoluma*)					
1	GTR+Γ	9	−22 263	—	—
2	GTR+Γ+COV	11	−22 124	139	—
3	GTR+Γ+COV/GTR+I+Γ	21	−21 504	759	620

several stop codons. EMBL accession numbers of all DNA sequences are shown in Appendix 1. The ITS matrix consisted of 942 characters and 59 terminals and included 2.7% missing data. It contained 466 variable characters of which 308 were parsimony informative. All six of the cloned ITS sequences from *N. polynesicum* were identical, while two different sequence types were found in *N. nadeaudii*. The duplicate sequences were omitted from the matrix so that only unique sequences were included. The cpDNA matrix consisted of 5097 characters and 56 taxa and included 12.9% missing data. It contained 724 variable characters of which 279 were parsimony informative. The combined ITS and cpDNA matrix consisted of 6039 aligned characters and 54 taxa and included 9.9% missing data. Of the 1171 characters that were variable, 566 were parsimony informative. The *AAT* matrix consisted of 834 characters and 70 terminals and contained 0.4% missing data. Of the 289 characters that were variable, 159 were parsimony informative.

Model selection—For the *AAT* data set, AICc and BIC both selected HKY+Γ, and therefore this model was used in the phylogenetic analyses of this data set. For the ITS data set, AICc selected GTR+I+Γ and BIC selected HKY+Γ. The performance of these two models was evaluated with Bayes factors, which showed positive evidence in favor of GTR+I+Γ over HKY+Γ (Table 1). For the cpDNA data set, both criteria favored GTR+Γ. When this model, which only allows rates to vary across sites, was compared to a covarion model, which allows rates to vary over time (Fitch and Markowitz, 1970; Fitch, 1971; Tuffley and Steel, 1998), Bayes factors strongly favored GTR+Γ+COV over GTR+Γ (Table 1). When the combined cpDNA and ITS data were treated as one homogeneously evolving entity, AICc and BIC both selected GTR+Γ. This model was compared to a simple covarion model (GTR+Γ+COV) and to a mixed model, for which the cpDNA and ITS data sets had their own set of parameters. The mixed model consisted of the best models for the two partitions (GTR+Γ+COV for cpDNA, and GTR+I+Γ for ITS). Bayes factors strongly favored the mixed model over both of the simple models (Table 1).

Phylogeny—The majority rule consensus trees from the Bayesian phylogenetic analyses of the cpDNA and ITS data (Fig. 1) are largely congruent, but there are some well supported (PPs > 0.95 and BPs > 70) contradictions between them. For example, ITS data resolves Sideroxyleae (Fig. 1, node P) as the sister group of a clade consisting of members of Chrysophylloideae together with representatives of the remainder of Sapotoideae (Fig. 1, node Y). This relationship, however, is not strongly supported by parsimony. This topology differs from that based on cpDNA, where Sapotoideae (Fig. 1, node Q) and Chrysophylloideae (Fig. 1, node S) are sister groups and Sideroxyleae is in Sapotoideae (see also Anderberg and Swenson, 2003; Swenson and Anderberg, 2005; Smedmark et al., 2006). ITS data analyzed with model-based methods also consistently give this different topology when the data include a more comprehensive taxon sample from Sapotoideae and Chrysophylloideae (results not shown). The two data sets also support incongruent relationships of a number of taxa within Sideroxyleae, for example the two *Nesoluma* species (nodes Z). In the cpDNA tree, they are nested within clade B (Fig. 1) while ITS data places them in clade U (Fig. 1). The two ITS alleles of *N. nadeaudii* are found to be paraphyletic with respect to the single one in *N. polynesicum*. Other incongruent relationships are those of *Sideroxylon tepicense*, which is nested within clade E1 in the ITS tree but not in the cpDNA tree, and *S. salicifolium*, which is nested within clade G in the cpDNA tree but not in the ITS tree. Also, *S. persimile* and *S. stenospermum* are the sister group of clade T in the cpDNA tree but not in the ITS tree where they are instead more closely related to clade U. *Sideroxylon rotundifolium* is in clade K5 in the cpDNA tree but not in the ITS tree, and *S. saxorum* is in clade B5 in the cpDNA tree but in clade B1 in the ITS tree.

The majority rule consensus tree from the Bayesian phylogenetic analysis of the *AAT* data set is shown in Fig. 3. For three of the species for which cpDNA and ITS data supported different relationships, *Sideroxylon tepicense* (Fig. 3, node e1), *S. salicifolium* (Fig. 3, node f1), and *S. rotundifolium* (Fig. 3, node k1), the different *AAT* clones form monophyletic groups. For two of the incongruent species, *S. persimile* (Fig. 3, node W) and *S. saxorum* (Fig. 3, node B5), the different *AAT* alleles are paraphyletic with respect to alleles from other species. In the case of *Nesoluma*, alleles occur in two strongly supported clades (Fig. 3, nodes Z′ and Z″). There are also two species, *S. obtusifolium* and *S. occidentale*, for which the results based on cpDNA and ITS were congruent, but which both have *AAT* alleles in two different clades (Fig. 3, *d).

The majority rule consensus trees from the Bayesian phylogenetic analysis of the combined cpDNA and ITS data set (Fig. 2) does not resolve the earliest divergences within Sideroxyleae with good support. The interrelationships among five clades (Fig. 2, nodes A, C, G, H, and N) and two species

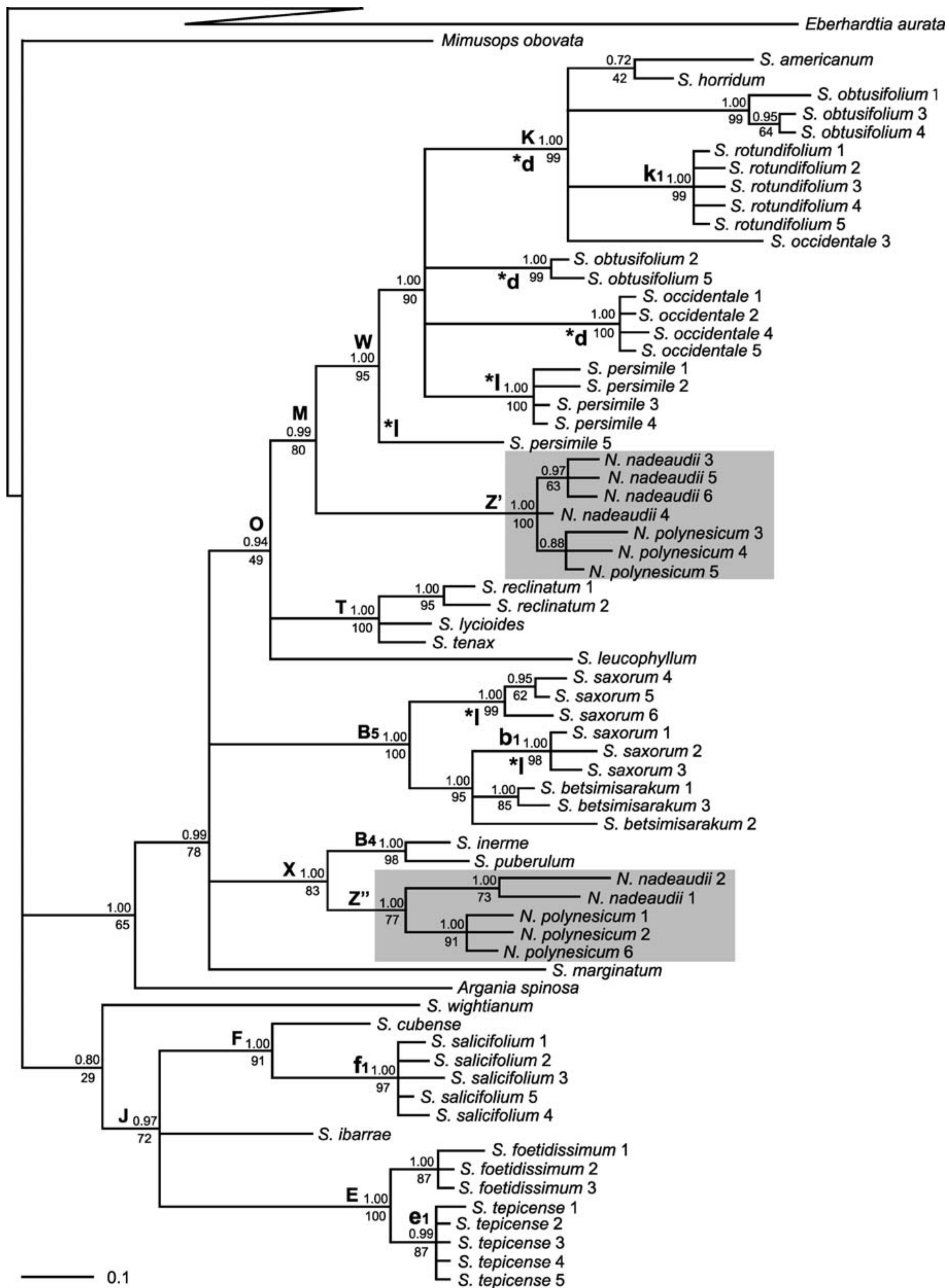

Fig. 3. The 70% majority rule consensus tree from Bayesian phylogenetic analysis of the *AAT* data set. Branch lengths are proportional to the amount of change. Posterior probabilities of clades are marked above branches and parsimony bootstrap proportions are below. Letters refer to clades mentioned in the text. Numbers behind species names are clone identifiers. The shaded boxes mark the different paralogues of the *Nesoluma* species.

TABLE 2. Posterior probabilities of eight ancestral areas for selected nodes.

Node	SE Asia	Africa	Indian Ocean	C Am	Caribbean	S USA	Trop Am	Pacific
A) cpDNA and ITS combined (Fig. 2)								
C	0.08	0.79	0.05	0.08	0.00	0.00	0.00	
D	0.12	0.72	0.03	0.13	0.00	0.00	0.00	
F	0.04	0.07	0.01	0.48	0.34	0.01	0.05	
F1	0.00	0.00	0.00	0.01	0.98	0.00	0.00	
G	0.01	0.07	0.00	0.92	0.00	0.00	0.00	
J	0.00	0.06	0.00	0.93	0.01	0.00	0.00	
K	0.00	0.00	0.00	0.29	0.67	0.00	0.04	
M	0.03	0.07	0.01	0.67	0.11	0.01	0.10	
N	0.04	0.15	0.01	0.71	0.03	0.03	0.03	
O	0.03	0.16	0.01	0.77	0.01	0.01	0.01	
P	0.10	0.68	0.01	0.21	0.00	0.00	0.00	
B) cpDNA (Fig. 1)								
B	0.03	0.82	0.13	0.02	0.00	0.00	0.00	0.00
C) ITS (Fig. 1)								
U	0.03	0.03	0.01	0.63	0.11	0.01	0.12	0.06

(Fig. 2, *Sideroxylon wightianum* and *S. leucophyllum*) are poorly resolved. Two of the clades are strongly supported American groups; one clade includes species that have been classified in *Dipholis* A.DC. and *Mastichodendron* (Engler) Lam (Fig. 2, node G, PP1.00, BP98), and the other clade (Fig. 2, node N, PP1.00, BP 100) includes species that have been classified in *Bumelia* Sw. (Aubréville, 1964). The African species are in two groups, one includes *Argania* and *S. mascatense* (Fig. 2, node A, PP 1.00, BP 100) and the other *Sideroxylon* s.s. (Fig. 2, node C, PP 1.00, BP 75) sensu Aubréville (1964). The combined analysis, like the separate analyses of cpDNA and ITS data (Fig. 1), shows that *S. oxyacanthum* is the sister of *Xantolis*. It is evident that this species does not belong in Sapotoideae but in the other large Sapotaceae clade, Chrysophylloideae. Despite the different sampling strategy, *AAT* data support several of the clades present in the tree from the combined analysis of ITS and cpDNA data (Fig. 3, nodes B5, E, F, I, K, M, O, T, Z).

Reconstruction of ancestral areas—Posterior probabilities of ancestral areas for some nodes are shown in Table 2. The geographical origin of Sideroxyleae is not resolved with strong support, but Africa is the region that gets the highest PP (0.68) for this node (Fig. 2, node P). Because basal nodes within Sideroxyleae are not strongly supported (Fig. 2, nodes D, J, O), and the analysis takes this topological uncertainty into account, there is not strong support for one single area for any of these nodes. Two areas, however, get higher PPs than any of the others for these three nodes, and these are Africa and Central America (Table 2). For both of the strongly supported clades *Bumelia* (Fig. 2, node N) and *Dipholis* + *Mastichodendron* (Fig. 2, node G), the analysis shows the highest PP (Table 2) for a most recent common ancestor in Central America, while the *Sideroxylon* s.s. clade (Fig. 2, node C) probably originated in Africa (Table 2).

For the most recent common ancestor of *Nesoluma* and its sister group in the ITS tree (Fig. 1, node U), the American regions get a total PP of 0.87, while Central America is the region that gets the highest PP (0.63). In the cpDNA tree (Fig. 1), *Nesoluma* is in a strongly supported clade (node B) consisting largely of species from the Indian Ocean Islands. Relationships within this clade are not resolved with good support, and the sister of *Nesoluma* is not known, but the ancestral area with the highest PP for the entire clade (Fig. 1, node B) is Africa (PP 0.82).

Bayesian molecular dating—Times of selected nodes, estimated as the mean of the posterior distribution, are given in Table 3. The uncertainty of age estimates, as measured by the 95% credibility intervals, is large (Table 3), probably because of the small amount of information in the *ndhF* data set.

If a hybridization has taken place in the evolutionary history of *Nesoluma*, as implied by the separate analyses of ITS and cpDNA, as well as by *AAT* data, the *Nesoluma* lineage should have diverged from its two parental lineages at the time of the hybridization. Therefore, we should expect *Nesoluma* to have split from its sister group at about the same time in each of the cpDNA and ITS trees. The most recent common ancestor of the *Nesoluma* clade and its American parental lineage is inferred to be older than node K (36.6 My) because *Nesoluma* is the sister

TABLE 3. Ages and credibility intervals of some nodes in Figs. 1 and 2 resulting from Bayesian molecular dating analysis (Thorne et al., 1998; Kishino et al., 2001) of *ndhF* data.

Node	Mean age (My)	95% Credibility interval (My)
A	43.0	13.0–77.0
B	43.0	22.2–68.8
C	52.4	32.0–78.4
D	57.6	38.0–83.8
F	21.2	4.0–44.4
F1	11.7	0.6–33.3
G	43.8	30.6–65.8
J	51.4	35.2–75.8
K	36.6	19.0–61.4
K_3	29.2	15.0–52.0
K_4	8.8	0.28–27.4
L	16.4	0.8–43.0
N	45.2	25.8–70.2
O	53.2	34.4–78.6
P	68.4	51.2–95.0
Q	74.0	54.0–101.6
T	26.2	7.6–50.6
V	79.8	57.6–108.2
Sapotaceae	98.8	69.8–128.6

of this clade in the ITS tree (Fig. 1). The analysis also indicates that this common ancestor was younger than node N (45.2 My), the next node one step down in the cpDNA tree (Fig. 1). Because the dating was based on cpDNA data, we cannot get a date for node U (Fig. 1), which is only present in the ITS tree and which would have provided a more accurate time estimate. The African ancestor of *Nesoluma*, on the other hand, should not be older than node B (43.0 My) in the cpDNA tree (Fig. 1) because *Nesoluma* belongs somewhere in this incompletely resolved clade. Based on these results, we infer that the *Nesoluma* lineage originated 43.0 to 36.6 Mya.

DISCUSSION

Our data indicate that many extant groups within Sideroxyleae began to diversify during a period (ca. 65–34 Mya) when the boreotropical flora is thought to have prevailed (Table 3, nodes A–C, D, G, J, K, N, O). While the estimation of ancestral areas gives inconclusive support for an Old World origin of Sideroxyleae, the analysis could not have pointed to an origin in the northern hemisphere because no extant species occur in this region. The difficulty in resolving early relationships within the group may result from a rapid radiation in the northern hemisphere in the early Tertiary, followed by extensive extinctions due to the deterioration of the climate that marked the end of the boreotropical flora. As we will argue in the following paragraphs, this study shows that the distribution of extant species in Sideroxyleae, like that of several other tropical and subtropical groups of angiosperms (e.g., Melastomataceae, Renner et al., 2001; Malpighiaceae, Davis et al., 2004; Burseraceae, Weeks et al., 2005; Moraceae, Zerega et al., 2005; Meliaceae, Muellner et al., 2006), is primarily the result of a gradual migration through the boreotropics followed by vicariance (caused by the barrier posed by the cooling and drying climate) rather than long-distance dispersal.

There are two American clades (Fig. 2, nodes J and O), which may represent one colonization of the New World or two separate colonization events. In the former case, the two clades would be sister groups or paraphyletic with respect to some other clade. In the latter case, the clades would be more closely related to other lineages within the group than to each other. The age estimates of these clades are, however, very similar. They are inferred to have begun to diversify in America 51.4 (Fig. 2, node J) and 53.2 Mya (Fig. 2, node O), which would make long-distance dispersal from Africa to America unlikely. The African and South American plates separated 96 Mya, and dispersals across the southern Atlantic continued into the earliest Paleogene (ca. 65 Mya), as evidenced from the simultaneous occurrence of pollen from the same taxa in the fossil record of Africa and South America (Morley, 2003). The credibility intervals of time estimates for these nodes are large (Table 3, node J, 35.2–75.8 My, node O, 34.4–78.6 My) and do not rule out dispersals across the southern Atlantic via island hopping later than 96 Mya. Had Sideroxyleae entered the New World by this route, we would, however, have expected to see South American taxa being sisters or paraphyletic with respect to the Central American, Caribbean, and North American taxa. There are, however, no endemic South American species in Sideroxyleae. Three *Sideroxylon* species occur in South America (*S. celastrinum*, *S. obtusifolium*, and *S. persimile*), but these are also present in Central America, Mexico, the Caribbean, and in one case also in the southern U.S. They are in three different clades (Fig. 2, nodes K3, K4, and L) and all represent younger lineages (Table 3, 29.2, 8.8, and 16.4 My, respectively), well-nested within one of the neotropical clades (Fig. 2, node O) that is inferred to have begun to diversify about 53.2 Mya (Table 3). That the species occurring in South America are comparatively young and also occur further north points to a colonization from the north. The extant radiations of the two New World clades (Fig. 2, clades J and O) are inferred to have begun in Central America and Mexico (Table 2), a region where North American boreotropical taxa found refuge (Morley, 2000), which further supports the hypothesis that Sideroxyleae entered the New World from the north rather than from the south. Because of the poor resolution, we cannot date the divergence between the Old and the New World, which should have taken place between 60 and 34 Mya if Sideroxyleae were part of the boreotropical flora. During this period, climate and high latitude land bridges made crossings between Eurasia and North America possible for subtropical as well as temperate taxa (Wolfe, 1975; Tiffney, 1985b). The inferred ages of the American clades (51.4 and 53.2 My) are, however, consistent with a crossing via a north Atlantic land bridge.

Our analyses suggest that the islands of the Caribbean were colonized at least twice from taxa in Central America. Clade O (Fig. 2, *Bumelia* s.l.) is inferred to have begun to diversify in Central America, and the younger ancestors at nodes N and M also most likely occurred in Central America (Table 2). Clade K, on the other hand, which is nested within clade O, probably had a common ancestor in the Caribbean islands (PP 0.67), indicating a colonization of this region. The age estimate of clade K is 36.6 My (Table 3). The other American clade is likewise inferred to have undergone its earliest bifurcations in Central America (Table 2, nodes J and G). Nested inside this group is the small clade F1 (Fig. 2), whose common ancestor occurred in the Caribbean (PP 0.98) about 11.7 Mya (Table 2). The geologic history of the region is complex, and there has been much controversy on the origin of the Caribbean islands (Morley, 2003). It does, however, seem that North and South America in the late Cretaceous were more or less connected by an island arc or land bridge, which moved east as the front of the Caribbean plate pushed through the opening between the two continents (Pindell et al., 1988). The eastward movement ultimately caused the Caribbean plate to separate from Yucatan ca. 49–39 Mya. Thirty-five Mya, tectonic uplift and a drop in sea level is thought to have exposed a land mass consisting of the Greater Antilles and the Aves ridge (GAARlandia, Iturralde-Vincent et al., 1999) that was connected to northern South America. The connection was again submerged subsequent to 33 Mya, and the existing Greater Antilles formed. Prior to this date, land had existed in different shapes at different times between North and South America, but the islands currently present in the Caribbean have only been continuously emergent from this time (Iturralde-Vincent et al., 1999). This means that the species that exist there today cannot have come earlier than 35 Mya, which is close to the inferred age of the Caribbean clade K (36.6 My) and may indicate that the origin of the modern Caribbean islands played an important role in the diversification of the group in this region. The other Caribbean clade (F1) is inferred to be much younger and thus represents a later colonization event of the Caribbean islands.

Although the cpDNA and nrITS phylogenies supported different ancestries of several taxa (discussed earlier), the nuclear low copy *AAT* phylogeny only confirmed one hybrid

origin, namely for *Nesoluma*. The two investigated species of *Nesoluma* both have maternally inherited cpDNA from a group now distributed in Africa and the Indian Ocean islands (Fig. 1, node B), whereas their nrITS, which is biparentally inherited but usually undergoes concerted evolution, is the sister of an American clade (Fig. 1, node K). The presence of two ITS alleles in *N. nadeaudii* indicates that homogenization of different ITS copies, which may originally have come from different parent species in a hybridization event, has not been complete. In the analysis, the two *N. nadeaudii* ITS alleles are paraphyletic with respect to the single one in *N. polynesicum*, which may be a sign of recent speciation or unclear species boundaries. The conflicting relationships supported by ITS and cpDNA data are strongly supported but difficult to explain, not only because it would require hybridization between geographically disjunct groups, but also because *Nesoluma* itself is restricted to some remote islands in the eastern Pacific, where no representative of either of the putatively hybridizing lineages occur today. In the nuclear low copy *AAT* gene phylogeny, each of the two *Nesoluma* species has two distinct copies of *AAT* (Fig. 3, nodes Z′ and Z″). One is the sister of the two African species *Sideroxylon inerme* and *S. puberulum* (Fig. 3, node X), which are in the same clade as *Nesoluma* in the cpDNA tree (Fig. 1, node B). The other copy is the sister of a large American clade (Fig. 3, node W), which includes species in its sister group in the ITS tree (Fig. 1, clade K) and also *S. persimile*. Thus the *AAT* phylogeny, despite being more scarcely sampled in terms of species, supports the same ancestry of the two *Nesoluma* species. That the two *AAT* copies have likely been retained in the genome for millions of years indicates that these species have two separate *AAT* loci. The simplest explanation for the retention of two distinct gene copies of different descent in the genome of a plant is allopolyploidy. Had homoploid hybridization rather than allopolyploidy taken place, recombination would have erased the information about the origin of the individual alleles. Unfortunately, we do not have any information about the ploidy levels for the taxa involved and therefore cannot be certain that the *Nesoluma* species are polyploids. We do, however, conclude that our data provide indirect evidence for this. The use of nuclear low copy genes for phylogeny reconstruction requires some caution because they may be affected by for example duplications, deletions, or lineage sorting, which may make homology assessments difficult (Doyle and Davis, 1998). The *AAT* phylogeny presented in this study contains at least one possible duplication (Fig. 3, *d) and two possible cases of lineage sorting (Fig. 3, *l). The two paralogues in each of *Sideroxylon obtusifolium* and *S. occidentale* may indicate that a gene duplication took place in the common ancestor of clade K (Fig. 2), while the paraphyly of copies in *S. persimile* (Fig. 3, node W) and *S. saxorum* (Fig. 3, node B5) may have been caused by lineage sorting.

The hybridization that is inferred to have coincided with the origin of *Nesoluma* is estimated to have occurred between 43.0 and 36.6 Mya (discussed earlier), which is at the end of the Eocene when the boreotropical flora was still prevalent in the northern hemisphere (Wolfe, 1975; Tiffney, 1985a). This indicates that the stem lineage of the American clade K (Fig. 2) and some representative of the African clade B (Fig. 2) coexisted in the same area, probably at high latitudes, before the drastic cooling that took place around the Eocene/Oligocene boundary (34 Mya) when tropical elements were forced toward the equator because of the decreased temperatures (Wolfe, 1978, 1980; Collinson et al., 1999). The large credibility intervals of the time estimates for nodes B (22.2–68.8) and K (19.0–61.4) do not eliminate the possibility that the hybridization could have taken place after the period of warm climate had ended. They do, however, indicate that it did not happen in the location where *Nesoluma* exists today. The three species are all found on relatively young islands in the Pacific Ocean. *Nesoluma polynesicum* is found in Maui-Nui, Oahu, and Kauai of the Hawaiian Islands, which are of volcanic origin and between 1.9 and 5.1 My old (Carson and Clague, 1995). This species is also present in two other volcanic islands much further South, Rapa Iti and Raivavae of the Austral Islands, which are 5.1 and 6.5 My of age (Bonneville et al., 2002). *Nesoluma nadeaudii* is found in Tahiti, which is about 1.4 My old (Hildenbrand et al., 2004). The third species, *N. st.-johnianum* is an endemic of Henderson Island, which is an uplifted coral atoll of the Pitcairn Islands that is only 380 000 yr old (Blake, 1995). Thus, our analyses indicate that the *Nesoluma* lineage arose about 30 My before the formation of any of the islands where the group occurs today. Most islands in the eastern Pacific formed as a result of hot spot activity (Wilson, 1963). Islands have existed continuously in this region, and new islands are formed successively at the hot spot as older islands gradually become submerged. It is possible that the *Nesoluma* lineage has existed on other islands prior to the ones where it is found today. A similar example is found in Begoniaceae, where the Hawaiian *Hillebrandia* lineage predates the currently existing Hawaiian Islands by about 20 million yr (Clement et al., 2004). There is evidence that *Hillebrandia*, which also has been suggested to be of boreotropic origin, has survived by island hopping (Clement et al., 2004).

Taxonomic considerations—Our phylogenetic results, showing that *Sideroxylon oxyacanthum* does not belong in *Sideroxylon* or in the more inclusive clade Sapotoideae, confirm Friis' (1978) observation that this species is probably not closely related to *S. mascatense* because they differ in several aspects of their morphology. Aubréville (1963) placed both *S. oxyacanthum* and *S. mascatense* in the genus *Spiniluma*. They are both spiny shrubs or trees that occur in Ethiopia and Somalia (*S. oxyacanthum*) or Ethiopia to Pakistan (*S. mascatense*), respectively. Friis (1978) noted that *S. oxyacanthum* and another species (*S. discolor*) differ from other species in *Sideroxylon* s.l. in having well-developed brachyblasts subtended by leafless spines, which we interpret to be transformed leaves. According to Friis, *S. discolor* is very similar to *S. oxyacanthum* and may even be conspecific. Unfortunately, *S. discolor* was not available for study. A future phylogenetic study, including this species, will show whether these two species do indeed form a group. In our analysis, *S. mascatense* is the sister of *Argania spinosa*, which occurs in northwestern Africa. These two species, like other *Sideroxylon* species, have thorns that are modified branches, developing in the axils of leaves and often carrying leaves and flowers in mature plants. Our results show that *S. oxyacanthum* is the sister of *Xantolis siamensis* (Figs. 1 and 2, node R) and thus member of a different lineage than *Sideroxylon*, i.e., Chrysophylloideae. *Xantolis* has thorns similar to those of *S. mascatense* and *Argania spinosa* and shares some plesiomorphic characters, such as number and arrangement of floral parts, with *S. mascatense*. *Sideroxylon oxyacanthum* is the type

species of *Spiniluma*, a genus name that may again be considered until the phylogenetic affinities within Chrysophylloideae are clarified.

The phylogenetic results from this study reveal problems with the current classification (Pennington, 1991), in which *Sideroxylon* is paraphyletic with respect to *Argania* and *Nesoluma*, but also with the previous classifications that recognized several smaller genera (e.g., Aubréville, 1964). The American clades are the most problematic. In the current study there is no support for a single, large, American clade. Instead there are two poorly supported clades (Fig. 2, nodes J and O). One of them includes *Bumelia* (Fig. 2, node N), which has pubescent ovaries and seeds with thin endosperm, and *S. leucophyllum*, which has divided corolla lobes and hairy ovaries like *Bumelia*, but differs in having seeds with a fleshy endosperm. The other American clade includes species that have been classified in *Dipholis* and *Mastichodendron*, which all have seeds with copious endosperm (Fig. 2, node G), and a group consisting of the two species, *Sideroxylon contrerasii* and *S. ibarrae*. All species in this clade (Fig. 2, node J) have glabrous ovaries. The delimitation of *Dipholis* and *Mastichodendron* is, however, somewhat unclear. In our analyses, *S. stevensonii*, which has been placed in *Dipholis* (Aubréville, 1964) but has a wood structure similar to that of *Mastichodendron* (Pennington, 1991), is the sister of *S. portoricense*, which has been classified in *Mastichodendron* (Aubréville, 1964). *Sideroxylon contrerasii* and *S. ibarrae* both have a long and broad adaxial seed scar but otherwise differ in their vegetative morphology (Pennington, 1991). We find no other way of recognizing genera in Sideroxyleae that are strongly supported to be monophyletic and morphologically distinct, than to include *Nesoluma* and *Argania* in *Sideroxylon*. Both *Argania* and *Nesoluma* are merely apomorphic derivatives with ancestors in *Sideroxylon*, and *Nesoluma* is of hybrid origin, descended from two different ancestral lineages within *Sideroxylon*. The single species of *Argania*, *A. spinosa* (L.) Skeels, was actually named *Sideroxylon spinosum* by Linnaeus in 1753, but for the three species of *Nesoluma* (*N. polynesicum*, *N. nadeaudii*, and *N. st.-johnianum*) formal recombinations under *Sideroxylon* are needed.

New combinations—*Sideroxylon polynesicum* (Hillebr.) Smedmark & Anderb. **comb. nov**.

Basionym: *Chrysophyllum polynesicum* Hillebr., Fl. Haw. Isl. 277 (1888).—Synonym: *Nesoluma polynesicum* (Hillebr.) Baill.

Sideroxylon st.-johnianum (H. J. Lam & Meeuse) Smedmark & Anderb. **comb. nov**.

Basionym: *Nesoluma st.-johnianum* H. J. Lam & Meeuse, Occas. Papers Bishop Mus., Honolulu 14: 153 (1938).

Sideroxylon nadeaudii (Drake) Smedmark & Anderb. **comb. nov**.

Basionym: *Palaquium nadeaudii* Drake, Fl. Polynes. Franc 120 (1891); Illustr. Ins. Mar. Pacif. 229 (1892). - Synonym: *Nesoluma nadeaudii* (Drake) Pierre ex H. J. Lam.

Conclusions—We conclude that the disjunct pantropical distribution of Sideroxyleae is a relict of a continuous range in the northern hemisphere during a period of warm climate in the early Tertiary. Extant species are found, e.g., in Central America, Mexico, southeast Asia, China, and Macaronesia, regions known to have acted as refugia for boreotropical taxa that migrated south in response to the cooling climate (Morley, 2000). The group most likely entered the New World from the north during the Eocene, via the north Atlantic land bridge, which allowed tropical and subtropical taxa to extend their ranges between Eurasia and North America. The Pacific clade *Nesoluma* is shown to have originated in a hybridization between an African and an American ancestral lineage during the time of the boreotropical flora. This indicates a common distribution range, probably in the northern hemisphere, for the two parental lineages as well as the hybrid at this time. This is one of few documented allopolyploids to be demonstrated in a tropical plant group. We also conclude that the circumscription of *Sideroxylon* should be amended to include the two small genera *Argania* and *Nesoluma*, both of which have evolved from ancestors in *Sideroxylon*, and that *Sideroxylon oxyacanthus* does not belong in *Sideroxylon* but is a close relative of *Xantolis* in the Chrysophylloideae.

LITERATURE CITED

AKAIKE, H. 1973. Information theory as an extension of the maximum likelihood principle. *In* B. N. Petrov and F. Csaki [eds.], Second international symposium on information theory, 267–281. Akademiai Kiado, Budapest, Hungary.

ANDERBERG, A. A., AND U. SWENSON. 2003. Evolutionary lineages in Sapotaceae (Ericales): a cladistic analysis based on *ndhF* sequence data. *International Journal of Plant Sciences* 164: 763–773.

ARNHEIM, N., M. KRYSTAL, R. SCHMICKEL, G. WILSON, O. RYDER, AND E. ZIMMER. 1980. Molecular evidence for genetic exchanges among ribosomal genes on nonhomologous chromosomes in man and apes. *Proceedings of the National Academy of Sciences, USA* 77: 7323–7327.

AUBRÉVILLE, A. 1963. Notes sur les Sapotacées. *Adansonia, new series* 3: 19.

AUBRÉVILLE, A. 1964. Sapotacées. *Adansonia*, Mémoire no. 1: 1–157.

BARTISH, I. V., U. SWENSON, J. MUNZINGER, AND A. A. ANDERBERG. 2005. Phylogenetic relationships among New Caledonian Sapotaceae (Ericales): molecular evidence for generic polyphyly and repeated dispersal. *American Journal of Botany* 92: 667–673.

BAEHNI, C. 1965. Mémoire sur les Sapotacées. Ill. Inventaire des genres. *Boissiera* 11: 1–262.

BERRY, E. W. 1939. A Miocene flora from the gorge of the Yumari river, Matanzas, Cuba. *Johns Hopkins University, Studies in Geology* 13: 95–135.

BLAKE, S. G. 1995. Late Quaternary history of Henderson Island, Pitcairn Group. *Biological Journal of the Linnean Society* 56: 43–62.

BONNEVILLE, A., R. LE SUAVE, L. AUDIN, V. CLOUARD, L. DOSSO, P. Y. GILLOT, P. JANNEY, K. JOHRDAHL, AND K. MAAMAATUAIAHUTAPU. 2002. Arago Seamount: the missing hotspot found in the Austral Islands. *Geology* 30: 1023–1026.

BREMER, K., E. M. FRIIS, AND B. BREMER. 2004. Molecular phylogenetic dating of asterid flowering plants shows early Cretaceous diversification. *Systematic Biology* 53: 496–505.

CARSON, H. L., AND D. A. CLAGUE. 1995. Geology and biogeography of the Hawaiian Islands. *In* W. L. Wagner and V. A. Funk [eds.], Hawaiian biogeography: evolution on a hot spot archipelago, 14–29. Smithsonian Institution Press, Washington, D.C., USA.

CHANDERBALI, A. S., H. VAN DER WERFF, AND S. S. RENNER. 2001. Phylogeny and historical biogeography of Lauraceae: evidence from the chloroplast and nuclear genomes. *Annals of the Missouri Botanical Garden* 88: 104–134.

CLEMENT, W. L., M. C. TEBBITT, L. L. FORREST, J. E. BLAIR, L. BROUILLET, T. ERIKSSON, AND S. M. SWENSEN. 2004. Phylogenetic position and biogeography of *Hillebrandia sandwicensis* (Begoniaceae): a rare Hawaiian relict. *American Journal of Botany* 91: 905–917.

COLLINSON, M. E., K. FOWLER, AND M. C. BOULTER. 1981. Floristic changes indicate a cooling climate in the Eocene of southern England. *Nature* 291: 315–317.

CRONQUIST, A. 1945. Studies in the Sapotaceae 3, *Dipholis* and *Bumelia*. *Journal of the Arnold Arboretum* 26: 435–447.

DAVIS, C. C., P. W. FRITSCH, C. D. BELL, AND S. MATHEWS. 2004. High-latitude Tertiary migrations of an exclusively tropical clade: evidence from Malpighiaceae. *International Journal of Plant Sciences* 165: S107–S121.

DEMESURE, B., B. COMPS, AND R. J. PETIT. 1995. A set of universal primers for amplification of polymorphic non-coding regions of mitochondrial and chloroplast DNA in plants. *Molecular Ecology* 4: 129–131.

DICK, C. W., K. ABUL-SALIM, AND E. BERMINGHAM. 2003. Molecular systematic analysis reveals cryptic tertiary diversification of a widespread tropical rain forest tree. *American Naturalist* 162: 691–703.

DOYLE, J. J., AND J. I. DAVIS. 1998. Homology in molecular phylogenetics: a parsimony perspective. *In* D. E. Soltis, P. S. Soltis, and J. J. Doyle [eds.], Molecular systematics of plants II DNA sequencing, 101–131. Kluwer, Boston, Massachusetts, USA.

DUBARD, M. 1912. Les Sapotacées de groupe des Sideroxylinées. Annales de l'Institut Botanico-Géologique Colonial de Marseille, serié 2, 10, 1.

ELDENÄS, P., M. KÄLLERSJÖ, AND A. A. ANDERBERG. 1999. Phylogenetic placement and circumscription of tribes Inuleae s. str. and Plucheae (Asteraceae): evidence from sequences of chloroplast gene *ndhF*. *Molecular Phylogenetics and Evolution* 13: 50–58.

FELSENSTEIN, J. 1984. Distance methods for inferring phylogenies: a justification. *Evolution* 38: 16–24.

FITCH, W. M., AND E. MARKOWITZ. 1970. An improved method for determining codon variability in a gene and its application to the rate of fixation of mutations in evolution. *Biochemical Genetics* 4: 579–593.

FITCH, W. M. 1971. Rate of change of concomitantly variable codons. *Journal of Molecular Evolution* 1: 84–96.

FRIIS, I. 1978. A reconsideration of the genera *Monotheca* and *Spiniluma* (Sapotaceae). *Kew Bulletin* 33: 91–98.

GOVAERTS, R., D. G. FRODIN, AND T. D. PENNINGTON. 2001. World checklist and bibliography of Sapotaceae. Royal Botanic Gardens, Kew, UK.

GRANT, V. 1971. Plant speciation. Columbia University Press, New York, New York, USA.

GREEN, P. 1999. Phrap, version 0.990329. Computer program and documentation available at www.phrap.org.

GREEN, P., AND B. EWING. 2002. Phred, version 0.020425c. Computer program and documentation available at www.phrap.org.

GRUAS-CAVAGNETTO, C. 1976. Étude palynologique du Paléogène du sud de l'Angleterre. *Cachiers de Micropaléontologie* 1: 1–49.

HAMILTON, M. B. 1999. Four primer pairs for the amplification of chloroplast intergenic regions with intraspecific variation. *Molecular Ecology* 8: 521–523.

HARLEY, M. 1991. Pollen morphology of the Sapotaceae. *In* T. D. Pennington [ed.], The genera of the Sapotaceae, 23–50. Royal Botanic Gardens, Kew, UK.

HILDENBRAND, A., P. Y. GILLOT, AND I. LE ROY. 2004. Volcano-tectonic and geochemical evolution of an oceanic intra-plate volcano: Tahiti-Nui (French Polynesia). *Earth and Planetary Science Letters* 217: 349–365.

HOLLICK, A. 1928. Paleobotany of Porto Rico. *Scientific Survey of Porto Rico and the Virgin Islands* 7: 177–393.

HUELSENBECK, J. P., AND J. P. BOLLBACK. 2001. Empirical and hierarchical Bayesian estimation of ancestral states. *Systematic Biology* 50: 351–366.

HUELSENBECK, J. P., AND F. RONQUIST. 2001. MrBayes: Bayesian inference of phylogenetic trees. *Bioinformatics* 17: 754–755.

HUGHES, C. E., C. D. BAILEY, AND S. A. HARRIS. 2002. Divergent and reticulate species relationships in *Leucaena* (Fabaceae) inferred from multiple data sources: insights into polyploid origins and nrDNA polymorphism. *American Journal of Botany* 89: 1057–1073.

ITURRALDE-VINCENT, M. A., AND R. D. E. MACPHEE. 1999. Paleogeography of the Caribbean region: implications for Cenozoic biogeography. *Bulletin of the American Museum of Natural History* 238: 1–95.

JOHNSON, A. W., AND J. G. PACKER. 1965. Polyploidy and environment in Arctic Alaska. *Science* 148: 237–239.

KASS, E. R., AND A. R. RAFTERY. 1995. Bayes factors. *Journal of the American Statistical Association* 90: 773–795.

KEDVES, M. 1967. Sur quelques problèmes de stratigraphie palynologique appliqué au Tertiaire inférieur en Europe. *Pollen et Spores* 9: 321–334.

KISHINO, H., J. L. THORNE, AND W. J. BRUNO. 2001. Performance of a divergence time estimation method under a probabilistic model of rate evolution. *Molecular Biology and Evolution* 18: 352–361.

KYNDT, T., B. VAN DROOGENBROECK, E. ROMEIJN-PEETERS, J. P. ROMERO-MOTOCHI, X. SCHELDEMAN, P. GOETGHEBEUR, P. VAN DAMME, AND G. GHEYSEN. 2005. Molecular phylogeny and evolution of Caricaceae based on rDNA internal transcribed spacers and chloroplast sequence data. *Molecular Phylogenetics and Evolution* 37: 442–459.

LAM, H. J. 1939. On the system of the Sapotaceae, with some remarks on taxonomical methods. *Recueil des Traveaux Botaniques Néerlandais* 36: 509–525.

LEE, C., AND J. WEN. 2004. Phylogeny of *Panax* using chloroplast *trnC-trnD* intergenic region and the utility of *trnC-trnD* in interspecific studies of plants. *Molecular Phylogenetics and Evolution* 31: 894–903.

LÖVE, A., AND D. LÖVE. 1957. Arctic polyploidy. *Proceedings of the Genetics Society of Canada* 2: 23–27.

MORLEY, R. J. 2000. Origin and evolution of tropical rainforests. Wiley, Chichester, UK.

MORLEY, R. J. 2003. Interplate dispersal paths for megathermal angiosperms. *Perspectives in Plant Ecology* 6: 5–20.

MUELLNER, A. N., V. SAVOLAINEN, R. SAMUEL, AND M. W. CHASE. 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. *Molecular Phylogenetics and Evolution* 40: 236–250.

NYLANDER, J. A. A. 2004. MrAIC.pl, version 1.4. Computer program distributed by the author. Evolutionary Biology Centre, Uppsala University, Uppsala, Sweden.

NYLANDER, J. A. A., F. RONQUIST, J. P. HUELSENBECK, AND J. L. NIEVES-ALDREY. 2004. Bayesian phylogenetic analysis of combined data. *Systematic Biology* 53: 47–67.

OLMSTEAD, R. G., AND J. A. SWEERE. 1994. Combining data in phylogenetic systematics: an empirical approach using three molecular data sets in the Solanaceae. *Systematic Biology* 43: 467–481.

PENNINGTON, T. D. 1991. The genera of the Sapotaceae. Royal Botanic Gardens, Kew, UK.

PINDELL, J. L., S. C. CANDE, W. C. PITMAN, D. B. ROWLEY, J. F. DEWEY, J. LABRECQUE, AND W. HAXBY. 1988. A plate-kinematic framework for models of Caribbean evolution. *Tectonophysics* 155: 121–138.

PIRIE, M. D., L. W. CHATROU, J. B. MOLS, R. H. J. ERKENS, AND J. OOSTERHOF. 2006. "Andean-centered" genera in the short-branch clade of Annonaceae: testing biogeographical hypotheses using phylogeny reconstruction and molecular dating. *Journal of Biogeography* 33: 31–46.

RAMBAUT, A. 1996. Se-Al: Sequence alignment editor, version 2.0a11. Computer program available at http://evolve.zoo.ox.ac.uk/.

RENNER, S. S. 2005. Relaxed molecular clocks for dating historical plant dispersal events. *Trends in Plant Science* 10: 550–558.

RENNER, S. S., G. CLAUSING, AND K. MEYER. 2001. Historical biogeography of Melastomataceae: the roles of Tertiary migration and long-distance dispersal. *American Journal of Botany* 88: 1290–1300.

RICHARDSON, J. E., L. W. CHATROU, J. B. MOLS, R. H. J. ERKENS, AND M. D. PIRIE. 2004. Historical biogeography of two cosmopolitan families of flowering plants: Annonaceae and Rhamnaceae. *Philosophical Transactions of the Royal Society of London, B, Biological Sciences* 359: 1495–1508.

RONQUIST, F. 2004. Bayesian inference of character evolution. *Trends in Ecology and Evolution* 19: 475–481.

RONQUIST, F., AND J. P. HUELSENBECK. 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 19: 1572–1574.

RUTSCHMANN, F. 2005. Bayesian molecular dating using PAML/multidivtime. A step-by-step manual. University of Zurich, Switzerland. Available at http://www.plant.ch.

RYDIN, C., K. R. PEDERSEN, AND E. M. FRIIS. 2004. On the evolutionary history of *Ephedra*: Cretaceous fossils and extant molecules. *Proceedings of the National Academy of Sciences, USA* 101: 16571–16576.

SANDERSON, M. J. 2002. r8s, version 1.50. Computer program and documentation available at http://phylo.ucdavis.edu/r8s/r8s.html.

SCHWARTZ, G. 1978. Estimating the dimensions of a model. *Annals of Statistics* 6: 461–464.

SMEDMARK, J. E. E., AND T. ERIKSSON. 2002. Phylogenetic relationships of *Geum* (Rosaceae) and relatives inferred from the nrITS and *trnL-trnF* regions. *Systematic Botany* 27: 303–317.

SMEDMARK, J. E. E., U. SWENSON, AND A. A. ANDERBERG. 2006. Accounting for variation of substitution rates through time in Bayesian phylogeny reconstruction of Sapotoideae (Sapotaceae). *Molecular Phylogenetics and Evolution* 39: 706–721.

SOLTIS, D. E., AND P. S. SOLTIS. 1993. Molecular-data and the dynamic nature of polyploidy. *Critical Reviews in Plant Sciences* 12: 243–273.

STEBBINS, G. L. 1947. Types of polyploids: their classification and significance. *Advances in Genetics* 1: 403–429.

STRAND, A. E., J. LEEBENS-MACK, AND B. G. MILLIGAN. 1997. Nuclear DNA-based markers for plant evolutionary biology. *Molecular Ecology* 6: 113–118.

SUH, Y., L. B. THIEN, H. E. REEVE, AND E. A. ZIMMER. 1993. Molecular evolution and phylogenetic implications of internal transcribed spacer sequences of ribosomal DNA in Winteraceae. *American Journal of Botany* 80: 1042–1055.

SWENSON, U., AND A. A. ANDERBERG. 2005. Phylogeny, character evolution, and classification of the Sapotaceae (Ericales). *Cladistics* 21: 101–130.

SWOFFORD, D. L. 2002. PAUP*: phylogenetic analysis using parsimony (* and other methods), version 4.0b10. Sinauer, Sunderland, Massachusetts, USA.

THORNE, J. L., AND H. KISHINO. 2002. Divergence time and evolutionary rate estimation with multilocus data. *Systematic Biology* 51: 689–702.

THORNE, J. L., H. KISHINO, AND I. S. PAINTER. 1998. Estimating the rate of evolution of the rate of molecular evolution. *Molecular Biology and Evolution* 15: 1647–1657.

TIFFNEY, B. H. 1985a. Perspectives on the origin of the floristic similarity between eastern Asia and eastern North America. *Journal of the Arnold Arboretum* 66: 73–94.

TIFFNEY, B. H. 1985b. The Eocene North Atlantic land bridge: its importance in Tertiary and modern phytogeography of the northern hemisphere. *Journal of the Arnold Arboretum* 66: 243–273.

TUFFLEY, C., AND M. STEEL. 1998. Modeling the covarion hypothesis of nucleotide substitution. *Mathematical Biosciences* 147: 63–91.

WASSERMAN, L. 2000. Bayesian model selection and averaging. *Journal of Mathematical Psychology* 44: 92–107.

WEEKS, A., D. C. CALY, AND B. B. SIMPSON. 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.

WILSON, J. T. 1963. A possible origin of the Hawaiian islands. *Canadian Journal of Physics* 41: 863–870.

WOLFE, J. A. 1975. Some aspects of plant geography of the northern hemisphere during the late Cretaceous and Tertiary. *Annals of the Missouri Botanical Garden* 62: 264–279.

WOLFE, J. A. 1978. A paleobotanical interpretation of Tertiary climate in the northern hemisphere. *American Scientist* 66: 694–703.

WOLFE, J. A. 1980. Tertiary climates and floristic relationships at high latitudes in the northern hemisphere. *Palaeogeography Palaeoclimatology Palaeoecology* 30: 313–323.

YANG, Z. 1997. PAML: a program package for phylogenetic analysis by maximum likelihood. *Cabios* 13: 555–556.

ZACHOS, J., M. PAGANI, L. SLOAN, E. THOMAS, AND K. BILLUPS. 2001. Trends, rhythms and aberrations in global climate 65 Ma to present. *Science* 292: 686–693.

ZALDIVAR-RIVERON, A., M. MORI, AND D. L. J. QUICKE. 2006. Systematics of the cyclostome subfamilies of braconid parasitic wasps (Hymenoptera : Ichneumonoidea): a simultaneous molecular and morphological Bayesian approach. *Molecular Phylogenetics and Evolution* 38: 130–145.

ZEREGA, N. J. C., W. L. CLEMENT, S. H. DATWYLER, AND G. D. WEIBLEN. 2005. Biogeography and divergence times in the mulberry family (Moraceae). *Molecular Phylogenetics and Evolution* 37: 402–416.

APPENDIX 1. List of species included in the study with information about EMBL accession numbers of DNA sequences, voucher specimens, geographic origin, and herbarium acronym. Sequences published elsewhere are marked with one (Smedmark et al., 2006), two (Anderberg and Swenson, 2003), three (I. V. Bartish et al., unpublished), or four (Bartish et al., 2005) asterisks. A dash indicates that the region was not sampled.

Taxon—EMBL accessions: *ndhF*, *trnC/petN*, *petN/psbM*, *psbM/trnD*, *trnH/psbA*, ITS, *AAT*; *Voucher specimen,* Locality.

Argania spinosa (L.) Skeels—AY230664**, AM179647*, AM179519*, AM179580*, AM179712*, AM408056, AM408009; *Nordenstam 9325*, Morocco, S.

Capurodendron androyense Aubrév.—AM179777*, AM179650*, AM179521*, AM179583*, AM179714*, AM408107, —; *Humbert 28855*, Madagascar, B.

Eberhardtia aurata (Pierre ex Dubard) Lecomte—AM179778*, AM179654*, AM179525*, AM179587*, AM179718*, AM408108, AM407961; *G. Hao 534*, Cult. South Cina Bot. Gard., S. ***Englerophytum natalense*** (Sond.) T. D. Penn.—AY230686**, —, —, —, DQ344111***, AY552150****, —; *C. K. 3483*, Tanzania, S.

Lecomtedoxa klaineana (Pierre ex Engl.) Pierre ex Dubard—AY230690**, AM179658*, AM179529*, AM179591*, AM179723*, AM408109, —; *Veldhuizen 1509*, Cult. Holland, WAG.

Mimusops obovata Sond.— —, —,—, —, —, —, AM407962; *Swenson & Karis 633*, South Africa, S.

Nesoluma nadeaudii (Drake) Pierre ex. H. J. Lam—AM407934, AM407794, AM407776, AM407756, AM407732, clones 1 AM408057/2 AM408058, clones 1 AM408001/2 AM408000/3 AM407996/4 AM407997/5 AM407998/6 AM407999; *Meyer 3050*, Tahiti, NOV. ***N. polynesicum*** (Hillebr.) Baill.—AY230704**, AM179680*, AM179549*, AM179614*, AM179746*, AM408059, clones 1 AM408005/2 AM408006/3 AM408002/4 AM408003/5 AM408004/6 AM408007; *Degener 20770*, Hawaii, S.

Omphalocarpum pachysteloides Mildbr. Ex Hutch. & Dalziel—AY230707**, —, —, —, DQ344122***, AY552150****, —; *Jongkind 2351*, Ghana, WAG.

Palaquium formosanum Hayata—AF421068**, AM179682*, AM179551*, AM179616*, AM179748*, AM408110, —; *Chung & Anderberg 1421*, Taiwan, HAST.

Sarcosperma laurinum (Benth.) Hooh f.—AF421080**, AM179646*, AM179518*, AM179579*, AM179711*, AM408055, —; *Saunders 2000*, Hong Kong, S. ***Sideroxylon americanum*** (Mill.) T. D. Penn.—AM179793*, AM179685*, AM179554*, AM179619*, AM179751*, AM408060, AM407963; Bahamas; Gillis 11576 (B). ***S. angustum*** T. D. Penn.—AM407935, AM407795, —, AM407757, AM407733, AM408061, —; *Ekman 4034*, Cuba, S. ***S. beguei*** Capuron ex Aubrév.—AM407936, —, —, AM407758, AM407734, AM408062, —; *McPherson et al. 14831*, Madagascar, K. ***S. betsimisarakum*** Lecomte—AY230729**, AM179686*, AM179555*, AM179620*, AM179752*, AM408063, clones 1 AM407991/2 AM407992/3 AM407993; *Schönenberger et al. A-102*, Madagascar, UPS. ***S.***

borbonicum A. DC.—AM407937, AM407796, —, —, —, AM408064, —; *Bosser 21325*, Reunion, P. ***S. capiri*** (A. DC.) Pittier—AM407938, AM407797, AM407777, AM407759, AM407735, AM408065, —; *García 1848*, Mexico, BM. ***S. capuronii*** Aubrév.—AM407939, AM407798, —, —, AM407736, AM408066, —; *Capuron 20151-SF*, Madagascar, P. ***S. celastrinum*** (Kunth) T. D. Penn—AM407940, AM407799, AM407778, AM407760, AM407737, AM408067, —; *Correll 50467*, Bahamas. BM. ***S. confertum*** C. Wright—AM407941, —, —, —, AM407738, AM408068, —; *Ekman 17405*, Cuba, S. ***S. contrerasii*** (Lundell) T. D. Penn—AM407942, AM407800, AM407779, AM407761, AM407739, AM408069, —; Lundell 20793, Guatemala, BM. ***S. cubense*** (Griseb.) T. D. Penn.—AM179794*, AM179687*, AM179556*, AM179621*, AM179753*, AM408070, AM408014; *Beurton & Mory 927*, Dominican Republic, B. ***S. floribundum*** Griseb.—AM407943, AM407801, AM407780, AM407762, AM407740, AM408071, —; *Lundell 20263*, Guatemala, BM. ***S. foetidissimum*** Jacq.—AY230730**, AM179688*, AM179557*, AM179622*, AM179754*, AM408072, clones 1 AM408021/2 AM408022/3 AM408023; *Lundin 638*, Cuba, S. ***S. galeatum*** (A. W. Hill) Baehni— —, —, —, —, —, AM408073, —; *Friedman 3288*, Rodrigues, P. ***S. gerrardianum*** (Hook f.) Aubrév.— —, —, —, —, —, AM408074, —; *Capuron 28826-SF*, Madagascar, P. ***S. grandiflorum*** A. DC.—AM407944, AM407802, AM407781, AM407763, AM407741, AM408075, —; *Friedman et al. 2653*, Mauritius, P. ***S. horridum*** (Griseb.) T. D. Penn.—AY230731**, AM179689*, AM179558*, AM179623*, AM179755*, AM408076, AM407964; *Gutiérrez & Nilsson 5*, Cuba, S. ***S. ibarrae*** (Lundell) T. D. Penn—AM407945, AM407803, AM407782, AM407764, AM407742, AM408077, AM408020; *Lundell 19752*, Guatemala, S. ***S. inerme*** L.—AY230732**, AM179690*, AM179559*, AM179624*, AM179756*, AM408078, AM407994; *Nielsen s.n.*, Cult. Denmark, S. ***S. lanuginosum*** Michx.—AY230733**, AM179691*, AM179560*, AM179625*, AM179757*, AM408079, —; *Correll & Ogden 28456*, Texas, S. ***S. leucophyllum*** S. Watson—AM407946, AM407804, AM407783, AM407765, AM407743, AM408080, AM408008; Carter 5706, Mexico, BM. ***S. lycioides*** L.—AM179795*, AM179692*, AM179561*, AM179626*, AM179758*, AM408081, AM408011; Radford et al. 11453, South Carolina, USA, B. ***S. majus*** (C. F. Gaertn.) Baehni—AM179796*, AM179693*, AM179562*, AM179627*, AM179759*, AM408082, —; *Capuron 28185_SF*, Reunion, B. ***S. marginatum*** (Decne. Ex Webb) Cout.—AM179797*, AM179694*, AM179563*, AM179628*, AM179760*, AM408083, AM408010; *Leyens CV-96–672*, Cape Verde, B. ***S. marmulano*** Banks ex Lowe—AY603783**, AM179695*, AM179564*, AM179629*, AM179761*, AM408084, —; *Swenson & Fernandez 581*, Canary Islands, S. ***S. mascatense*** (A. DC.) T. D. Penn.—AF421066**, AM179696*, AM179565*, AM179630*, AM179762*, AM408085, —; *Thulin, Beier & Hussein 9774*, Yemen, UPS. ***S. obovatum*** Lam.—AM407947, AM407805, AM407784, AM407766, AM407744, AM408086, AM407947; *García et al. 5586*, Dominican Republic, S. ***S. obtusifolium*** (Roem. & Schult.) T. D. Penn.—AM179798*, AM179697*, AM179566*, AM179631*, AM179763*, —, clones 1 AM407960/2 AM407966/3 AM407965/4 AM407967/5 AM408029; *Alvarez et al. 28772*, Mexico, B. ***S. occidentale*** (Hemsl.) T. D. Penn.—AM179799*, AM179698*, AM179567*, AM179632*, AM179764*, AM408088, clones 1 AM407973/2 AM407974/3 AM407975/4 AM407976/5 AM407977; *Carter & Sharsmith 4268*, Mexico, B. ***S. oxyacanthum*** Baill.—AM407948, AM407806, AM407785, AM407767, AM407745, AM408089, —; *Wood Y/75/388*, Yemen, BM. ***S. palmeri*** Rose (T. D. Penn.)—AM407949, AM407807, AM407786, AM407768, AM407746, AM408090, —; *Palmer 1513*, Mexico, S. ***S. persimile*** (Hemsl.) T. D. Penn.—AM407950, AM407808, AM407787, AM407769, AM407747, AM408091, clones 1 AM407978/2 AM407979/3 AM407980/4 AM407981/5 AM407982; *Véliz 99.7038*, Guatemala, BM. ***S. picardae*** (Urb.) T. D. Penn.—AM407951, AM407809, AM407788, AM407770, AM407748, AM408092, —; *Ekman 15576*, Hispaniola, S. ***S. portoricense*** Urb.—AM407952, AM407810, AM407789, AM407771, AM407749, AM408093, —; *Mathew 1*, Jamaica, BM. ***S. puberulum*** A. DC.—AM407953, AM407811, —, —, —, AM408094, AM407995; *Coode 4121*, Mauritius, P. ***S. reclinatum*** Michx.—AY230734**, AM179699*, AM179568*, AM179633*, AM179765*, AM408095, clones 1 AM407983/2 AM407984; *Traverse 592*, USA, GB. ***S. repens*** (Urb. & Ekman) T. D. Penn—AM179800*, AM179700*, AM179569*, AM179634*, AM179766*, AM408096, —; *Greuter & Rankin 24954*, Dominican Republic, B. ***S. rotundifolium*** (Sw.) T. D. Penn.—AM407954, AM407812, AM407790, AM407772, AM407750, AM408097, clones 1 AM407968/2 AM407969/3 AM407970/4 AM407971/5 AM407972; *Webster et al. 8458*, Jamaica, S. ***S. salicifolium*** (L.) Lam—AY230735**, AM179701*, AM179570*, AM179635*, AM179767*, AM408098, clones 1 AM408015/2 AM408016/3 AM408017/4 AM408019/5 AM408018; *Gutiérrez & Nilsson 14*, Cuba, S. ***S. saxorum*** Lecomte—AY230736**, AM179702*, AM179571*, AM179636*, AM179768*, AM408099, clones 1 AM407988/2 AM407989/3 AM407990/4 AM407985/5 AM407986/6 AM407987; *Jongkind 3500*, Madagascar, WAG. ***S. sessiliflorum*** (Poir.) Capuron ex Aubrév.—AM407955, AM407813, AM407791, AM407773, AM407751, AM408100, —; *Lorence & Edgerley 2706*, Mauritius, P. ***S. stenospermum*** (Standl.) T. D. Penn.—AM407956, —, —, —, AM407752, AM408101, —; *Stevens 22935*, Nicaragua, BM. ***S. stevensonii*** (Standl.) Standl. & Steyerm.—AM407957, AM407814, AM407792, AM407774, AM407753, AM408102, —; *Lundell & Contreras 19057*, Guatemala, S. ***S. tamblokoko*** Aubrév.—AM407958, AM407815, —, —, AM407754, AM408103, —; *Capuron 22388-SF*, Madagascar, P. ***S. tenax*** L.—AM179801*, AM179703*, AM179572*, AM179637*, AM179769*, AM408104, AM408012; *Radford & Leonard 11519*, South Carolina, B. ***S. tepicense*** (Standl.) T. D. Penn.—AM407959, AM407816, AM407793, AM407775, AM407755, AM408105, clones 1 AM408024/2 AM408025/3 AM408026/4 AM408027/5 AM408028; *Ctentry 2931*, Mexico, S. ***S. wightianum*** Hook. & Arn.—AM179802*, AM179704*, AM179573*, AM179638*, AM179770*, AM408106, AM408013; *G. Hao 532*, Cult. South Cina Bot. Gard., S.

Xantolis siamensis (Fletcher) P. Royen—AY230744**, —, —, AM179645*, DQ344151***, AY552154****, —; *Smitairi 1*, Thailand, L.