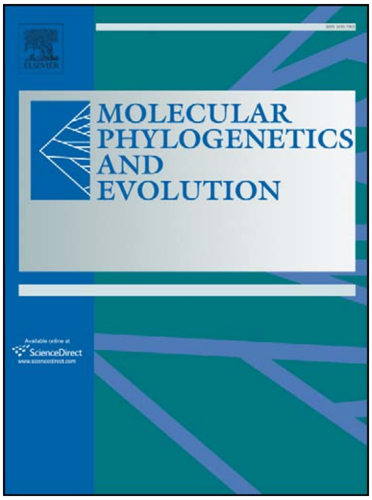


Plastid and nuclear DNA markers reveal intricate relationships at subfamilial and tribal levels in the soapberry family (Sapindaceae)

Sven Buerki, Félix Forest, Pedro Acevedo-Rodríguez, Martin W. Callmander, Johan A.A. Nylander, Mark Harrington, Isabel Sanmartín, Philippe Küpfer, Nadir Alvarez

PII: S1055-7903(09)00017-7
DOI: 10.1016/j.ympev.2009.01.012
Reference: YMPEV 3130

To appear in: *Molecular Phylogenetics and Evolution*

Received Date: 21 May 2008
Revised Date: 27 November 2008
Accepted Date: 23 January 2009

Please cite this article as: Buerki, S., Forest, F., Acevedo-Rodríguez, P., Callmander, M.W., Nylander, J.A.A., Harrington, M., Sanmartín, I., Küpfer, P., Alvarez, N., Plastid and nuclear DNA markers reveal intricate relationships at subfamilial and tribal levels in the soapberry family (Sapindaceae), *Molecular Phylogenetics and Evolution* (2009), doi: 10.1016/j.ympev.2009.01.012

Plastid and nuclear DNA markers reveal intricate relationships at subfamilial and tribal levels in the soapberry family (Sapindaceae)

Sven Buerki [a,*], Félix Forest [b], Pedro Acevedo-Rodríguez [c], Martin W. Callmander [d,e], Johan A. A. Nylander [f], Mark Harrington [g], Isabel Sanmartín [h], Philippe Küpfer [a] and Nadir Alvarez [a]

[a] Institute of Biology, University of Neuchâtel, Rue Emile-Argand 11, CH-2009 Neuchâtel, Switzerland.

[b] Molecular Systematics Section, Jodrell Laboratory, Royal Botanic Gardens, Kew, Richmond, Surrey TW9 3DS, United Kingdom.

[c] Department of Botany, Smithsonian Institution, National Museum of Natural History, NHB-166, Washington, D.C. 20560, U.S.A.

[d] Missouri Botanical Garden, P.O. Box 299, 63166-0299, St. Louis, MO, U.S.A.

[e] Conservatoire et Jardin botaniques de la ville de Genève, ch. de l'Impératrice 1, CH-1292 Chambésy, Switzerland.

[f] Department of Botany, Stockholm University, SE-10691, Stockholm, Sweden.

[g] School of Marine and Tropical Biology, James Cook University, P.O. Box 6811, Cairns, Queensland 4870, Australia.

[h] Department of Biodiversity and Conservation, Royal Botanic Garden – CSIC, Plaza de Murillo 2, 28014 Madrid, Spain.

[*] Corresponding author. Phone: +41327182360; Fax: +41327183001; E-mail address: sven.buerki@unine.ch.

Abstract

The economically important soapberry family (Sapindaceae) comprises about 1900 species mainly found in the tropical regions of the world, with only a few genera being restricted to temperate areas. The infrafamilial classification of the Sapindaceae and its relationships to the closely related Aceraceae and Hippocastanaceae – which have now been included in an expanded definition of Sapindaceae (i.e., subfamily Hippocastanoideae) – have been debated for decades. Here we present a phylogenetic analysis of Sapindaceae based on eight DNA sequence regions from the plastid and nuclear genomes and including 85 of the 141 genera defined within the family. Our study comprises 997 new sequences of Sapindaceae from 152 specimens. Despite presenting 18.6% of missing data our complete data set produced a topology fully congruent with the one obtained from a subset without missing data, but including fewer markers. The use of additional information therefore led to a consistent result in the relative position of clades and allowed the definition of a new phylogenetic hypothesis. Our results confirm a high level of paraphyly and polyphyly at the subfamilial and tribal levels and even contest the monophyletic status of several genera. Our study confirms that the Chinese monotypic genus *Xanthoceras* is sister to the rest of the family, in which subfamily Hippocastanoideae is sister to a clade comprising subfamilies Dodonaeoideae and Sapindoideae. On the basis of the strong support demonstrated in Sapindoideae, Dodonaeoideae and Hippocastanoideae as well as in 14 subclades, we propose and discuss informal groupings as basis for a new classification of Sapindaceae.

Key words: Aceraceae; classification; Hippocastanaceae; molecular phylogeny; paraphyly; polyphyly; Sapindaceae; *Xanthoceras*.

1. Introduction

The soapberry family (Sapindaceae: Sapindales) comprising c. 1900 species (Acevedo-Rodríguez, pers. comm.), has a predominantly pantropical distribution with the occurrence of some taxa in temperate areas (e.g., *Acer*, *Aesculus*, *Atalaya*, *Diplopeltis, Dodonaea*). Sapindaceae include many economically important species used for their fruits [e.g., guarana (*Paullinia cupana*), litchi (*Litchi chinensis*), longan (*Dimocarpus longan*), pitomba (*Talisia esculenta*) and rambutan (*Nephelium lappaceum*)], wood [e.g., buckeyes (*Aesculus*)] or as ornamentals (*Koelreuteria*, *Ungnadia*).

The circumscription of the family as well as the relationships among subfamilial entities have been widely challenged since the very first worldwide treatment of Sapindaceae sensu stricto (s.s.) (including subfamilies Sapindoideae and Dodonaeoideae) proposed by Radlkofer (1890, 1933; for a review see Harrington et al., 2005). For instance, several genera within the Sapindoideae (e.g., *Tinopsis* and *Plagioscyphus* from Madagascar; Capuron, 1969) were shown to be morphologically transitional between tribes described by Radlkofer (1933), which prevented the recognition of unequivocal tribes. Within Sapindaceae s.s. the higher taxonomic entities (subfamilies and tribes) were originally defined by Radlkofer (1933) based on the number and type of ovules per locule, the fruit morphology, the presence or not of an arillode, the leaf type and the cotyledon shape. On the basis of macromorphological and palynological characters, Müller and Leenhouts (1976) revised the classification of Radlkofer (1933). They recognized eight major pollen types (A to H) and several subtypes (e.g., type-A1), mainly based on their shape and characteristics of the aperture (Fig. 1). The pollen grains in Sapindaceae are triporate [the diporate type-D pollen of *Lophostigma* recognized by Müller and Leenhouts (1976) was wrongly identified; see Acevedo-Rodríguez (1993a)]. Spherical pollen shape occurs in the majority of species (e.g., types A, B and H), whereas a triangular (type-C) or oblate (type-A1) shape is more restricted. The colpi may be absent (e.g., type-G) or parasyncolporate (e.g., type-A) to syncolporate (e.g., type-B) (Fig. 1). Based on those characters Müller and Leenhouts (1976) rearranged the nine tribes of Sapindoideae recognized by Radlkofer (1933) into three taxonomically unranked groups characterized by their distribution, the presence or absence of an arillode surrounding the seed and the pollen types [i.e.,. group A comprised Sapindeae, Lepisantheae (incl. Aphanieae) and Melicocceae; group B comprised Schleichereae, Nephelieae and Cupanieae; group C comprised Paullinieae and Thouinieae]. They did not, however, modify the classification within the Dodonaeoideae and maintained the five tribes described by Radlkofer (i.e., Cossinieae, Dodonaeeae, Doratoxyleae, Harpullieae and Koelreuterieae; 1933). Furthermore, Müller and Leenhouts (1976) kept the predominantly temperate families Aceraceae and Hippocastanaceae separate from the rest of Sapindaceae. The circumscription of Sapindaceae has been debated ever since. Takhtajan (1987), Cronquist (1988) and Dahlgren (1989) maintained Aceraceae and Hippocastanaceae separate from Sapindaceae, whereas broader concepts of the family have been adopted by several workers (e.g., Umadevi and Daniel, 1991; Judd et al., 1994; Gadek et al., 1996; Savolainen et al., 2000; Thorne, 2000, 2007; APGII, 2003).

Building on a large-scale molecular phylogenetic analysis of Sapindales (Gadek et al., 1996), Harrington et al. (2005) published the first molecular phylogeny of Sapindaceae sensu lato (s.l.) (including Aceraceae and Hippocastanaceae) inferred from the plastid genes *rbcL* and *matK*. Their phylogeny recognized the subdivision of Sapindaceae s.l. into four supported lineages, a monotypic Xanthoceroideae, Hippocastanoideae (including Aceraceae, Hippocastanaceae and *Handeliodendron*), a more narrowly defined Dodonaeoideae and Sapindoideae (including *Koelreuteria* and *Ungnadia*). Relationships between these four lineages remained weakly supported. Confirming previous works based on morphological features, Harrington et al. (2005) highlighted the paraphyletic or polyphyletic nature of several tribes described by Radlkofer (1933).

According to the new assessment of the Sapindaceae s.l. proposed by Thorne (2007; mainly based on Harrington et al., 2005) and a broad review of currently described taxa, it is now widely accepted that the c. 1900 species of this cosmopolitan family are divided into 141 genera (see Table 1; Acevedo-Rodríguez, pers. comm.). Even if Harrington et al. (2005) covered worldwide representatives of Sapindaceae s.l., the sampling (64 of the 141 genera, i.e., 45.4%) and the number of markers were not sufficient to assess the relationships among and within the major lineages of the family with confidence. In this study we provide a new assessment of the phylogenetic relationships within Sapindaceae s.l. based on 60.3% of the generic diversity (85 of the 141 genera) and including the previously unsampled tribe Cossinieae. The analysis is based on a combination of one nuclear (ITS region; *ITS1*, *5.8S*, *ITS2*) and seven plastid (coding *matK* and *rpoB*; non coding *trnL* intron and intergenic spacers *trnD-trnT*, *trnK-matK*, *trnL-trnF* and *trnS-trnG*) markers. Coding plastid regions have proven to be useful in addressing phylogenetic relationships at higher taxonomic levels (e.g., Clayton et al., 2007; Muellner et al., 2006, 2007; Harrington et al., 2005), whereas noncoding regions (introns and intergenic spacers) were shown to be more useful at lower taxonomic ranks (Baldwin, 1992; Soltis and Soltis, 1998). The combination of several markers from both nuclear and plastid genomes as well as coding and non coding regions are expected to improve the resolution of phylogenetic relationships within the family. In this study, our objectives are (1) to examine the relationships between the traditionally defined Aceraceae and Hippocastanaceae with the rest of Sapindaceae, (2) to evaluate the tribal concepts of Radlkofer (1933) and Müller and Leenhouts (1976), (3) to examine phylogenetic relationships among taxa in light of characters traditionally used to define the higher level groupings in Sapindaceae s.l. (e.g., number of ovules per locule, pollen morphology, leaf type and presence/absence of an arillode) and (4) to propose a new preliminary infrafamilial classification for Sapindaceae s.l.

In addition of being a challenging family at the taxonomic level, the amplification of molecular markers in Sapindaceae s.l. is made difficult by several mutations occurring in flanking regions of widely used plastid and nuclear regions such as *matK* (Harrington et al., 2005) and ITS (Edwards and Gadek, 2001). Those mutations complicate the compilation of multilocus data sets without missing data. Maximizing taxa and markers representation to provide a reliable phylogenetic hypothesis inferred from nuclear and plastid genomes is required to propose a new classification for family Sapindaceae. This was achieved by analysing two data sets based on the same taxa, but including different levels of missing data (i.e., different number of markers). While the inclusion of missing data was widely recognized as a major drawback in phylogenetic analyses during the early 90's (e.g., Huelsenbeck, 1991; Wiens and Reeder, 1995), recent simulations (Wiens, 1998, 2003, 2006) and empirical analyses (Bapste et al., 2002 ; Driskell et al., 2004; Philippe et al., 2004) have shown that taxa comprising high levels of missing data could be accurately placed in phylogenies. Moreover, adding incomplete taxa to a phylogenetic analysis was even shown to improve the accuracy of a given topology, e.g. by subdividing misleading long branches (Wiens, 2005). However, there is a strong heterogeneity in the ability of the different phylogenetic algorithms for managing data sets with substantial levels of missing data (Wiens, 2006), with maximum parsimony performing poorly compared to model-based algorithms such as maximum likelihood and Bayesian inference (Wiens, 2005, 2006).

2. Material and Methods

2.1. Taxon sampling

Species names, voucher information, and GenBank accession numbers for all sequences are provided in the Appendix. The sampling strategy was designed to encompass the majority of subfamilies, tribes and genera of the family as recognized by the existing classifications of Radlkofer (1933), Müller and Leenhouts (1976) and Thorne (2007). Ingroup sampling comprised 152 specimens representing 60.3% of the generic diversity (85 of the 141 genera; 28 of the 57 missing genera in this analysis are monospecific; Table 1). The outgroup included Anacardiaceae (*Sorindeia* sp.; defined as outgroup in all analyses; Savolainen et al., 2000 and Muellner et al., 2007) and Simaroubaceae (*Harrisonia abyssinica*). Silica-gel dried samples (Chase and Hills, 1991) were collected in the field by the authors and complemented with materials from the DNA banks of the Missouri Botanical Garden (St. Louis, U.S.A.), the Royal Botanic Gardens, Kew (London, U.K.) and the James Cook University (Cairns, Australia).

2.2. DNA sequencing

Samples from the collections of the Missouri Botanical Garden and field collected samples were extracted in the laboratory of Evolutionary Botany at the University of Neuchâtel (Switzerland) using the QIAGEN DNeasy plant kit (Qiagen, Hilden, Germany) and following the manufacturer's protocol. Samples from the collections of the Royal Botanic Gardens, Kew, were extracted using the 2X cetyltrimethylammonium bromide (CTAB) procedure of Doyle and Doyle (1987) with minor modifications (see Muellner et al., 2005) followed by additional purification using a caesium chloride/ethidium bromide gradient (1.55 g/ml) and a dialysis procedure. The samples from James Cook University (Cairns, Australia) were extracted with the CTAB procedure of Doyle and Doyle (1987).

Seven plastid DNA regions and one nuclear ribosomal DNA region were amplified. Primers for the plastid regions are those described in Edwards and Gadek (2001) for *matK* (specific primer for the Dodonaeoideae were designed by Harrington et al., 2005) and the *trnK-matK* intergenic spacer (IGS), the DNA barcoding project (http://www.kew.org/barcoding/update.html) for *rpoB*, Demesure et al. (1995) for the *trnD-trnT* IGS, Taberlet et al. (1991) for *trnL* intron and *trnL-trnF* IGS, and Hamilton (1995) for *trnS-trnG* IGS. Primers for the ITS region are described in White et al. (1990) and additional primers were designed by Edwards and Gadek (2001) for Sapindaceae s.l.

Amplification of selected regions were achieved in a 25 μl reaction mixture containing 5 μl 5X PCR buffer, 1.5 μl 25mM $MgCl_2$, 0.5 μl 10mM dNTPs, 0.5 μl 10mM primers, 0.2 μl GoTaq polymerase (5u/μl) (Promega, Madison, WI, USA), and 14.5 μl ddH_2O. The amplification of the *matK* region was improved by the addition of 4% DMSO in the total volume of the PCR mix. PCR was performed in a Biometra ® T3 thermocycler. Initial denaturation was programmed for 2 min at 95^{o}C, followed by 35 cycles at 95^{o}C for 45 s, 50^{o}C for 45 s, 72^{o}C for 1 min, plus a final extension of 10 min at 72^{o}C. PCR products were purified using the QIAquick PCR purification kit (Qiagen, Hilden, Germany) and fluorescent sequencing was performed by Macrogen, Inc. (Seoul, South Korea) with the same primers used for PCR amplification.

2.3. Alignment

The program Sequencher version 4.1 (Gene Codes Corp., Ann Arbor, Michigan, USA) was used to assemble complementary strands and verify software base-calling. The eight regions where initially aligned individually with ClustalX (Thompson et al., 1997), and thereafter manually adjusted with the program Bioedit (Hall, 1999) using the similarity criterion (Morrison, 2006). The program Concatenate (Alexis Criscuolo, http://www.lirmm.fr/~criscuol/) was used to construct two combined matrices, differing in the number of markers considered and in the level of missing data (see below).

2.4. Phylogenetic analyses

2.4.1. Single-gene analyses

Individual phylogenetic analyses and their corresponding bootstrap analyses were performed using the maximum likelihood (ML) and maximum parsimony (MP) criteria. Each partition and the combined data sets were analyzed using parsimony ratchet (Nixon, 1999) as implemented in PAUPrat (Sikes and Lewis, 2001). Based on recommendations by Nixon (1999), ten independent searches were performed with 200 iterations and 15% of the parsimony informative characters perturbed. The shortest equally most parsimonious trees were combined to produce a strict consensus tree. To assess the support at each node, non parametric bootstrap analyses (Felsenstein, 1985) were performed using PAUP* version 4.0b10 (Swofford, 2002) with 1000 replicates, SPR branch swapping, simple sequence addition, MULTREES and holding 10 trees per replicate. We used SPR branch swapping because it has been shown to be twice as fast as TBR and results in support percentages that are not significantly different (Salamin et al., 2003).

Model selection for each partition was assessed using Modeltest version 3.7 (Posada and Crandall, 1998) and the Akaike information criterion (Akaike, 1973). ML analyses were performed using RAxML version 7.0.0 (Stamatakis, 2006; Stamatakis et al., 2008) with a 1000 rapid bootstrap analyses followed by the search of the best-scoring ML tree in one single run. This analysis was done using the facilities offered by the CIPRES portal in San-Diego, U.S.A. (http://8ball.sdsc.edu:8888/cipres-web/home).

In this study, nodes with bootstrap supports (BS) below 50% are considered not supported, 50-74% are considered weakly supported, 75-89% are moderately supported and 90-100% are strongly supported. Topological differences between single-gene phylogenetic trees were compared using TreeJuxtaposer (Munzner et al., 2003), taking into account the level of resolution of each marker and their bootstrap supports. In this study, topological differences having a bootstrap support inferior to 75% were not considered.

2.4.2. Combined analyses

The impact of missing data on combined MP and ML phylogenetic analyses was tested based on two different combined matrices. The first matrix (hereafter named “4 markers” data set) was composed of specimens for which sequence information was available for the nuclear ribosomal ITS region and for three of the seven plastid regions (*rpoB*, *trnL* intron and *trnL-trnF* IGS). In this combined matrix, the four remaining plastid markers were not included in order to have a complete matrix without missing data. The second combined matrix (hereafter named “4+4 markers” data set) comprised the same set of taxa as the “4 markers” data set, but also included the other four plastid markers (*matK*, *trnD-trnT* IGS, *trnK-matK* IGS and *trnS-trnG* IGS). This data set was designed to evaluate the effect of additional information on the resolution and support of topologies in comparison to the “4 markers” analyses. Taxa for which no sequences were available for a given marker were coded as missing data for the corresponding cells in the combined matrix (sensu Wiens and Reeder, 1995).

Total evidence trees (sensu Kluge, 1989) were determined using both ML and MP criteria on the two data sets using the same settings as in the single-gene analyses. Non parametric bootstrap analyses were performed for the data sets following the same settings as for the single-gene analyses. Before computing total evidence trees, an incongruence length difference (ILD) test (Farris et al., 1994) was performed as implemented in PAUP* version 4.0b10 (Swofford, 2002) with 100 replicates.

2.5. Topological congruence and impact of missing data on combined analyses

Based on analyses of the combined matrices (i.e., “4 markers” and “4+4 markers” data sets), the impact of of missing data on MP and ML phylogenetic analyses was investigated (i) by assessing topological distances among trees obtained using different data sets and algorithms and (ii) by comparing taxa groupings (and clade supports) in each topology. The explicitly agree distance (Estabrook et al., 1985; Estabrook, 1992; EA distance) was calculated to evaluate the extent to which total evidence trees were compatible with each other. The EA distance quantifies the differences between trees of the same size (i.e., comprising the same number of terminal taxa). It evaluates the proportion of triplets that are resolved identically in two trees (see Wilkinson et al., 2005). EA distances were calculated using DARWIN 5 (Perrier et al., 2003). The congruence of topological groupings in analyses obtained from different data sets and algorithms was evaluated using TreeJuxtaposer (Munzner et al., 2003) and bootstrap supports of each main clade were compared.

3. Results

3.1. Alignment

The number of sequences included in each single-gene partition varied from 69 in *trnS-trnG* IGS to 154 in *rpoB*, *trnL* intron and *trnL-trnF* IGS (Table 2). For the ITS region, all specimens were sequenced, except the outgroup species *Sorindeia* sp. (i.e., 153 sequences were produced). All sequences used to reconstruct the phylogenetic tree of Sapindaceae s.l. were produced for this study. The alignment length ranged from 363 bp in *rpoB* to 2156 bp in *trnS-trnG* IGS (Table 2). The ITS region had the highest number of variable characters (51.4%), whereas *trnS-trnG* IGS had the lowest (23.8%), even less than the coding regions *matK* and *rpoB* (29.1% and 37.2%, respectively). The same trend was recorded for the percentage of potentially parsimony-informative characters (37.8% for the ITS region and 9.0% for the *trnS-trnG* IGS; Table 2).

The combined data sets consisted respectively of 615 sequences (154 specimens; no missing data in ingroup taxa) for the “4 markers” data set, and 997 sequences (154 specimens; 18.6 % missing data) for the “4+4 markers” data set (Table 2). The alignment length of the two data sets was respectively 3031 bp (“4 markers”) and 9657 bp (“4+4 markers”). The “4 markers” data set had a highest percentage of variable characters (44.7%) than the “4+4 markers” data set (37.0%). The same observations were recorded for the percentage of potentially parsimony-informative characters (30.3% for the “4 markers” and 21.2% for the “4+4 markers” data sets; Table 2). However, when considering the total amount of phylogenetic information averaged by the number of taxa, the “4 markers” data set shown a value more than twice higher than did the “4+4 markers” data set (Table 2).

3.2. Phylogenetic analyses

3.2.1. Single-gene analyses

The best-fit model for all partitions was the general time reversible (GTR) with an alpha parameter for the shape of the gamma distribution to account for among-site rate heterogeneity (Yang, 1993). The only exception was for the ITS region for which a proportion of invariable sites was added. Although the MP and ML single-gene analyses provided topologies with different levels of resolution within Sapindaceae s.l. (e.g., the MP trees were usually not resolved in several parts of the tree), no moderately to strongly supported differences (> 75%) were observed between single-gene trees. In addition, the ILD test was not significant (P=0.9) and indicated that the eight data sets were congruent. Those results allowed the combination of the partitions in a total evidence approach. Statistics (number of most parsimonious trees; tree length; consistency and retention indices) for each analysis are reported in Table 2.

3.2.2. Combined analyses

The most parsimonious trees for the two combined analyses under the MP criterion were respectively 5889 ("4 markers" data set) and 9843 ("4+4 markers" data set) steps. Under the ML criterion, the best-fit model for the combined matrices was GTR with a proportion of invariable sites and an alpha parameter for the shape of the gamma distribution to account for among-site rate heterogeneity (Yang, 1993). This model was used to perform the ML search (log likelihoods were -34322.2 for the "4 markers" data set and -69253.8 for the "4+4 markers" data set) followed by rapid bootstrap analyses.

3.3. Topological congruence and impact of missing data on combined analyses

The congruence (expressed by 1 - EA distance) between total evidence trees compiled under the ML criterion was higher (98% of common triplets between total evidence trees based on "4 markers" and "4+4 markers" data sets) than between total evidence trees obtained under MP criterion (90% of common triplets between total evidence trees based on "4 markers" and "4+4 markers" data sets) (Table 3). The MP "4 markers" total evidence tree exhibits the highest EA distances with the other total evidence trees (Table 3).

Each of the four total evidence analyses showed support for the monophyly of Sapindaceae s.l. as defined by Thorne (2007) including Aceraceae and Hippocastanaceae (Table 4). No matter which data set or algorithm were considered, the family was subdivided into three moderately to strongly-supported lineages and a fourth lineage only consisting of *Xanthoceras sorbifolia*, with the following relationships: (*Xanthoceras sorbifolia*, (clade A, (clade B, clade C))) (Table 4; Fig. 2). Despite strong support for each clade, the sister position of the monotypic *Xanthoceras* was not supported in any analyses (see clade A + clade B + clade C in Table 4). This lineage corresponded to subfamily Xanthoceroideae as described by Thorne (2007). Clade A corresponded to subfamily Hippocastanoideae (including the previous recognized families Aceraceae and Hippocastanaceae) as described by Harrington et al. (2005) and Thorne (2007). Clade B corresponded to subfamily Dodonaeoideae as described by Harrington et al. (2005) and Thorne (2007) with the addition of *Euphorianthus* (Cupanieae; Sapindoideae). Clade C corresponded to subfamily

Sapindoideae (Thorne, 2007; Harrington et al., 2005) plus one representative from Dodonaeoideae, *Conchopetalum*, included in tribe Harpullieae. Clade C was moderately to strongly supported as monophyletic and divided into ten groups, but not in the MP "4 markers" total evidence tree (only one exception: clade V nested in clade VI; Table 4). The bootstrap supports of each clade obtained under the ML algorithm are consistent in both data sets (Table 4), whereas support slightly increases in MP analyses, in parallel to an increase in missing data (Table 4).

The "4 markers" and "4+4 markers" topologies recognized all the classical tribes (except the Paullinieae) as paraphyletic or polyphyletic. However, phylogenetic status of tribes Cossinieae and Koelreuterieae were not tested because only one genus per tribe was considered. In total 5 of the 67 non-monotypic sampled genera (7.5%) are paraphyletic or polyphyletic (*Cupaniopsis*, *Guioa*, *Haplocoelum*, *Matayba*, *Sarcotoechia*). However, the phylogenetic status of some of these genera needs to be treated with caution because of weak bootstrap supports and limited sampling (e.g., *Guioa*).

4. Discussion

4.1. Congruence of topologies with and without missing data

Our results indicate a high level of congruence among topologies obtained using data sets with and without missing data and based on different algorithms. Considering the "4 markers" data set (without missing data), MP and ML algorithms however produced slightly different topologies regarding clades C-V and C-VI (i.e., in the MP "4 markers" tree clade C-VI is paraphyletic with the inclusion of the clade C-V, whereas all other topologies considered this clade as monophyletic; Table 4). This could be explained mostly by the small amount of phylogenetic information in the "4 markers" data set that prevent the MP algorithm to find a proper solution (averaged over the number of terminal taxa; Table 2). Although the addition of 4 markers to the data set generated 18.6% of missing data (27.3% of missing nucleotides) in the "4+4 markers" data set, the added information doubled the mean amount of potentially parsimonious-informative characters per terminal taxa and increased the bootstrap support for several nodes in the total evidence trees (Tables 2, 4). Since our results highlight a high congruence level among topologies obtained with different data sets and algorithms, only the ML total evidence tree inferred from the "4+4 markers" data set will be discussed in order to maximize phylogenetic information (Figs. 2-6).

4.2. Phylogenetic relationships

Our results support (1) the paraphyly of the currently defined Dodonaeoideae and Sapindoideae as defined by Thorne (2007); (2) the polyphyly of all tribes (tribes Cossinieae and Koelreuterieae are not considered because only one genus per tribe was sampled) with the possible exception of Paullinieae – whose monophyletic status shall be evaluated by the inclusion of three missing genera *Houssayanthus*, *Lophostigma* and *Thinouia* in future analyses – and (3) the paraphyly or polyphyly of 5 of the 67 non-monotypic sampled genera (7.5%) included in this study (Table 1).

In light of these results, a new infrafamilial classification for Sapindaceae s.l. is required. However, we recommend caution in formally proposing new tribes until (i) non-molecular synapomorphies supporting putative new tribal delimitations are identified and (ii) the inclusion of missing genera in future phylogenetic analyses. In order to provide efficient guidelines for a new classification of the family, the phylogenetic framework obtained here is discussed according to several key morphological characters such as leaf type (including phyllotaxy), wood anatomy, number of ovules per locule, fruit type and pollen (Fig. 1), as well as geographical distribution. Hereafter, the definition of Dodonaeoideae and Sapindoideae will be expanded to include *Euphorianthus* in the former and *Conchopetalum* in the latter.

Xanthoceroideae and Hippocastanoideae occur mostly in temperate regions [except *Billia* (not included here), which occurs from Mexico to tropical South America], whereas Dodonaeoideae have a temperate (e.g., south of Australia) and tropical pattern of distribution. On the other hand, Sapindoideae have mainly radiated in tropical regions. Within Sapindaceae s.l., a trend towards the reduction of the number of ovule per locule is observed: from six to eight (Xanthoceroideae) to two (Hippocastanoideae and most of the

Dodonaeoideae) and finally one (Sapindoideae except *Conchopetalum*). All four subfamilies recognized by Thorne (2007) are discussed separately below.

4.3. Subfamily Xanthoceroideae (Fig. 2)

The phylogenetic position of the monotypic Chinese *Xanthoceras* in relation to the other three main lineages of Sapindaceae remains unsupported (BS<50) (Fig. 2; Table 4). Nevertheless, this species was moderately supported as the earliest-diverging lineage in Sapindaceae s.l. in earlier studies (*matK*, *rbcL*, Harrington et al. 2005; *rbcL*, Savolainen et al. 2000; 18S rDNA, *atpB*, *rbcL*, Soltis et al. 2000). In the first molecular phylogeny of Sapindaceae s.l., Harrington et al. (2005) argued that an increased sampling of other monotypic Southeast Asian genera of Harpullieae (e.g., *Arfeuillea*, *Delavaya, Eurycorymbus*) and Koelreuterieae (*Sinoradlkofera*) might help break up possible long-branch attraction and stabilize the position of this taxon. However, our study shows that even when considering 60.3% of the generic diversity and including *Arfeuillea*, *Delavaya* and *Eurycorymbus*, the phylogenetic position of this genus remains unchanged. This small shrub is characterized by unusual features in Sapindaceae such as deciduous imparipinnate leaves (vs. deciduous simple leaves or sempervirent imparipinnate or paripinnate leaves in other Sapindaceae), six to eight fertile ovules per locule (generally 1 or 2 ovules per locule in the rest of the family) and the presence of orange horn-like appendages protruding from the disk (absent in other genera). Moreover, this species exhibits a type-A pollen which was expected to be ancestral in Sapindaceae by Müller and Leenhouts (1976) (Fig. 1). However, this pollen type is widespread across the taxa sampled in our phylogeny and is consequently of limited systematic utility.

4.4. Subfamily Hippocastanoideae (Clade A; Figs. 2)

The inclusion of Aceraceae and Hippocastanaceae in Sapindaceae has been debated for decades (e.g., Radlkofer, 1933; Müller and Leenhouts, 1976; Umadevi and Daniel, 1991; Judd et al., 1994) and both are currently included in Sapindaceae by the

Angiosperm Phylogeny Group (APGII, 2003). However, the final decision regarding the taxonomic level of this well-supported clade (BS 100; Fig. 2) is somewhat dependant on the placement of *Xanthoceras sorbifolia*. Although *Billia* and *Handeliodendron*, thought to be close relative of *Aesculus* (Xiang et al., 1998; Forest et al., 2001), were not sampled here, the analysis confirms the definition of Hippocastanoideae as previously suggested by Judd et al. (1994) and Harrington et al. (2005). This temperate clade is characterized by deciduous opposite simple leaves (generally palmatilobate), two ovules per locule and a type-A pollen (Biesboer, 1975; Müller and Leenhouts, 1976; Fig. 1).

4.5. Subfamily Dodonaeoideae (Clade B; Figs. 2, 3)

The improved sampling for subfamily Dodonaeoideae (i.e., the addition of genera *Arfeuillea*, *Averrhoidium*, *Doratoxylon*, *Euphorianthus*, *Eurycorymbus*, *Llagunoa* and *Majidea)* allows the recognition of two moderately to well-supported clades (Fig. 3; Table 4). This topology was partially recovered by Harrington et al. (2005), but the addition of new taxa allow their delimitation based on fruit morphology: clade I (Doratoxylon group) occurs from Africa, Madagascar to Australasia and is characterized by indehiscent berry-like fruits, whereas clade II (Dodonaea group) is distributed in South America, Madagascar, Australasia and the Pacific islands (*Dodonaea viscosa* had a worldwide distribution) and comprises species with dehiscent fruits. In addition to the widespread type-A pollen occurring in both clades, specialized pollen types characterizing specific taxa occur in clade II [i.e., type-F (*Diplopeltis huegli*) and type-H (*Harpullia cupanoides*)] (George and Erdtman, 1969; Müller and Leenhouts, 1976; Fig. 1). Clades I and II have generally two ovules per locule; however a reduction to one ovule per locule occurs independently in the two clades (*Filicium* in clade I and *Euphorianthus* in clade II). Moreover, a few species of *Harpullia* (clade II), such as *H. arborea*, have 1-2 ovules per locule (Adema et al., 1994).

4.6. Subfamily Sapindoideae (clade C; Figs. 2, 4-6)

4.6.1. Early-diverging lineages (Fig. 4)

Subfamily Sapindoideae is by far the most diverse lineage in terms of species. Based on our analyses, we propose to divide it into ten groups that are discussed in light of their morphological features, geographical distribution and compared to tree topologies obtained by Harrington et al. (2005) (Figs. 4-6). The *Delavaya* group is the first lineage to diverge in Sapindoideae (clade I). Only the Chinese monotypic genus *Delavaya* is included in the present study. Results from Harrington et al. (2005) highlighted the Mexican and Texan genus *Ungnadia* (from which nuclear sequences were unavailable) as the most basal lineage in Sapindoideae. Combined plastid analyses (S. Buerki, unpubl. data) revealed a close-relationship between those two genera as suggested by Judd et al. (1994; based on morphological characters); however this relationship must be further examined using nuclear sequences. The *Delavaya* group is characterized by elongated petal base appendages and glabrous stamens (Judd et al., 1994) and the wood anatomy within the group is identical to the Cupanieae (Klaassen, 1999). The *Koelreuteria* group (clade II; BS 100), here comprising only *Koelreuteria*, is distributed in southern China and western Pacific. The study of Harrington et al. (2005) revealed a close-relationship between this genus and *Smellophylum* and *Stadmania*, distributed in East-Africa, Madagascar and the Mascarene archipelago. When a broad definition is considered, the *Koelreuteria* group shows both ancestral (type-A pollen; Müller and Leenhouts, 1976; Fig. 1) and derived characters (one ovule per locule in *Smellophylum* and *Stadmania*) and is characterized by the presence of trichomes on the anther. Since these two lineages show a disjunct distribution and transitional character states, they might be relicts of early diversification events in the subfamily (caused by long distance dispersals for example). The *Schleichera* group, which is partially recovered by Harrington et al. (2005), here with the inclusion of *Amesiodendron* (Cupanieae), is a well-supported (BS 100) tropical Asian clade (clade III; Fig. 4). This clade is characterized by a Cupanieae-like wood anatomy (Klaassen, 1999) and type-B pollen (Müller and Leenhouts, 1976; Fig. 1).

4.6.2. The Litchi group (Figs. 4, 5)

This clade (clade IV; BS 98; Fig. 4) is divided into two well-supported groups (*a* and *b*; Fig. 5). Clade *a* (BS 100) partially corresponds to the *Dimocarpus* group proposed by Müller and Leenhouts (1976; traditionally comprising *Cubilia*, *Dimocarpus*, *Litchi*, *Nephelium*, *Pometia* and *Xerospermum*) and a heterogeneous group comprising mostly African genera as well as the Indian and Australian *Lepidopetalum*. Our study also confirms the close relationships of *Pometia* (characterized by type-C1 pollen; Müller and Leenhouts, 1976; van der Ham, 1990; Fig. 1) with the other member of the *Dimocarpus* group as expected by Müller and Leenhouts (1976). The Lepisantheae-type wood anatomy of *Eriocoelum* (Cupanieae; Klaassen, 1999) confirms its relationships with the other genera of Lepisantheae from this clade. A more comprehensive analysis of this clade is currently being undertaken (S. Buerki, unpubl. data).

Clade *b* (BS 100) partially corresponds to group A of Müller and Leenhouts (1976) with the addition of *Pseudima* (Cupanieae). The inclusion of the South American *Pseudima* is supported by type-A pollen (Müller and Leenhouts, 1976; Fig. 1) and similar wood anatomy shared with other Sapindeae (Klaassen, 1999). Our results highlight the close affinities of *Lepisanthes*, *Sapindus* and *Atalaya*, but the understanding of relationships within this group will require additional data. The monophyly of the African-Malagasy *Deinbollia* is supported by molecular analyses and type-A1 pollen (Müller and Leenhouts, 1976; Fig. 1).

4.6.3. The Macphersonia group (Fig. 4)

Our study reveals for the first time relationships between southeast African and Malagasy genera (BS 100; Fig. 4). Two strongly supported clades were formed by South African *Pappea capensis* and Malagasy *Plagioscyphus* (BS 100) and Malagasy *Beguea*, *Conchopetalum* and *Haplocoelum perrieri*, as well as east African and Malagasy *Macphersonia* (BS 100). *Pappea* was previously thought to be related to other Nephelieae (*Alectryon*, *Podonephelium*, *Smelophyllum* and *Stadmania*) by Müller and Leenhouts (1976), and placed without support as sister to Paullinieae and Thouinieae by Harrington et al. (2005). The position of *Conchopetalum*, characterized by inflated fruits without arillode, in the traditional core Malagasy Schleichereae, defined by indehiscent fruits and a fleshy arillode surrounding the seed, was an unexpected result (Capuron, 1969). This clade is characterized by actinomorphic flowers, one ovule per locule (except two in *Conchopetalum*) and is distributed throughout Madagascar and southeast Africa.

4.6.4. The Cupania group (Figs. 4, 6)

The Australasian and Malagasy/South American clade VI (BS 75; Fig. 4) encloses the majority of Cupanieae genera (23 of the 32 sampled genera) and is divided into two main groups (Figs. 4, 6). In the Australasian clade *a* (BS 100), the monophyly of *Elattostachys* is well supported and the expected close relationship between the New Caledonian *Podonephelium* and Australasian and Pacific *Alectryon* is confirmed by this phylogenetic analysis and the shared type-A pollen (Müller and Leenhouts, 1976; Fig. 1). Only one non Cupanieae taxon belongs to clade *b* (BS 100): *Tinopsis apiculata* (Schleichereae). The Malagasy *Tinopsis* was first described as part of the Cupanieae (Radlkofer, 1933) and later transferred to the Schleichereae based on the indehiscence of the fruit and the presence of a fleshy arillode (Capuron, 1969). However, no floral or vegetative characters have been identified to discriminate this genus from the Malagasy Cupanieae genera *Tina* and *Neotina*. This study confirms the close relationships between these genera and supports Radlkofer’s (1933) hypothesis. This example and others encountered in clades II and V provide strong arguments supporting the convergent evolution of fruit morphology and consequently its limited systematic utility. The plasticity of fruit types has been demonstrated in several phylogenetic studies performed on a wide range of taxa (e.g., van Welzen, 1990; Adema, 1991; Muellner et al., 2003). The Cupania group is characterized by type-B pollen (except *Alectryon* and *Podonephelium* which have type-A pollen; Müller and Leenhouts, 1976; Fig. 1). In general, taxa within clade *b* present low genetic distances among them while having long terminal branches (especially the Australasian representatives such as *Cupaniopsis*, *Gongrodiscus* and *Toechima*).

4.6.5. The Paullinia group and allies (Tristiropsis, Blomia and Melicoccus groups) (Figs. 4)

Although strongly supported in general (except for the *Blomia* group; Table 4), the relationships between these four groups remain unclear (Fig. 4). The monophyly of the Australasian clade VII and the Mexico/East African clade VIII are weakly to well-supported (BS 100 and BS 61, respectively; Fig. 4). To date, no morphological characters have been identified that circumscribe these lineages. The monophyly of the South American clade IX is well supported (BS 100; Fig. 4) and confirms the suggested affinities between *Melicoccus* and *Talisia* argued by Acevedo-Rodríguez (2003) based on morphology and pollen characters.

The pantropical clade X (Figs. 4) is strongly supported (BS 100) and corresponds both to the Nomophyllae group defined by Radlkofer (1933) and to the group C proposed by Müller and Leenhouts (1976) containing Paullinieae and Thouinieae. Although no representatives of genus *Allophylus* (Thouinieae) were included here, our study confirms the results of the morphological cladistic analyses of the two tribes conducted by Acevedo-Rodríguez (1993b) and the molecular analyses of Harrington et al. (2005), which show a monophyletic Paullinieae nested in a paraphyletic Thouinieae. Our analysis indicates that the enigmatic species *Sapindus oligophyllus* has affinities with genera in this clade (Fig. 7). The generic position of this taxon has puzzled taxonomists for decades. It was first described as a member of *Aphania* and subsequently transferred in *Sapindopsis*, *Howethoa*, *Sapindus* (see Rauschert, 1982 for review) and recently merged, although informally, in *Lepisanthes* by Xia and Gadek (2007). The increase of sampling and the inclusion of *Allophylus* species might help to circumscribe the position of this taxon. Type-A pollen and the tree life-form are shared by the most basal lineages in this clade (*Athyana weinmannifolia*, *Diatenopteryx sorbifolia* and *Bridgesia incisifolia*; Acevedo-Rodriguez, 1993b; Figs. 1, 7), whereas the other taxa have a highly specialized pollen type (type-C2-3; Müller and Leenhouts, 1976; Fig. 1) and a tendency towards liana habit. Species with subtype-C pollen do not form a monophyletic group and consequently this character is of limited systematic value (e.g., type-C3 is encountered in *Thouinia* and *Paullinia*; Müller and Leenhouts, 1976; Acevedo-

Rodriguez, 1993b; Figs. 1, 4). Clade X is characterized by zygomorphic flowers, petals with a prominent scale, an unilateral disk and imparipinnate leaves. The liana habit and the development of tendrils and stipules constitute synapomorphies for Paullinieae (Fig. 4).

4.7. Informal tribal groupings within Sapindaceae

The phylogenetic analysis inferred from eight nuclear and plastid regions provides a robust assessment of the relationships within Sapindaceae s.l. (although the relationships between the subfamilies remain weakly supported) (Fig. 2). Nevertheless, the tribal delimitations as currently defined (and based largely on fruit morphology) must be revised because of the plasticity of fruit characters in this group. When Richardson et al. (2000a, b) assessed the tribal classification of Rhamnaceae (also defined by fruit morphology), they encountered the same taxonomic difficulty and proposed a new classification based on molecular data in combination with morphological characters. We follow a similar approach and propose here an informal grouping that could serve as basis for a formal reclassification of Sapindaceae s.l. based on molecular and morphological data. The family is subdivided into four subfamilies (as recognized by Thorne, 2007) and 14 groups: Xanthoceroideae, Hippocastanoideae (two groups); Dodonaeoideae (two groups) and Sapindoideae (10 groups) (Figs. 2-4). The groups within subfamilies might represent circumscriptions for the definition of future tribes.

Subfamily Xanthoceroideae

It includes the monotypic Chinese *Xanthoceras sorbifolia*, this deciduous shrub is characterized by alternate imparipinnate leaves, 6-8 ovules per locule and orange horn-like appendages protruding from the disk (Fig. 2).

Subfamily Hippocastanoideae

Temperate deciduous shrubs and trees (except *Billia* found from Mexico to tropical South America) with simple generally palmatilobate opposite leaves and 2 ovules per locule (Fig.

2). Although our sampling is limited for this subfamily, results from other studies (Judd et al., 1996; Harrington et al., 2005; S. Buerki, unpubl. data) allows us to suggest two groups within subfamily Hippocastanoideae, corresponding to the formerly recognized families Aceraceae and Hippocastanaceae (Fig. 4):

- *Acer* group (*Acer* and *Dipteronia*): leaves palmately lobed to 3-foliate or imparipinnate, or entire; actinomorphic unisexual or bisexual flowers and samara;
- *Aesculus* group (*Aesculus*, *Billia* and *Handeliodendron*): leaves palmately divided into 3-5 leaflets; zygomorphic andromonoecious flowers and dehiscent fruit with one seed.

Subfamily Dodonaeoideae

This subfamily is expanded to include *Euphorianthus*, formerly placed in Sapindoideae. The Dodonaeoideae as defined by Radlkofer (1890, 1933) are characterized by the presence of two or rarely more apotropous and upright ovules per locule, or rarely one ovule that is epitropous and pendulous. However, this does not hold anymore because of the inclusion of the above mentioned genus of Sapindoideae showing one campylotropous ovule per locule. This subfamily is divided into two groups (Fig. 3):

- *Doratoxylon* group (Doratoxyleae, without *Averrhoidium*): indehiscent berry-like fruits;
- *Dodonaea* group (Cossinieae, Dodonaeeae, *Arfeuillea*, *Averrhoidium*, *Eurycorymbus*, *Euphorianthus*, *Harpullia* and *Majidea*): dehiscent fruits.

Subfamily Sapindoideae

The subfamily Sapindoideae should be expanded to include *Conchopetalum*, formerly placed into the Dodonaeoideae. This subfamily as defined by Radlkofer (1933) are characterized by a single apotropous and upright or ascending ovule per locule; however the inclusion of several genera with two ovules per locule [*Conchopetalum* (this study), *Delavaya*, *Koelreuteria* and *Ungnadia*; Harrington et al., 2005; Thorne, 2007; this study] renders this key-character obsolete. Based on our phylogenetic analysis, ten groups are now recognized (Fig. 4):

- *Delavaya* group (*Delavaya* and *Ungnadia*): two ovules per locule; type-A pollen; elongated basal petals appendages; glabrous stamens and Cupanieae wood anatomy.
- *Koelreuteria* group (*Koelreuteria*, *Smelophyllum* and *Stadmania*): type-A pollen and trichomes on anthers.
- *Schleichera* group (*Amesiodendron*, *Paranephelium* and *Schleichera*): type-B pollen and Cupanieae-type IV wood anatomy.
- *Litchi* group [Lepisantheae, Nephelieae (without *Alectryon*, *Pappea*, *Podonephelium*, *Stadmania*, *Smelophyllum*), Sapindeae (without *Sapindus oligophyllus*)*, Blighia*, *Eriocoelum*, *Haplocoelopsis*, *Laccodiscus*, *Lecaniodiscus*, *Lepidopetalum* and *Pseudima)*]: to date, no morphological characters characterizing this group have been identified.
- *Macphersonia* group (*Beguea*, *Conchopetalum*, *Haplocoelum perrieri*, *Macphersonia*, *Pappea*, and *Plagioscyphus*): actinomorphic flowers and one ovule per locule (except two in *Conchopetalum*).
- *Cupania* group (Cupanieae [without *Amesiodendron*, *Blighia*, *Blomia*, *Dictyoneura*, *Eriocoelum*, *Haplocoelopsis*, *Laccodiscus*, *Lepidopetalum*, *Pseudima*], *Alectryon*, *Podonephelium* and *Tinopsis*): type-B pollen (except *Alectryon* and *Podonephelium*, which demonstrate type-A pollen).
- *Tristiropsis* group (*Dictyoneura* and *Tristiropsis*): to date, no morphological characters characterizing this group have been identified.
- *Blomia* group (*Blomia* and *Haplocoelum foliosum*): to date, no morphological characters characterizing this group have been identified.
- *Melicoccus* group (*Talisia* and *Melicoccus*): pollen type-A and Melicocceae wood anatomy.

- *Paullinia* group (Paullinieae, Thouinieae and *Sapindus oligophyllus*): imparipinnate leaves; zygomorphic flowers; petals with a prominent scale and a unilateral disk.

4.8. Conclusions

This study based on eight nuclear and plastid regions and 60.3% of the generic diversity of the Sapindaceae s.l. (152 samples and 139 species) (1) provides strong support for the monophyly of the family when *Xanthoceras sorbifolia*, Aceraceae and Hippocastanaceae are included (although relationships among subfamilies are still weakly supported), (2) highlights a high degree of paraphyly and polyphyly at subfamilial and tribal level, especially in Sapindaceae s.s. (subfamilies Dodonaeoideae and Sapindoideae) and (3) proposes a new informal classification for infrafamilial arrangements. Increased sampling, filled sequence gaps and the compilation of an extensive morphological matrix are now required to establish strong synapomorphies for each phylogenetic clade. A particular attention might be given to inflorescence types (and breeding systems) and floral morphology (e.g., shape and type of petal scale, type of disk, number of carpels, pubescence on the anthers, toxicity of the arillode). This might lead to a new formal infrafamilial classification for Sapindaceae s.l., based on the patterns highlighted in this study.

Acknowledgments

The first author would like to thank Professor Giorgina Bernasconi from the University of Neuchâtel (Switzerland) for help and financial support. We are grateful to the following people and institutions for the supply of plant material: the Conservatoire et Jardin botaniques de la ville de Genève (Switzerland); the Malagasy staff of the Missouri Botanical Garden in Antananarivo (Madagascar); Caroline Weckerle (University of Zürich, Switzerland); Laura Vary (University of California-Irvine; U.S.A) and Jean-Noël Labat (Muséum d'Histoire Naturelle, Paris, France). The first author thank Yong-Ming Yuan (University of Neuchâtel, Switzerland) for providing laboratory support and plant material. We thank Nils Arrigo (University of Neuchâtel; Switzerland), Ken Wurdack (Smithsonian

Institution, Washington, U.S.A.), Jérôme Munzinger (Laboratorie de Botanique, Centre IRD de Nouméa, Nouméa, New Caledonia), Peter Lowry II and Pete Phillipson (Missouri Botanical Garden, Saint-Louis; U.S.A.) for valuable comments on the manuscript and herbarium determinations. We thank the Associate Editor and two anonymous reviewers for numerous constructive comments that greatly improved the manuscript.

References

Acevedo-Rodríguez, P., 1993a. A revision of *Lophostigma* (Sapindaceae). Syst. Bot. 18, 379–388.

Acevedo-Rodríguez, P., 1993b. Systematics of *Serjania* (Sapindaceae). I. A revision of *Serjania* sect. *Platycoccus*. New York Botanical Garden, New York.

Acevedo-Rodríguez, P., 2003. Melicocceae (Sapindaceae): *Melicoccus* and *Talisia*. Flora Neotropica Monograph 87, 1–179.

Adema, F., 1991. *Cupaniopsis* Radlk. (Sapindaceae): a monograph. Leiden Botanical Series 15, 1–190.

Adema, F., Leenhouts, P.W., van Welzen, P.C., 1994. Sapindaceae. Flora Malesiana Ser. I 11, 419–768.

Akaike, H., 1973. Information theory and an extension of the maximum likelihood principle. In: Akademiai Kiado (Ed.), Second International Symposium on Information Theory. Budapest, pp. 267–281.

Angiosperm Phylogeny Group (APG II), 2003. An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APGII. Bot. J. Linn. Soc. 141, 399–436.

Baldwin, B.G., 1992. Phylogenetic utility of the internal transcribed spacers of nuclear ribosomal DNA in plants: an example from the Compositae. Mol. Phylogenet. Evol. 1, 3-16.

Bapteste, E., Brinkmann, H., Lee, J. A., Moore, D. V., Sensen, C. W., Gordon, P., Durufle, L., Gaasterland, T., Lopez, P., Müller, M., Philippe, H., 2002. The analysis of 100 genes supports the grouping of three highly divergent amoebae: *Dictyostelium*, *Entamoeba*, and *Mastigamoeba*. Proc. Natl. Acad. Sci. USA 99, 1414–1419.

Biesboer, D.D., 1975. Pollen morphology of the Aceraceae. Grana 15, 19-27.

Capuron, R., 1969. Révision des Sapindacées de Madagascar et des Comores. Mém. Mus. Natl. Hist. Nat. Ser. B Bot. 19, 1-189.

Chase, M.W., Hills, H.H., 1991. Silica-gel - An ideal material for field preservation of leaf samples for DNA studies. Taxon 40, 215-220.

Clayton, J. W., 2007. Molecular phylogeny of the tree-of-heaven family (Simaroubaceae) based on chlorplast and nuclear markers. Int. J. Plant Sci. 168, 1325-1339.

Cronquist, A., 1988. The evolution and classification of flowering plants, ed. 2. New York: New York Botanic Gardens.

Dahlgren, G., 1989. An updated system of classification. Bot. J. Linn. Soc. 100, 197–203.

Davies, F.G., 1997. A new genus *Haplocoelopsis* (Sapindaceae) from East and Central Africa. Kew Bull. 52: 719-723.

Davies, F.G., Verdcourt, B., 1998. Flora of Tropical East Africa: Sapindaceae. Rotterdam.

Demesure, B., Sodzi, N., Petit, R.J., 1995. A set of universal primers for amplification of polymorphic noncoding regions of mitochondrial and chloroplast DNA in plants. Mol. Ecol. 4, 129-131.

Doyle, J.J., Doyle, J.L., 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochem. Bull. 19, 11–15.

Driskell, A.C., Ané, C., Burleigh, J.G., McMahon, M.M., O'Meara, B.C., Sanderson, M.J., 2004. Prospects for building the Tree of Life from large sequence databases. Science 306,1172–4.

Edwards, K.J., Gadek, P.A., 2001. Evolution and biogeography of *Alectryon* (Sapindaceae). Mol. Phylogenet. Evol. 20, 14-26.

Estabrook, G. F., 1992. Evaluating undirected positional congruence of individual taxa between two estimates of the phylogenetic tree for a group of taxa. Syst. Biol. 41, 172-177.

Estabrook, G. F., McMorris, F. R., Meacham, C. A., 1985. Comparison of undirected phylogenetic trees based on subtrees of four evolutionary units. Syst. Zool. 34, 193-200.

Farris, J. S., Källersjö, M., Kluge, A.G., Bult, C., 1994. Testing significance of incongruence. Cladistics, 10, 315–319.

Felsenstein, J., 1985. Confidence-limits on phylogenies - an approach using the bootstrap. Evolution 39, 783-791.

Ferrucci, M.S., 1991. Flora del Paraguay. Vol 16. Sapindaceae. Héliographia SA, Genève.

Ferrucci, M.S., 1998. Flora fanerogamica Argentina. Sapindaceae. 52, 1-165.

Forest, F., Drouin, J.N., Charest, R., Brouillet, L., Bruneau, A., 2001. A morphological phylogenetic analysis of *Aesculus* L. and *Billia* Peyr. (Sapindaceae). Can. J. Bot. 79, 154–169.

Gadek, P.A., Fernando, E.S., Quinn, C.J., Hoot, S.B., Terrazas, T., Sheahan, M.C., Chase, M.W., 1996. Sapindales: molecular delimitation and infraordinal groups. Am. J. Bot. 83, 802–811.

George, A. S., Erdtman, G., 1969. A revision of the genus *Diplopeltis* Endl. (Sapindaceae). Grana Palynologica 9, 92-109.

Hall, T.A., 1999. Bioedit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucl. Acids. Symp. Ser. 41:95-98.

Hamilton, M.B., 1999. Four primer pairs for the amplification of chloroplast intergenic regions with intraspecific variation. Mol. Ecol. 8, 521-523.

Harrington, M.G., Edwards, K.J., Johnson, S.A., Chase, M.W., Gadek, P.A., 2005. Phylogenetic inference in Sapindaceae sensu lato using plastid *matK* and *rbcL* DNA sequences. Syst. Bot. 30, 366-382.

Huelsenbeck, J. P., 1991. When are fossils better than extant taxa in phylogenetic analysis? Syst. Zool. 40, 458–469.

Judd, W.S., Sanders, R.W., Donoghue, M.J., 1994. Angiosperm family pairs: preliminary phylogenetic analyses. Harv. Pap. Bot. 5, 1–51.

Klaassen, R., 1999. Wood anatomy of the Sapindaceae. Int. Assoc. Wood. Anat. J. 2 (suppl), 1–214.

Kluge, A.G., 1989. A Concern for evidence and a phylogenetic hypothesis of relationships among *Epicrates* (Boidae, Serpentes). Syst. Zool. 38, 7-25.

Mabberley, D.J., 2008. Mabberley's plant-book. Cambridge University Press, Cambridge.

Morrison, D.A., 2006. Multiple sequence alignment for phylogenetic purposes. Aust. Syst. Bot. 19, 479-539.

Muellner, A.N., Samuel, R., Johnson S.A., Cheek, M., Pennington, T.D., Chase M.W., 2003. Molecular phylogenetics of Meliaceae (Sapindales) based on nuclear and plastid DNA sequences. Am. J. Bot. 90, 471-480.

Muellner, A.N., Samuel, R., Chase, M.W., Pannell, C.M., Greger, H., 2005. *Aglaia* (Meliaceae): an evaluation of taxonomic concepts based on DNA data and secondary metabolites. Am. J. Bot. 92, 534–543.

Muellner, A.N., Savolainen, V., Samuel, R., Chase, M.W., 2006. The mahogany family "out-of-Africa": Divergence time estimation, global biogeographic patterns inferred from plastid *rbcL* DNA sequences, extant, and fossil distribution of diversity. Mol. Phylogenet. Evol. 40, 236–250.

Muellner, A.N., Vassiliades, D.D., Renner, S. S., 2007. Placing Biebersteiniaceae, a herbaceous clade of Sapindales, in a temporal and geographic context. Pl. Syst. Evol. 266, 233-252.

Müller, J., Leenhouts, P. W., 1976. A general survey of pollen types in Sapindaceae in relation to taxonomy. In: Ferguson, I.K., Müller, J. (Eds.), The evolutionary significance of the exine. London, Academic Press, pp. 407–445.

Munzner, T., Guimbretiere, F., Tasiran, S., Zhang, L., Zhou, Y., 2003. TreeJuxtaposer: scalable tree comparison using Focus+ context with guaranteed visibility. In: SIGGRAPH. Pp. 453-462.

Nixon, K.C., 1999. The Parsimony Ratchet, a new method for rapid parsimony analysis. Cladistics 15, 407-414.

Perrier, X., Flori, A., Bonnot, F., 2003. Data analysis methods. In: Hamon, P., Seguin, M., Perrier X., Glaszmann, J. C. (Eds.), Genetic diversity of cultivated tropical plants. Montpellier, Enfield, Science Publishers, pp. 43-76.

Phillipe, H., Snell, E.A., Bapteste, E., Lopez, P., Holland, P.W.H., Casane, D., 2004. Phylogenomics of eukaryotes: impact of missing data on large alignments. Mol. Biol. Evol. 21, 1740–52.

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

Radlkofer, L., 1890. Ueber die Gliederung der Familie der Sapindaceen. Sitz. - Ber. Akad. Wiss. Munchen 20, 105–379.

Radlkofer, L., 1933. Sapindaceae. In: Engler, A. (Ed.), Das Pflanzenreich IV, 165 (Heft 98a-h),. Leipzig, Verlag von Wilhelm Engelmann.

Rauschert, S., 1982. Nomina nova generica et combinationes novae Spermatophytorum et Pteridophytorum. Taxon 31, 554-563.

Richardson, J.E., Fay, M.F., Cronk, Q.C.B., Bowman, D., Chase M.W., 2000a. A phylogenetic analysis of Rhamnaceae using *rbcL* and *trnL-F* plastid DNA sequences. Am. J. Bot. 87, 1309-1324.

Richardson, J.E., Fay, M.F., Cronk, Q.C.B., Chase M.W., 2000b. A revision of the tribal classification of Rhamnaceae. Kew Bull. 55, 311-340.

Salamin, N., Chase, M.W., Hodkinson, T.R., Savolainen, V., 2003. Assessing internal support with large phylogenetic DNA matrices. Mol. Phylogenet. Evol. 27, 528–539.

Savolainen, V., Fay, M.F., Albach, D.C., Backlund, A., van der Bank, M., Cameron, K.M., Johnson, S.A., Lledó, M.D., Pintaud, J.-C., Powell, M., Sheahan, M.C., Soltis, D.E., Soltis, P.S., Weston, P., Whitten, W.M., Wurdack, K.J., Chase, M.W., 2000. Phylogeny of the eudicots: a newly complete familial analysis based on *rbcL* gene sequences. Kew Bull. 55, 257–309.

Sikes, D.S., Lewis, P.O., 2001. PAUPRAT, University of Connecticut, Storrs.

Soltis, D.E., Soltis, P. S., 1998. Choosing an approach and appropriate gene for phylogenetic analysis. In: Soltis, D.E., Soltis, P.E., Doyle, J. J. (Eds.), Molecular systematics of plants II: DNA sequencing. Kluwer, Boston, Massachusetts, USA, pp. 1–42.

Soltis, D.E., Soltis, P.S., Chase, M.W., Mort, M.E., Albach, T.D., Zanis, M., Savolaninen, V., Hahn, W.H., Hoot, S.B., Fay, M.F., Axtell, M., Swensen, S.M., Prince, L.M., Kress, W.J., Nixon, K.C., and Farris, J.S., 2000. Angiosperm phylogeny inferred from 18S rDNA, rbcL, and atpB sequences. Bot. J. Linn. Soc. 133, 381-461

Stamatakis, A., 2006. RAxML-VI-HPC: Maximum Likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics 22, 2688-2690.

Stamatakis, A., Hoover, P., Rougemont, J., 2008. A rapid bootstrap algorithm for the RAxML web-servers. Syst. Biol. 57, 758-771

Swofford, D.L., 2002. PAUP*: Phylogenetic analysis using parsimony (*and other methods), version 4. Sinauer Associates, Sunderland, Massachusetts.

Taberlet, P., Gielly, L., Pautou, G., Bouvet, J., 1991. Universal primers for amplification of three noncoding regions of chloroplast DNA. Plant Mol. Biol. 17, 1105-1109.

Takhtajan, A., 1987. Systema Magnoliophytorum. Leningrad: Soviet Sciences Press.

Thomas, D.W., Harris, D.J., 1999. New Sapindaceae from Cameroon and Nigeria. Kew Bull. 54, 951-957.

Thompson, J.D., Gibson, T. J., Plewniak, F., Jeanmougin, F., Higgins, D.G., 1997. The Clustal X Windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. 25, 4876-4882.

Thorne, R. F., 2000. The classification and geography of the flowering plants: dicotyledons of the class Angiospermae. Bot. Rev. 66, 441–647.

Thorne, R.F., 2007. An update classification of the class Magnoliopsida (“Angiospermae”). Bot. Rev. 73, 67-182.

Umadevi, I., Daniel, M., 1991. Chemosystematics of the Sapindaceae. Feddes Repertorium 102, 607–612.

van der Ham, R.W., 1990. Nephelieae pollen (Sapindaceae): form, function, and evolution. Leiden Botanical Series 13, 1–255.

van Welzen, P.C., 1990. *Guioa* Cav. (Sapindaceae): Taxonomy, phylogeny, and historical biogeography. Leiden Botanical Series 12, 1–315.

.

White, T.J., Bruns, T., Lee, S., Taylor, J., 1990. Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. In: Innis, M., Gelfand, D., Sninsky, J., White, T. (Eds.), PCR protocols: a guide to methods and applications. San Diego, Academic Press, pp. 315-322.

Wiens, J.J., 1998. Does adding characters with missing data increase or decrease phylogenetic accuracy? Syst. Biol. 47, 625–40.

Wiens, J.J., 2003. Missing data, incomplete taxa, and phylogenetic accuracy. Syst. Biol. 52, 528–38.

Wiens, J.J., 2005. Can incomplete taxa rescue phylogenetic analyses from long-branch attraction? Syst. Biol. 54, 731-742.

Wiens, J.J., 2006. Missing data and the design of phylogenetic analyses. J. Biomed. Inform. 39, 34–42.

Wiens, J.J., Reeder, T.W., 1995. Combining data sets with different numbers of taxa for phylogenetic analysis. Syst. Biol. 44, 548-558.

Wilkinson, M., Cotton, J. A., Creevey, C., Eulenstein, O., Harris, S. R., Lapointe, F. J., Levasseur, C., McInerney, J. O., Pisani, D., Thorley, J. L., 2005. The shape of supertrees to come: Tree shape related properties of fourteen supertree methods. Syst. Biol. 54, 419-431.

Xia, N., Gadek, P. A., 2007. Flora of China 12, 1-6 (www.efloras.org).

Xiang, Q.-Y., Crawford, D.J., Wolfe, A.D., Tang, Y.-C., DePamphilis, C.W., 1998. Origin and biogeography of *Aesculus* L. (Hippocastanaceae): a molecular phylogenetic perspective. Evolution, 52, 988–997.

Yang, Z.H., 1993. Maximum-likelihood-estimation of phylogeny from DNA-sequences when substitution rates differ over sites. Mol. Biol. Evol. 10, 1396–1401.

Figures legends

Fig. 1. Schematic representation of pollen types in Sapindaceae following Müller and Leenhouts (1976). See text for explanations regarding the morphological differentiation between pollen types.

Fig. 2. Best maximum likelihood phylogenetic tree for Sapindaceae s.l. inferred from eight nuclear and plastid nucleotide sequences. Bootstrap supports are indicated above branches. The revised infrafamilial classification based on molecular and morphological characters is indicated in grey. Abbreviations: Cossinieae (COS); Cupanieae (CUP); Dodonaeeae (DOD); Doratoxyleae (DOR); Koelreuterieae (KOE); Harpullieae (HAR); Lepisantheae (LEP); Melicocceae (MEL); Nephelieae (NEP); Paullinieae (PAU); Sapindeae (SAP); Schleichereae (SCH) and Thouinieae (THO).

Fig. 3. Relationships within subfamilies Hippocastanoideae (clade A) and Dodonaeoideae (clade B). Bootstrap supports are indicated above branches. The revised infrafamilial classification based on molecular and morphological characters is indicated in grey. See Fig. 2 for abbreviations of tribes.

Fig. 4. Relationships within subfamily Sapindoideae (clade C). Bootstrap supports are indicated above branches. The revised infrafamilial classification based on molecular and morphological characters is in grey. See Fig. 2 for abbreviations of tribes.

Fig. 5. Phylogenetic relationships within the *Litchi* group (clade C-IV; see Fig. 4). Bootstrap supports are indicated above branches. See Fig. 2 for abbreviations of tribes.

Fig. 6. Phylogenetic relationships within the *Cupania* group (clade C-VI; see Fig. 4). Bootstrap supports are indicated above branches. See Fig. 2 for abbreviations of tribes.

Table 1

Table 1. Infrafamilial classification of Sapindaceae sensu lato (Thorne, 2007). Information on number of taxa, habit and distribution of genera were taken from literature (Radlkofer, 1933; Acevedo-Rodríguez, 1993, 2003; Adema et al., 1994; Ferrucci 1991, 1998; Davies, 1997; Davies and Verdcourt, 1998; Klaassen, 1999; Thomas and Harris, 1999; Xia and Gadek, 2007; Mabberley, 2008). Abbreviations are as follows: s = shrub, st = small tree, t = tree and l = liana. Genera sampled for the phylogenetic analysis of Sapindaceae are indicated in bold and genera found to be either paraphyletic or polyphyletic are identified by an asterisk (*).

Sapindaceae Jussieu
104/141 genera, 205/1886 species

	Genera	Author	Taxa	Habit	Distribution
Dodonaeoideae Burnett					
Cossinieae Bl. (Cos)	*Cossinia*	Comm. ex Lam.	4	s-st	Mascarenes, New Caledonia, E Australia, Fiji
2/2 genera, 3/7 species	***Llagunoa***	**Ruíz & Pavón**	3	s-st	W tropical South America
Dodonaeeae Kunth (Dod)	***Diplopeltis***	**Endl.**	5	s-t	NW Australia
3/5 genera, 5/78 species	*Distichostemon*	F. Muell.	6	s	Australia
	Dodonaea	**Miller**	c. 65	s-st	Mainly in Australia, Malesia, New Guinea, Carribean and Madagascar
	Hirania	Thulin	1	s	Somalia
	Loxodiscus	**Hook. f.**	1	s	New Caledonia
Doratoxyleae Radlk. (Dor)	***Averrhoidium***	**Baillon**	2	t	South America
6/9 genera, 8/22 species	***Doratoxylon***	**Thou. ex Hook. f.**	5	st-t	Madagascar and Mascarenes Islands
	Euchorium	Eckman & Radlk.	1	t	Cuba
	Exothea	Macfad.	3	t	West Indies, Central America and Florida
	Filicium	**Thw ex Hook. f.**	3	s-st	E Africa, Madagascar and SE India
	Ganophyllum	**Blume**	2	t	W and C Africa, Andamans and Nicobars to NE Australia and Solomon Islands to Malesia
	Hippobromus	Ecklon & Zeyher	1	t	South Africa

Table 1. Continued

	Genera	Author		Taxa	Habit	Distribution
Dodonaeoideae Burnett	*Hypelate*	P. Browne		1	s-st	West Indies and Florida
Harpullieae Radlk. (Har)	*Zanha*	Hiern		4	t	Tropical Africa and Madagascar
6/6 genera, 8/34 species	***Arfeuillea***	**Pierre ex Radlk.**		1	t	SE Asia
	Conchopetalum	**Radlk.**		2	st-t	Madagascar
	Eurycorymbus	**Handel-Mazzetti**		1	t	China
	Harpullia	**Roxb.**		26	s-st	India, SE China, Malesia to Australia, New Caledonia and Pacific Islands
	Magonia	A. St. Hil.		1	t	South America
	Majidea	**J. Kirk ex Oliver**		3	t	Tropical Africa and Madagascar
Hippocastanoideae Burnett	***Acer***	**L.**		111	s-t	N temperate & tropical mountains
5/5 genera, 18/129 species	***Aesculus***	**L.**		13	t	SE Europe, India, E Asia and N America
	Billia	Peyr.		2	s-t	S Mexico to Tropical South America
	Dipteronia	**Oliver**		2	s-st	C & S China
	Handeliodendron	Rehder		1	s-t	China - deciduous
Sapindoideae Burnett						
Cupanieae Reichenb. (Cup)	***Amesiodendron***	**Hu**		1	t	China, Indo-China and Malesia
36/48 genera, 79/462 species	*Aporrhiza*	Radlk.		6	t	Tropical Africa
	Arytera	**Blume**	c.	28	s-t	Indo-Malesia to E Australia and Pacific
	Blighia	**Koenig**		4	t	Tropical Africa
	Blighiopsis	Van der Vecken		1	t	Tropical Africa
	Blomia	**Miranda**		1	t	Mexico
	Cnesmocarpon	Adema		4	s-st	Australia and Papua New Guinea
	Cupania	**L.**	c.	45	s-t	Tropical America
	Cupaniopsis*	**Radlk.**		60	s-st	Malesia, New Guinea, N-E Australia, Pacific islands, New Caledonia
	Dictyoneura	**Blume**		3	s-st	Malesia
	Dilodendron	Radlk.		1	t	South America
	Diploglottis	**Hook.f.**		12	t	NE Australia and Papua New Guinea
	Diplokeleba	N.E. Br.		2	st	South America
	Elattostachys	**(Blume) Radlk.**	c.	20	s-t	Malesia to Australia, W Pacific
	Eriocoelum	**Hook.f.**	c.	10	t	Tropical Africa

Table 1. Continued

Sapindoideae Burnett

Cupanieae Reichenb. (Cup)

Genera	Author		Taxa	Habit	Distribution
Euphorianthus	**Radlk.**		1	t	E Malesia
Gloeocarpus	Radlk.		1	t	Philippines
Gongrodiscus	**Radlk.**		3	s-t	New Caledonia
Gongrospermum	Radlk.		1	t	Philippines
Guioa*	**Cav.**		65	s-t	SE Asia, Malesia to E Australia; Pacific and New Caledonia
Haplocoelopsis	**F.G. Davies**		1	s-t	E Africa
Jagera	**Blume**		2	t	New Guinea and Australia
Laccodiscus	**Radlk.**		4	s-st	W Africa
Lepiderema	**Radlk.**		8	t	Australia and New Guinea
Lepidopetalum	**Blume**		7	s-t	India, NE Australia and Solomon Islands
Lynchodiscus	Radlk.		6	t	W Tropical Africa
Matayba*	**Aublet.**	c.	56	s-t	Tropical America
Mischarytera	(Radlk.) H.Turner		3	t	Australia, Papua New Guinea
Mischocarpus	**Blume**		15	s-t	SE Asia, Malesia to Australia
Molinaea	**Comm. ex. Juss.**		9	s-t	Madagascar, Mascarenes
Neotina	**Capuron**		2	t	Madagascar
Paranephelium	**Miq.**		4	s-t	SE Asia and W Malesia
Pavieasia	Pierre		3	t	S China, N Vietnam
Pentascyphus	Radlk.		1	t	Guyana
Phyllotrichum	Thorel ex Lecompte		1	t	SE Asia
Pseudima	**Radlk.**		3	t	South America
Rhysotoechia	**Radlk.**		14	s-t	Australia, New Guinea, Malesia
Sarcopteryx	**Radlk.**		12	s-t	Malesia, New Guinea and E Australia
Sarcotoechia*	**Radlk.**		11	t	NE Australia and New Guinea
Scyphonychium	Radlk.		1	t	NE Brazil
Sisyrolepis	Radlk.		1	s-st	Thailand
Storthocalyx	**Radlk.**		4	s	New Caledonia
Synima	**Radlk.**		2	t	Australia and SE New Guinea
Tina	**Roem. & Schult.**		6	s-st	Madagascar
Toechima	**Radlk.**		7	t	Australia and New Guinea
Trigonachras	Radlk.		8	t	Malesia
Tripterodendron	Radlk.		1	t	Brazil
Vouarana	**Aublet.**		1	t	NE South America

Table 1. Continued

	Genera	Author		Taxa	Habit	Distribution
Sapindoideae Burnett						
Koelreuterieae Radlk. (Koe)	*Erythrophysa*	E. Mey ex Arnott		9	s	Africa and Madagascar
2/4 genera, 2/15 species	***Koelreuteria***	**Laxmann**		3	t	S China, Japan
	Sinoradlkofera	F.G. Mey		2	st	China and N Vietnam
	Stocksia	Benth.		1	s	E Iran, Afghanistan
Lepisantheae Radlk. (Lep)	*Chonopetalum*	Radlk.		1	t	Tropical W Africa
4/10 genera, 7/97 species	***Chytranthus***	**Hook f.**	c.	30	st	Africa
	Glenniea	**Hook.f.**		8	t	Tropical Africa, Madagascar, Sri Lanka, Malesia
	Lepisanthes	**Blume**		24	s-t	Tropical Africa, Madagascar, S-SE Asia, Malesia and NW Australia
	Namataea	D.W. Thomas & D.J. Harris		1	st	Cameroon
	Pancovia	**Willd.**	c.	13	st	Tropical Africa
	Placodiscus	Radlk.	c.	15	t	Tropical W Africa
	Pseudopancovia	Pellegrin		1	t	Tropical W Africa
	Radlkofera	Gilg.		1	s-st	Tropical Africa
	Zollingeria	Kurz		3	t	SE Asia and Malesia
Melicocceae Blume (Mel)	*Castanospora*	F. Muell.		1	t	NE Australia
5/5 genera, 8/67 species	***Melicoccus***	**P. Browne**		10	t	Tropical America
	Talisia	**Aublet**		52	s-t	Tropical America
	Tristira	Radlk.		1	t	Malesia
	Tristiropsis	**Radlk.**		3	t	Pacific Ocean, Australia, Solomon Islands and Malesia
Nephelieae Radlk. (Nep)	***Alectryon***	**Gaertn.**	c.	30	s-st	E Malesia, Australia, New Zealand, New Caledonia, to Hawaii
11/12 genera, 15/77 species	***Cubilia***	**Blume**		1	t	Malesia
	Dimocarpus	**Lour.**		6	s-t	S and SE Asia and Australia
	Litchi	**Sonn.**		1	t	Tropical China to W Malesia
	Nephelium	**L.**		22	t	SE Asia and Malesia
	Otonephelium	Radlk.		1	t	India

Table 1. Continued

	Genera	Author	Taxa	Habit	Distribution
Sapindoideae Burnett	***Pappea***	**Eckl. & Zeyh.**	1	s-t	Tropical E to S Africa
Nephelieae Radlk. (Nep)	***Podonephelium***	**Baillon**	4	s-t	New Caledonia
	Pometia	**Forst. & Forst.**	2	t	Malesia and Pacific Islands
	Smelophyllum	Radlk.	1	t	South Africa
	Stadmania	Lam.	6	t	Tropical E Africa, S Africa, Madagascar and Mascarenes Islands
	Xerospermum	**Blume**	2	s-st	Indochinese Peninsula and Malesia
Paullinieae Kunth (Pau)	***Cardiospermum***	**L.**	c. 12	l	Tropical and subtropical America; 1 sp. extending to Africa
4/7 genera, 15/466 species	*Houssayanthus*	Hunz.	3	s-l	South America
	Lophostigma	Radlk.	2	l	South America
	Paullinia	**L.**	c. 200	l	Tropical America and one pantropical sp.
	Serjania	**Miller**	c. 226	l	Tropical America
	Thinouia	Triana & Planchon	9	l	Tropical America
	Urvillea	**Kunth**	14	l	Tropical America
Sapindeae DC (Sap)	***Atalaya***	**Blume**	12	st	Australia, New Guinea and S Africa
3/7 genera, 12/89 species	***Deinbollia***	**Schumach. & Thonn.**	c. 40	t	Tropical Africa and Madagascar
	Hornea	Baker	1	s-t	Mauritius
	Porocystis	Radlk.	2	s-t	Tropical South America
	Sapindus	**L.**	13	t	Tropical to warm temperate regions
	Thouinidium	Radlk.	7	s-t	Mexico and West Indies
	Toulicia	Aublet	14	t	South America
Schleichereae Radlk. (Sch)	***Beguea***	**Capuron**	1	t	Madagascar
8/12 genera, 12/55 species	*Bizonula*	Pellegrin	1	t	Tropical Africa
	Camptolepis	Radlk.	4	t	E Africa and Madagascar
	Chouxia	Capuron	6	s-st	Madagascar
	Haplocoelum*	**Radlk.**	c. 6	st-t	Tropical Africa and Madagascar
	Lecaniodiscus	**Planch. ex Benth.**	3	st	Tropical Africa
	Macphersonia	**Blume**	8	s-t	Tropical E Africa and Madagascar
	Plagioscyphus	**Radlk.**	10	st-t	Madagascar
	Pseudopteris	Baill.	3	s	Madagascar

Table 1. Continued

	Genera	Author	Taxa	Habit	Distribution
Sapindoideae Burnett	***Schleichera***	**Willd.**	1	t	Tropical SE Asia to Indo-China and Malesia
Schleichereae Radlk. (Sch)	***Tinopsis***	**Radlk.**	11	t	Madagascar
	Tsingya	Capuron	1	t	Madagascar
Thouinieae Bl. (Tho)	*Allophylus*	L.	c. 250	s-st-l	Pantropical
6/6 genera, 10/285 species	***Athyana***	**(Griseb.) Radlk.**	1	t	South America
	Bridgesia	**Bertero ex Cambess.**	1	s-st	Chile
	Diatenopteryx	**Radlk.**	2	t	South America
	Guindilia	Hook & Arn.	3	s	South America
	Thouinia	**Poit.**	28	l	Mexico and West Indies
Sapindoideae unplaced taxa	***Delavaya***	**Franchet**	1	s-st	SW China and N Vietnam
2/2 genera, 2/2 species	*Ungnadia*	Endl.	1	s-st	S North America
Xanthoceroideae Thorne & Reveal 1/1 genera, 1/1 species	***Xanthoceras***	**Bunge**	1	s-st	N-NE China and Korea

Table 2

Table 2. Characteristics of partitions used in the phylogenetic analyses of the Sapindaceae s.l.. See text for explanations regarding the compilation of combined data sets (i.e. “4 markers” and “4+4 markers”). IGS, intergenic spacer; the asterisk (*) indicates markers included in the combined “4 markers” phylogenetic analysis. MP, maximum parsimony.

	Single-gene analysis								**Combined analyses**	
Phylogenetic information	*ITS**	*matK*	*rpoB**	*trnD-trnT* IGS	*trnK-matK* IGS	*trnL* intron*	*trnL-trnF* IGS*	*trnS-trnG* IGS	4 markers*	4*+4 markers
No. of ingroup sampled species/genera	139/84	110/69	139/84	79/62	100/63	139/84	139/84	67/47	139/84	139/84
No. sequences incl. outgroup (in brackets, total number of samples for the combined analyses)	153	119	154	85	109	154	154	69	615 (154)	997 (154)
Sequence length range	650-705	1074-1242	357-363	1086-1425	705-753	510-522	380-430	1311-1365	-	-
Alignment length	1234	1614	363	1925	931	773	661	2156	3031	9657
Missing data (percentage of ingroup sequences; in brackets percentage of nucleotides for the combined analyses)	0	21.7	0	44.1	28.3	0	0	54.6	0 (0)	18.6 (27.3)
No. constant characters (%)	599 (48.5)	1144 (70.9)	228 (62.8)	1096 (56.9)	530 (56.9)	489 (63.3)	359 (54.3)	1643 (76.2)	1675 (55.3)	6088 (63.0)
No. variable characters (%)	635 (51.4)	470 (29.1)	135 (37.2)	829 (43.1)	401 (43.1)	284 (36.7)	302 (45.7)	513 (23.8)	1356 (44.7)	3569 (37.0)
No. potentially parsimony-informative (PI) characters (%)	467 (37.8)	295 (18.3)	95 (26.2)	405 (21.0)	230 (24.7)	166 (21.5)	190 (28.7)	195 (9.0)	918 (30.3)	2043 (21.2)
Mean amount of phylogenetic information per sample (averaged by variable sites number/PI sites number)	-	-	-	-	-	-	-	-	8.8/6.0	23.2/13.3
No. trees retained (MP)	525	1991	1997	2001	1786	1707	1023	1190	1138	1010
Tree length (MP ; step)	4365	837	246	1447	790	578	576	798	5889	9843
Consistency Index (MP)	0.282	0.701	0.707	0.731	0.675	0.681	0.680	0.741	0.372	0.504
Retention Index (MP)	0.640	0.849	0.899	0.811	0.790	0.879	0.861	0.665	0.698	0.726

Table 3. Level of topological agreement (based on EA distances) between total evidence trees inferred from the “4 markers” and “4+4 markers” data sets. See text for explanations regarding the compilation of these data sets. MP, maximum parsimony; ML, maximum likelihood.

	1	2	3	4
1- ML "4 markers"	-			
2- MP "4 markers"	0.177	-		
3- ML "4+4 markers"	0.021	0.189	-	
4- MP "4+4 markers"	0.028	0.173	0.027	-

Table 4

Table 4. Summary of the bootstrap support for each clade recovered in the four total evidence trees (two data sets and two algorithms). Bootstrap supports for clade C-I are not indicated because this lineage is only composed by *Delavaya yunnanensis*. Note: Although monophyletic, clade C-V is nested into clade C-VI, the latter is not recovered by the MP analysis based on the “4 markers” data set. MP, maximum parsimony; ML, maximum likelihood.

	ML		**MP**	
Combined data sets	**4 markers**	**4+4 markers**	**4 markers**	**4+4 markers**
Sapindaceae s.l.	94	91	97	97
Clade A + Clade B + Clade C	65	58	60	57
Clade A	100	100	99	99
Clade B	94	91	99	99
B-I	100	100	100	100
B-II	88	77	86	86
Clade C	92	87	96	98
C-II	100	100	100	100
C-III	100	100	100	100
C-IV	77	98	<50	65
C-IV-a	73	100	73	100
C-IV-b	100	100	<50	83
C-V	100	100	99	100
C-VI	60	75	-	<50
C-VI-a	69	89	<50	65
C-VI-b	65	83	<50	58
C-VII	100	100	100	100
C-VIII	60	61	50	70
C-IX	100	100	99	100
C-X	100	100	93	100

Appendix

Appendix. Voucher information and GenBank accession numbers for taxa used in the phylogenetic analysis of family Sapindaceae s.l. (including outgroups). Abbreviations: ANH = Andong National University, South Korea; BBG = Bogor Botanic Garden, India, living collections; CSIRO = CSIRO Arboretum, Australia; G = Conservatoire et Jardin Botanique de la ville de Genève, Switzerland; JCT = James Cook University of Northern Queensland, Australia; K = Royal Botanic Gardens, Kew, U.K.; NEU = Neuchâtel, Switzerland; MO = Missouri Botanical Garden, U.S.A.; P = Muséum d'Histoire Naturelle, France; RBG = Royal Botanic Gardens, Kew, U.K., living collections; US = Smithsonian Institution, U.S.A and Z = University of Zürich, Switzerland.

						GenBank accession number							
Genera	**species**	**Author**	**Voucher**	**Herbarium**	**Country**	**ITS**	**matK**	**rpoB**	**trnD-trnT**	**trnK-matK**	**trnL**	**trnL-F**	**trnS-trnG**
Ingroup													
Acer	*erianthum*	Schwer.	*Chase 19983*	K	China	EU720501	-	EU720843	EU720980	-	EU721271	EU721459	-
Acer	*saccharum*	Marshall	*Chase 106*	K	Cult. source, Orange Co.	EU720502	-	EU720844	-	-	EU721272	EU721460	-
Aesculus	*indica*	(Wall. ex Cambess.) Hook.	*Chase 19987*	K	India	EU927392	-	EU720845	EU720981	-	EU721273	EU721461	-
Alectryon	*connatus*	Radlk.	*Chase 2047*	K	Australia	EU720415	EU720577	EU720732	EU720928	EU721025	EU721169	EU721357	EU721534
Amesiodendron	*chinensis*	(Merr.) Hu	*Yuan s.n.*	NEU	China	EU720403	-	EU720718	EU720917	-	EU721155	EU721344	EU721525
Arfeuillea	*arborescens*	Pierre	*Chase 2122*	K	Bogor, BG	EU720461	EU720629	EU720793	EU720962	EU721067	EU721229	EU721417	-
Arytera	*littoralis*	Blume	*Yuan s.n.*	NEU	China	EU720405	EU720566	EU720720	EU720919	EU721018	EU721157	EU721346	EU721527
Arytera	*littoralis*	Blume	*Chase 2123*	K	Bogor, BG	EU720462	EU720630	EU720794	EU720963	EU721068	EU721230	EU721418	-
Atalaya	*alata*	(Sim) H. Forbes	*Edwards KE228*	JCT	South Africa	EU720425	EU720593	EU720748	EU720939	EU721036	EU721184	EU721372	EU721543
Atalaya	*capense*	R.A. Dyer	*Edwards KE 509*	JCT	South Africa	EU720429	-	EU720752	-	-	EU721188	EU721376	-
Athyana	*weinmannifolia*	(Griseb.) Radlk.	*Pennington 17581*	MO	Peru	EU720487	EU720649	EU720824	EU720975	EU721086	EU721257	EU721445	EU721576
Averrhoidium	*dalyi*	Acev.-Rodr. & Ferrucci	*Weckerle 00/03/18-1/1*	Z	Peru	EU720495	-	EU720836	-	-	EU721268	EU721456	-
Beguea	*apetala*	Capuron	*Buerki 149*	NEU	Madagascar	EU720491	EU720652	EU720828	EU720978	EU721089	EU721261	EU721449	-
Beguea	*apetala*	Capuron	*Vary 40*	MO	Madagascar	EU720512	EU720663	EU720856	-	EU721100	EU721281	EU721469	-
Blighia	*sapida*	K.D. Koenig	*Edwards KE86*	JCT	West Africa	EU720416	EU720578	EU720733	EU720929	EU721026	EU721170	EU721358	EU721535
Blomia	*prisca*	(Standl.) Lundell	*Acevedo 12242*	US	Mexico, Yucatan	EU720444	EU720611	EU720772	-	EU721050	EU721208	EU721396	-
Bridgesia	*incisifolia*	Bertero ex Cambess.	*Killip & Pisano 39778*	K	Chile	EU720476	EU720645	EU720811	EU720973	EU721082	EU721247	EU721435	-
Cardiospermum	sp.		*Yuan s.n.*	NEU	China	EU720399	-	EU720713	EU720912	-	EU721150	EU721339	-
Chytranthus	*carneus*	Radlk.	*Chase 2868*	RBG	-	EU720477	EU720646	EU720812	EU720974	EU721083	EU721248	EU721436	EU721575
Conchopetalum	*brachysepalum*	Capuron	*Rabarimanarivo 8*	MO	Madagascar	EU720530	EU720680	EU720877	-	EU721117	EU721299	EU721487	EU721586
Cubilia	*cubili*	(Blanco) Adelb.	*Chase 2125*	K	Bogor, BG	EU720463	EU720631	EU720795	EU720964	EU721069	EU721231	EU721419	EU721567

						GenBank accession number							
Genera	**species**	**Author**	**Voucher**	**Herbarium**	**Country**	**ITS**	**Genera**	**species**	**Author**	**Voucher**	**Herbarium**	**Country**	**ITS**
Cupania	*dentata*	DC.	*Acevedo 12241*	US	Mexico, Yucatan	EU720523	EU720670	EU720867	EU720988	EU721107	EU721289	EU721477	EU721581
Cupania	*hirsuta*	Radlk.	*Acevedo 1101*	US	French Guiana	EU720521	EU720668	EU720865	-	EU721105	EU721287	EU721475	-
Cupania	*rubiginosa*	(Poir.) Radlk.	*Mori 8868*	MO	French Guiana	EU720481	-	EU720817	-	-	EU721251	EU721439	-
Cupania	*scrobiculata*	Rich.	*Acevedo 11100*	US	French Guiana	EU720524	EU720671	EU720868	EU720989	EU721108	EU721290	EU721478	-
Cupaniopsis	*anacardioides*	Radlk.	*Chase 217*	K	Australia	EU720438	EU720605	EU720763	EU720946	EU721045	EU721199	EU721387	EU721552
Cupaniopsis	*flagelliformis*	(Bailey) Radlk.	*Edwards KE42*	JCT	Australia	EU720432	EU720598	EU720755	EU720942	-	EU721191	EU721379	EU721547
Cupaniopsis	*fruticosa*	Radlk.	*Munzinger 564*	MO	New Caledonia	EU720533	-	EU720881	-	EU721119	EU721302	EU721490	-
Cupaniopsis	sp.		*Munzinger 710*	MO	New Caledonia	EU720532	-	EU720880	EU720996	-	EU721301	EU721489	EU721587
Cupaniopsis	sp.		*Munzinger 1103*	MO	New Caledonia	EU720507	EU720660	EU720851	-	EU721097	EU721278	EU721466	-
Deinbollia	*borbonica*	Scheff.	*Edwards KE197*	JCT	Tanzania	EU720412	EU720574	EU720729	-	-	EU721166	EU721354	EU721532
Deinbollia	*macrocarpa*	Capuron	*H. Razafindraibe 118*	MO	Madagascar	EU720535	EU720683	EU720883	-	EU721121	EU721304	EU721492	EU721589
Deinbollia	*macrocarpa*	Capuron	*Buerki 144*	NEU	Madagascar	EU720503	EU720656	EU720847	-	EU721093	EU721275	EU721463	-
Deinbollia	*oblongifolia*	(E. Mey. ex Arn.) Radlk.	*Edwards KE233*	JCT	South Africa	EU720427	EU720595	EU720750	-	-	EU721186	EU721374	EU721545
Deinbollia	*pervillei*	(Blume) Radlk.	*Phillipson 5919*	MO	Madagascar	EU720395	EU720560	EU720708	-	EU721012	EU721145	EU721334	-
Deinbollia	*pervillei*	(Blume) Radlk.	*Callmander 688*	MO	Madagascar	EU720514	-	EU720858	-	-	EU721283	EU721471	-
Delavaya	*yunnanensis*	Franch.	*Forrest 20682*	MO	China, Yunnan	EU720484	-	EU720821	-	-	EU721254	EU721442	-
Diatenopteryx	*sorbifolia*	Radlk.	*Zardini 43371*	MO	Paraguay	EU720534	EU720682	EU720882	-	EU721120	EU721303	EU721491	EU721588
Dictyoneura	*obtusa*	Blume	*Edwards KE142*	JCT	Australia	EU720428	-	EU720751	-	-	EU721187	EU721375	-
Dimocarpus	*australianus*	Leenh.	*Edwards KE34*	JCT	Australia	EU720433	-	EU720757	-	-	EU721193	EU721381	-
Diploglottis	*campbelli*	Cheel	*Chase 2048*	K	Australian, BG	EU720457	EU720624	EU720788	EU720960	EU721062	EU721224	EU721412	-
Diplopeltis	*huegelii*	Endl.	*Chase 2192*	K	Australia	EU720473	EU720642	EU720807	EU720971	EU721079	EU721243	EU721431	-
Dipteronia	*sinensis*	Oliv.	*Chase 502*	RBG	-	EU720445	EU720612	EU720774	-	-	EU721210	EU721398	-
Dodonaea	*madagascariensis*	Radlk.	*Bocksberger GB028*	NEU	Madagascar	EU720518	-	EU720862	EU720984	-	EU721284	EU721472	-
Dodonaea	*viscosa*	Jacq.	*Razafitsalama 956*	MO	Madagascar	EU720519	EU720666	EU720863	EU720985	EU721103	EU721285	EU721473	-

						GenBank accession number							
Genera	**species**	**Author**	**Voucher**	**Herbarium**	**Country**	**ITS**	**Genera**	**species**	**Author**	**Voucher**	**Herbarium**	**Country**	**ITS**
Dodonaea	*viscosa*	Jacq.	*Merello 1077*	MO	Peru	EU720536	EU720684	EU720884	EU720997	EU721122	EU721305	EU721493	-
Dodonaea	*viscosa*	Jacq.	*Yuan s.n.*	NEU	China	EU720406	EU720567	EU720721	EU720920	EU721019	EU721158	EU721347	-
Doratoxylon	*chouxii*	Capuron	*Labat JNL3543*	P	Madagascar	EU720394	EU720559	EU720707	EU720908	EU721011	EU721144	EU721333	-
Doratoxylon	*chouxii*	Capuron	*Callmander 679*	MO	Madagascar	EU720513	EU720664	EU720857	-	EU721101	EU721282	EU721470	-
Elattostachys	*apetala*	Radlk.	*Munzinger 692*	MO	New Caledonia	EU720537	EU720685	EU720885	EU720998	EU721123	EU721306	EU721494	EU721590
Elattostachys	*apetala*	Radlk.	*McPherson 18184*	MO	New Caledonia	EU720538	EU720686	EU720886	EU720999	EU721124	EU721307	EU721495	EU721591
Elattostachys	*microcarpa*	S.T. Reynolds	*Edwards KE98*	JCT	Australia	EU720409	EU720571	EU720726	-	-	EU721163	EU721351	-
Elattostachys	*nervosa*	(F. Muell.) Radlk.	*Chase 2022*	K	Australian, BG	EU720455	EU720622	EU720786	EU720959	EU721060	EU721222	EU721410	EU721563
Elattostachys	sp.		*Lowry 5650A*	MO	New Caledonia	EU720529	EU720679	EU720876	EU720994	EU721116	EU721298	EU721486	EU721585
Eriocoelum	*kerstingii*	Gilg ex Engl.	*Merello 1586*	MO	Ghana	EU720539	EU720687	EU720887	EU721000	EU721125	EU721308	EU721496	EU721592
Eriocoelum	*microspermum*	Radlk.	*Bradley 1025*	MO	Gabon	EU720540	EU720688	EU720888	EU721001	EU721126	EU721309	EU721497	EU721593
Euphorianthus	*longifolius*	Radlk.	*Chase 2126*	K	Bogor, BG	EU720464	-	EU720796	-	-	EU721232	EU721420	-
Eurycorymbus	*cavalerieri*	(H. Lév.) Rehder & Hand.-Mazz.	*Yuan s.n.*	NEU	China	EU720404	EU720565	EU720719	EU720918	EU721017	EU721156	EU721345	EU721526
Filicium	*decipiens*	(Wight & Arn.) Thwaites	*Chase 2128*	K	Bogor, BG	EU720466	EU720633	EU720798	-	-	EU721234	EU721422	-
Filicium	*longifolium*	(H. Perrier) Capuron	*Rabenantonadro 1113*	MO	Madagascar	EU720541	-	EU720889	-	-	EU721310	EU721498	-
Filicium	*thouarsianum*	(A. DC.) Capuron	*Antilahimena 5021*	MO	Madagascar	EU720493	-	EU720832	-	-	EU721265	EU721453	-
Ganophyllum	*falcatum*	Blume	*Chase 2129*	K	Bogor, BG	EU720467	EU720634	EU720799	-	EU721071	EU721235	EU721423	-
Glenniea	*pervillei*	(Baill.) Leenh.	*Andriamihajarivo 1053*	MO	Madagascar	EU720490	EU720651	EU720827	EU720977	EU721088	EU721260	EU721448	-
Gongrodiscus	*bilocularis*	H.Turner	*Munzinger 749*	MO	New Caledonia	EU720542	EU720689	EU720890	-	EU721127	EU721311	EU721499	-
Guioa	*glauca*	Radlk.	*McPherson 18230*	MO	New Caledonia	EU720545	EU720692	EU720893	-	EU721130	EU721315	EU721503	-
Guioa	*microsepala*	Radlk.	*Munzinger 744*	MO	New Caledonia	EU720546	EU720693	EU720894	-	EU721131	EU721316	EU721504	EU721596
Guioa	*semiglauca*	(F. Muell.) Radlk.	*Chase 2058*	K	Australian, BG	EU720458	EU720625	EU720789	-	EU721063	EU721225	EU721413	-
Guioa	*villosa*	Radlk.	*McPherson 18040*	MO	New Caledonia	EU720544	EU720691	EU720892	EU721003	EU721129	EU721314	EU721502	EU721595

						GenBank accession number							
Genera	**species**	**Author**	**Voucher**	**Herbarium**	**Country**	**ITS**	**Genera**	**species**	**Author**	**Voucher**	**Herbarium**	**Country**	**ITS**
Guioa	sp.		*Munzinger 945*	MO	New Caledonia	EU720505	EU720658	EU720849	-	EU721095	EU721277	EU721465	-
			Edwards KE276	JCT	Tanzania								
Haplocoelopsis	*africana*	F.G. Davies				EU720441	EU720608	EU720767	EU720949	-	EU721203	EU721391	EU721555
Haplocoelum	*foliosum*	(Hiern) Bullock	*Friis 1894*	MO	Ethiopia	EU720479	-	EU720815	-	-	EU721250	EU721438	-
			Edwards KE195	JCT	Tanzania								
Haplocoelum	*foliosum subsp. foliosum*	(Hiern) Bullock				EU720410	EU720572	EU720727	EU720924	-	EU721164	EU721352	EU721530
Haplocoelum	*perrieri*	Capuron	*Rakotomalaza 1165*	MO	Madagascar	EU720396	-	EU720709	EU720909	-	EU721146	EU721335	EU721519
Harpullia	*arborea*	(Blanco) Radlk.	*Chase 1353*	K	Bogor, BG	EU720448	-	EU720779	-	-	EU721215	EU721403	-
Jagera	*javanica*	(Blume) Blume ex Kalkman	*Chase 2130*	K	Bogor, BG	EU720468	EU720635	EU720800	-	EU721072	EU721236	EU721424	EU721569
			Edwards KE178	JCT	Australia								
Jagera	*javanica* subsp. *australiana*	Leenh.				EU720442	-	EU720769	-	-	EU721205	EU721393	EU721556
Koelreuteria	*paniculata*	Laxm.	*Harder 5668*	MO	Vietnam	EU720548	EU720695	EU720896	-	EU721133	EU721318	EU721506	-
Koelreuteria	*paniculata*	Laxm.	*Yuan CN2006-3*	NEU	China	EU720397	EU720561	EU720710	-	EU721013	EU721147	EU721336	EU721520
Koelreuteria	sp.		*Harder 5724*	MO	Vietnam	EU720547	EU720694	EU720895	EU721004	EU721132	EU721317	EU721505	-
Laccodiscus	*klaineanus*	Pierre ex Engl.	*Walters 1269*	MO	Gabon	EU720549	EU720696	EU720897	-	EU721134	EU721319	EU721507	-
			Edwards KE194	JCT	Tanzania								
Lecaniodiscus	*fraxinifolius*	Baker				EU720418	EU720580	EU720735	EU720931	EU721028	EU721172	EU721360	EU721536
			Edwards KE36	JCT	Australia								
Lepiderema	*hirsuta*	S.T. Reynolds				EU720435	EU720601	EU720759	-	EU721041	EU721195	EU721383	EU721549
Lepiderema	*pulchella*	Radlk.	*Chase 2020*	K	Australian, BG	EU720454	-	EU720785	EU720958	-	EU721221	EU721409	-
				JCT									
Lepidopetalum	*fructoglabrum*	Welzen	*Edwards KE139*		Australia	EU720408	-	EU720724	EU720922	-	EU721161	EU721349	EU721528
Lepisanthes	*alata*	(Blume) Leenh.	*Chase 1355*	K	Bogor, BG	EU720450	EU720618	EU720781	-	EU721056	EU721217	EU721405	-
Lepisanthes	*feruginea*	(Radlk.) Leenh.	*Chase 1354*	K	Bogor, BG	EU720449	EU720617	EU720780	-	EU721055	EU721216	EU721404	-
Lepisanthes	*rubiginosa*	(Roxb.) Leenh.	*Chase 1350*	K	Bogor, BG	EU720446	EU720614	EU720776	EU720952	EU721052	EU721212	EU721400	EU721558
Lepisanthes	*senegalensis*	(Poir.) Leenh.	*Callmander 627*	MO	Madagascar	EU720492	EU720654	EU720830	EU720979	EU721091	EU721263	EU721451	EU721577
Litchi	*chinensis*	Sonn.	*Yuan s.n.*	NEU	China	EU720400	EU720564	EU720715	EU720914	EU721016	EU721152	EU721341	EU721522
Llagunoa	*mollis*	Kunth	*Jaramillollejia 3199*	MO	Colombia	EU720482	-	EU720818	-	-	EU721252	EU721440	-
Llagunoa	*nitida*	Ruiz & Pav.	*Pennington 17552*	MO	Peru	EU720486	-	EU720823	-	-	EU721256	EU721444	-
Loxodiscus	*coriaceus*	Hook. f.	*Bradford 1136*	MO	New Caledonia	EU720488	-	EU720825	-	-	EU721258	EU721446	-

						GenBank accession number							
Genera	**species**	**Author**	**Voucher**	**Herbarium**	**Country**	**ITS**	**Genera**	**species**	**Author**	**Voucher**	**Herbarium**	**Country**	**ITS**
Macphersonia	*chapelieri*	(Baill.) Capuron	*Buerki 138*	NEU	Madagascar	EU720459	EU720627	EU720791	EU720961	EU721065	EU721227	EU721415	EU721566
Macphersonia	*gracilis*	O. Hoffm.	*Rabenantoandro 1081*	MO	Madagascar	EU720550	EU720697	EU720898	EU721005	EU721135	EU721320	EU721508	EU721597
Majidea	*zanguebarika*	Kirk ex Oliv.	*TH275*	MO	Madagascar	EU720552	-	EU720900	EU721006	-	EU721322	EU721510	-
Matayba	*apetala*	Radlk.	*Acevedo 11929*	US	Jamaica	EU720526	EU720674	EU720871	-	EU721111	EU721293	EU721481	EU721583
Matayba	cf. *opaca*	Radlk.	*Acevedo 11118*	US	French Guiana	EU720522	EU720669	EU720866	EU720987	EU721106	EU721288	EU721476	EU721580
Matayba	*domingensis*	(DC.) Radlk.	*Taylor 11819*	MO	Caribbean	EU720551	EU720698	EU720899	-	EU721136	EU721321	EU721509	EU721598
Matayba	*elaeagnoides*	Radlk.	*Zardini 43278*	MO	Paraguay	EU720553	EU720699	EU720901	-	EU721137	EU721323	EU721511	-
Matayba	*guianensis*	Aubl.	*Acevedo 12342*	US	French Guiana	EU720527	EU720675	EU720872	-	EU721112	EU721294	EU721482	-
Matayba	*laevigata*	Radlk.	*Acevedo 12357*	US	French Guiana	EU720528	EU720676	EU720873	EU720992	EU721113	EU721295	EU721483	-
Melicoccus	*bijugatus*	Jacq.	*Acevedo s.n.*	US	Puerto Rico	EU927391	EU720610	EU720771	-	EU721049	EU721207	EU721395	-
Melicoccus	*lepidopetalus*	Radlk.	*Acevedo 11128*	US	Bolivia	EU720443	-	EU720770	-	-	EU721206	EU721394	-
Mischarytera	sp.	-	*Edwards KE159*	JCT	Australia	EU720417	EU720579	EU720734	EU720930	EU721027	EU721171	EU721359	-
Mischocarpus	*exangulatus*	(F. Muell.) Radlk.	*Edwards KE30*	JCT	Australia	EU720434	EU720600	EU720758	EU720943	EU721040	EU721194	EU721382	-
Mischocarpus	*grandissumus*	Radlk.	*Edwards KE37*	JCT	Australia	EU720437	EU720604	EU720762	EU720945	EU721044	EU721198	EU721386	EU721551
Mischocarpus	*pentapetalus*	(Rox.) Radlk.	*Chase 2133*	K	Bogor, BG	EU720470	EU720637	EU720802	EU720966	EU721074	EU721238	EU721426	EU721571
Mischocarpus	*pyriformis*	(F. Muell.) Radlk.	*Chase 2059*	K	Australian, BG	EU720460	EU720628	EU720792	-	EU721066	EU721228	EU721416	-
Molinaea	*petiolaris*	Radlk.	*Rabenantoandro 1448*	MO	Madagascar	EU720554	EU720700	EU720902	EU721007	EU721138	EU721324	EU721512	-
Molinaea	sp. nov.		*Antilahimena 4301*	MO	Madagascar	EU720510	EU720662	EU720854	EU720983	EU721099	EU721280	EU721468	EU721578
Neotina	*coursii*	Capuron	*H. Razafindraibe 119*	MO	Madagascar	EU720543	EU720690	EU720891	EU721002	EU721128	EU721313	EU721501	EU721594
Nephelium	*lappaceum (=N. chryseum)*	L.	*Yuan s.n.*	NEU	China	EU720401	-	EU720716	EU720915	-	EU721153	EU721342	EU721523
Pancovia	*golungensis*	(Hiern) Exell & Mendonça	*Edwards KE231*	JCT	Tanzania	EU720411	EU720573	EU720728	EU720925	EU721022	EU721165	EU721353	EU721531
Pappea	*capensis*	Eckl. & Zeyh.	*Edwards KE232*	JCT	South Africa	EU720424	EU720592	EU720747	EU720938	EU721035	EU721183	EU721371	EU721542
Paranephelium	*macrophyllum*	King	*Chase 1356*	K	Bogor, BG	EU720451	EU720619	EU720782	EU720955	EU721057	EU721218	EU721406	-

						GenBank accession number							
Genera	**species**	**Author**	**Voucher**	**Herbarium**	**Country**	**ITS**	**Genera**	**species**	**Author**	**Voucher**	**Herbarium**	**Country**	**ITS**
Paranephelium	*xestophyllum*	Miq.	*Edrwards KE503*	JCT	Asia	EU720420	EU720582	EU720737	-	EU721029	EU721174	EU721362	-
Paullinia	*pinnata*	L.	*Edwards KE199*	JCT	Tanzania	EU720413	EU720575	EU720730	EU720926	EU721023	EU721167	EU721355	-
Paullinia	*subauriculata*	Radlk.	*Weckerle 00/03/19-1/1*	Z	Peru	EU720494	-	EU720833	-	-	EU721266	EU721454	-
Plagioscyphus	aff. *louvelii*	Danguy & Choux	*Lowry 6034*	MO	Madagascar	EU720555	EU720701	EU720903	EU721008	EU721139	EU721325	EU721513	EU721599
Plagioscyphus	*unijugatus*	Capuron	*Buerki 145*	NEU	Madagascar	EU720475	EU720644	EU720809	EU720972	EU721081	EU721245	EU721433	EU721574
Podonephelium	*homei*	Radlk.	*Pillon 156*	MO	New Caledonia	EU720489	EU720650	EU720826	EU720976	EU721087	EU721259	EU721447	-
Pometia	*pinnata*	J.R. Forst. & G. Forst.	*Chase 2135*	K	Bogor, BG	EU720471	EU720638	EU720803	EU720967	EU721075	EU721239	EU721427	EU721572
Pometia	*pinnata*	J.R. Forst. & G. Forst.	*Yuan s.n.*	NEU	China	EU720402	-	EU720717	EU720916	-	EU721154	EU721343	EU721524
Pseudima	sp.		*McPherson 15867*	MO	Panama	EU720556	EU720702	EU720904	EU721009	EU721140	EU721326	EU721514	EU721600
Rhysotoechia	*mortoniana*	(F. Muell.) Radlk.	*Edwards KE117*	JCT	Australia	EU720414	EU720576	EU720731	EU720927	EU721024	EU721168	EU721356	EU721533
Sapindus	*oligophyllus (=Aphania oligophylla)*	Merr. & Chun	*Yuan s.n.*	NEU	China	EU720407	EU720568	EU720722	EU720921	EU721020	EU721159	EU721159	-
Sarcopteryx	*martyana*	(F. Muell.) Radlk.	*Irvine IRV1810*	CSIRO	Australia	EU720426	EU720594	EU720749	EU720940	EU721037	EU721185	EU721373	EU721544
Sarcopteryx	*reticulata*	S.T. Reynolds	*Gray BG1137*	CSIRO	Australia	EU720421	EU720587	EU720741	-	EU721033	EU721178	EU721366	EU721539
Sarcopteryx	sp.	-	*Edwards KE49*	JCT	Australia	EU720439	EU720607	EU720765	EU720948	EU721047	EU721201	EU721389	EU721554
Sarcotoechia	*serrata*	S.T. Reynolds	*Edwards KE31*	JCT	Australia	EU720436	EU720603	EU720761	EU720944	EU721043	EU721197	EU721385	EU721550
Sarcotoechia	*villosa*	S.T. Reynolds	*Edwards KE102*	JCT	Australia	EU720419	EU720581	EU720736	-	-	EU721173	EU721361	-
Schleichera	*oleosa*	(Lour.) Oken	*Chase 2137*	K	Bogor, BG	EU720423	EU720591	EU720746	EU720937	-	EU721182	EU721370	EU721541
Serjania	*altissima*	(Poepp.) Radlk.	*Weckerle 00/07/02-1/4*	Z	Peru	EU720498	-	EU720840	-	-	EU721269	EU721457	-
Serjania	*communis*	Cambess.	*Chase 2138*	K	Bogor, BG	EU720472	EU720640	EU720805	EU720969	EU721077	EU721241	EU721429	-
Serjania	*glabrata*	Kunth	*Merello 1058*	MO	Peru	EU720557	EU720703	EU720905	EU721010	EU721141	EU721327	EU721515	-
Storthocalyx	sp.		*Munzinger 960*	MO	New Caledonia	EU720504	EU720657	EU720848	-	EU721094	EU721276	EU721464	-
Synima	*macrophylla*	S.T. Reynolds	*Edwards KE19*	JCT	Australia	EU720430	EU720596	EU720753	EU720941	-	EU721189	EU721377	EU721546
Talisia	*angustifolia*	Radlk.	*Zardini 43668*	MO	Paraguay	EU720558	EU720705	EU720907	-	EU721143	EU721328	EU721516	-
Talisia	*nervosa*	Radlk.	*Pennington 628*	MO	-	EU720474	EU720643	EU720808	-	EU721080	EU721244	EU721432	-
Talisia	*obovata*	A.C. Sm.	*R.Lombello 13*	MO	Brazil	EU720485	EU720648	EU720822	-	EU721085	EU721255	EU721443	-

						GenBank accession number							
Genera	**species**	**Author**	**Voucher**	**Herbarium**	**Country**	**ITS**	**Genera**	**species**	**Author**	**Voucher**	**Herbarium**	**Country**	**ITS**
Thouinia	*acuminata*	S. Watson	*Liston 633-2*	MO	Mexico, Jalisco	EU720478	EU720647	EU720814	-	EU721084	EU721249	EU721437	-
Tina	*isaloensis*	Drake	*Ranirison PR827*	G	Madagascar	EU720520	EU720667	EU720864	EU720986	EU721104	EU721286	EU721474	EU721579
Tina	*striata*	Radlk.	*Vary 45*	MO	Madagascar	EU720509	EU720661	EU720853	-	EU721098	EU721279	EU721467	-
Tinopsis	*apiculata*	Radlk.	*Buerki 131*	NEU	Madagascar	EU720422	EU720589	EU720744	EU720936	EU721034	EU721180	EU721368	EU721540
Toechima	*erythrocarpum*	(F. Muell.) Radlk.	*Edwards KE20*	JCT	Australia	EU720431	EU720597	EU720754	-	EU721038	EU721190	EU721378	-
Toechima	*plurinerve*	Radlk.	*Chase 1357*	K	Bogor, BG	EU720452	EU720620	EU720783	EU720956	EU721058	EU721219	EU721407	EU721561
Toechima	*tenax*	(Cunn. ex Benth.) Radlk.	*Chase 2046*	K	Australian, BG	EU720456	EU720623	EU720787	-	EU721061	EU721223	EU721411	EU721564
Toechima	*tenax*	(Cunn. ex Benth.) Radlk.	*Chase 2132*	K	Bogor, BG	EU720469	EU720636	EU720801	EU720965	EU721073	EU721237	EU721425	EU721570
Tristiropsis	*acutangula*	Radlk.	*Chase 1358*	K	Bogor, BG	EU720453	EU720621	EU720784	EU720957	EU721059	EU721220	EU721408	EU721562
Urvillea	*ulmaceae*	Kunth	*Weckerle 00/07/05-1/1*	Z	Peru	EU720499	EU720655	EU720841	-	EU721092	EU721270	EU721458	-
Vouarana	*guianensis*	Aubl.	*Lucas 109*	MO	French Guiana	EU720525	EU720673	EU720870	EU720991	EU721110	EU721292	EU721480	EU721582
Xanthoceras	*sorbifolium*	Bunge	*Yuan CN2006*	NEU	China	EU720398	EU720562	EU720711	EU720910	EU721014	EU721148	EU721337	-
Outroup													
Sorindeia	sp.		*Buerki 137*	NEU	Madagascar	-	-	EU720831	-	-	EU721264	EU721452	-
Harrisonia	*abyssinica*	Oliv.	*Edwards KE510*	JCT	Tanzania	EU720440	-	EU720766	-	-	EU721202	EU721390	-

Figure 1

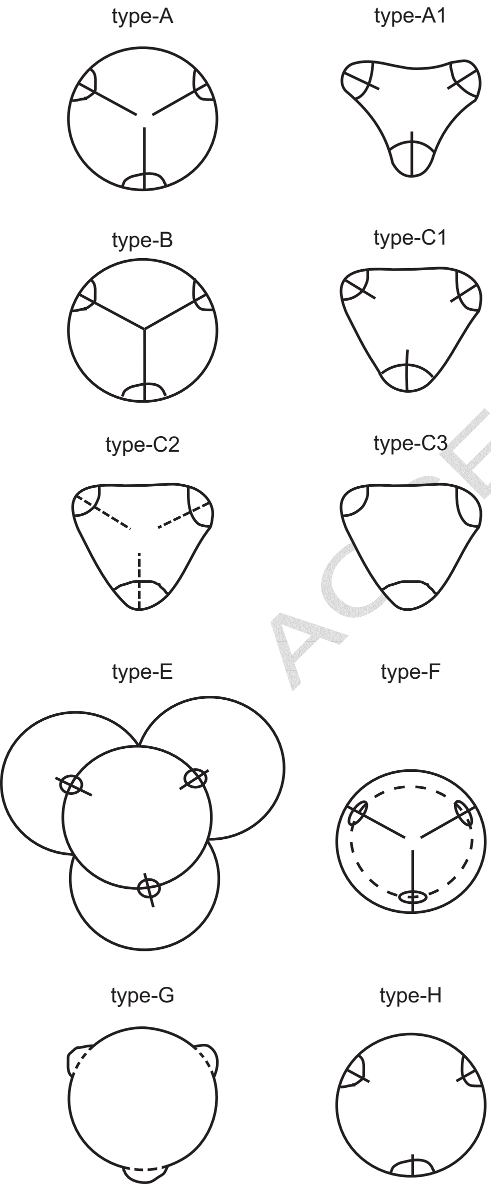

type-A
type-A1
type-B
type-C1
type-C2
type-C3
type-E
type-F
type-G
type-H

Figure 2

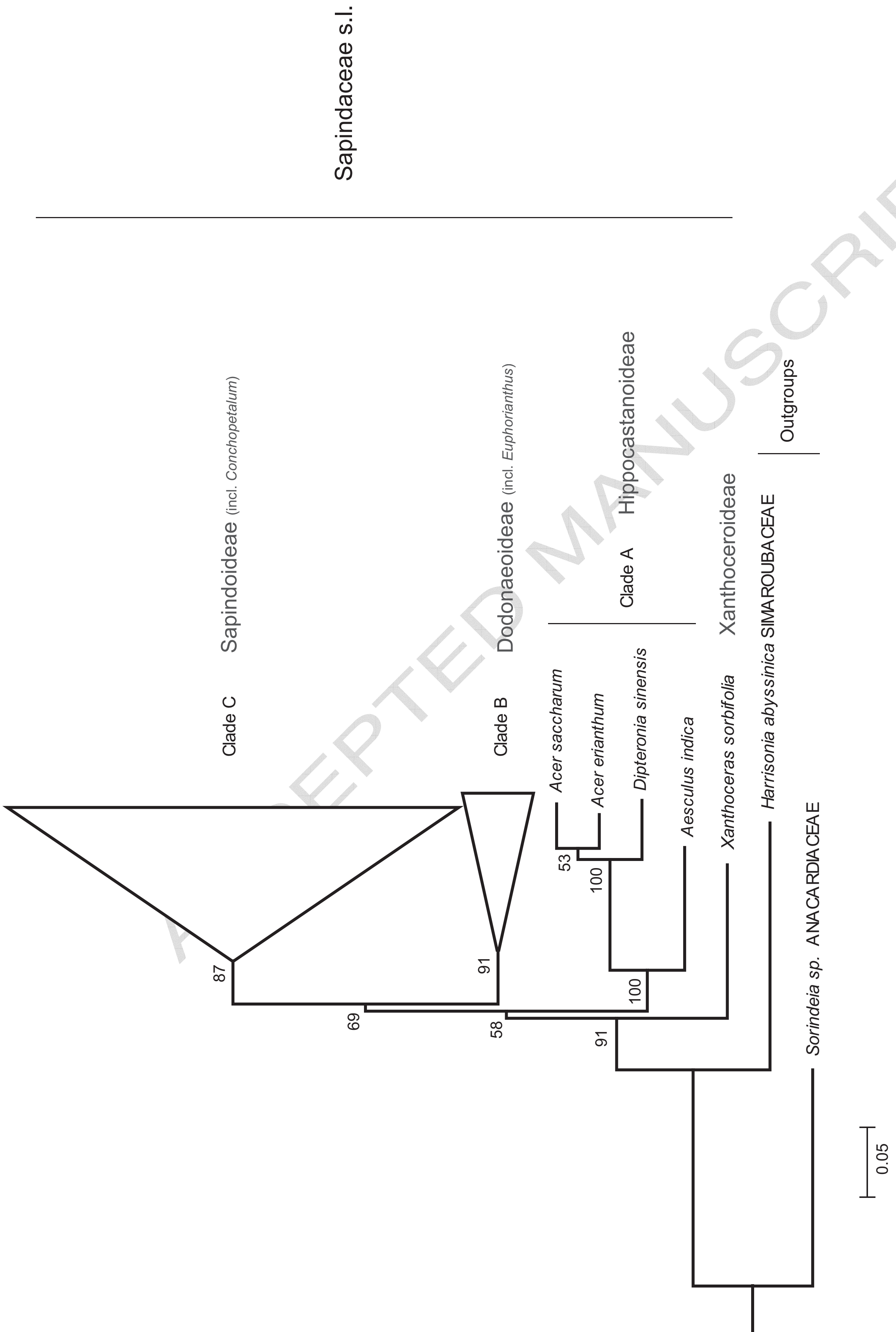

Sapindaceae s.l.
Clade C
Sapindoideae (incl. *Conchopetalum*)
Clade B
Dodonaeoideae (incl. *Euphorianthus*)
Clade A
Hippocastanoideae
Acer saccharum
Acer erianthum
Dipteronia sinensis
Aesculus indica
Xanthoceroideae
Xanthoceras sorbifolia
Outgroups
Harrisonia abyssinica SIMAROUBACEAE
Sorindeia sp. ANACARDIACEAE
87
91
69
58
53
100
100
91
0.05

Figure 3

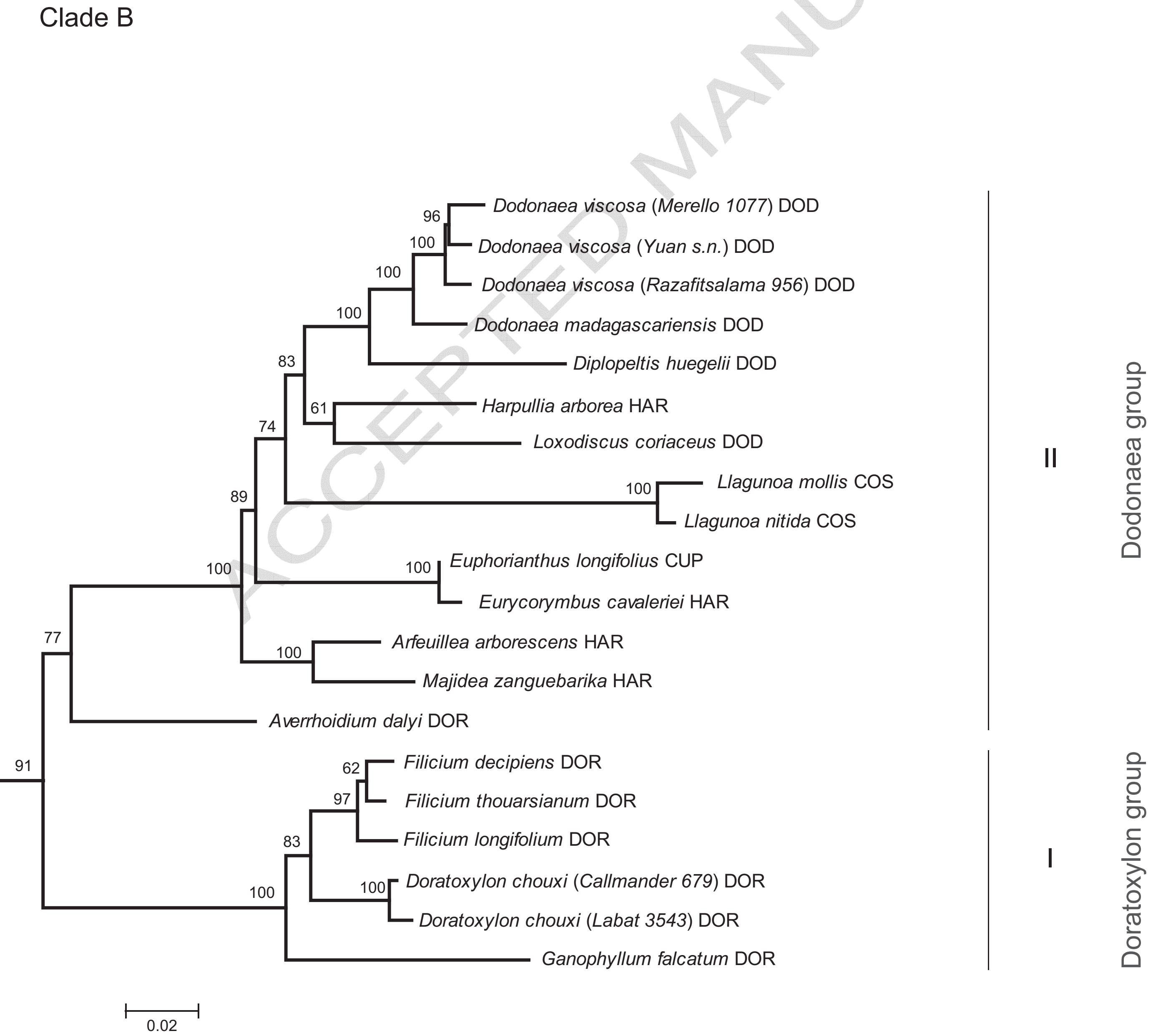

Clade B
96
100
100
100
83
61
74
89
100
77
91
100
100
100
62
97
83
100
100
Dodonaea viscosa (Merello 1077) DOD
Dodonaea viscosa (Yuan s.n.) DOD
Dodonaea viscosa (Razafitsalama 956) DOD
Dodonaea madagascariensis DOD
Diplopeltis huegelii DOD
Harpullia arborea HAR
Loxodiscus coriaceus DOD
Llagunoa mollis COS
Llagunoa nitida COS
Euphorianthus longifolius CUP
Eurycorymbus cavaleriei HAR
Arfeuillea arborescens HAR
Majidea zanguebarika HAR
Averrhoidium dalyi DOR
Filicium decipiens DOR
Filicium thouarsianum DOR
Filicium longifolium DOR
Doratoxylon chouxi (Callmander 679) DOR
Doratoxylon chouxi (Labat 3543) DOR
Ganophyllum falcatum DOR
II
Dodonaea group
I
Doratoxylon group
0.02

Figure 4

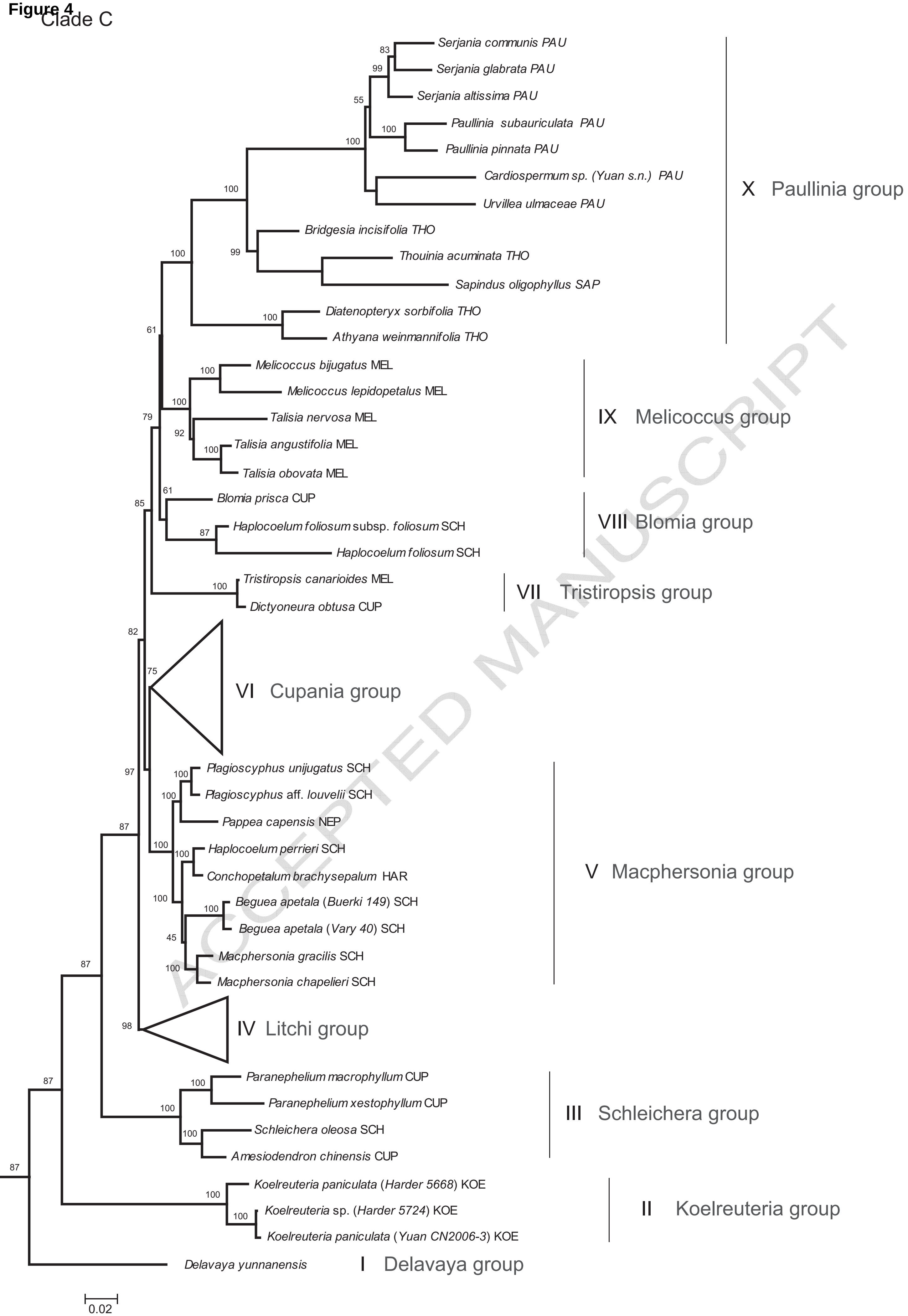
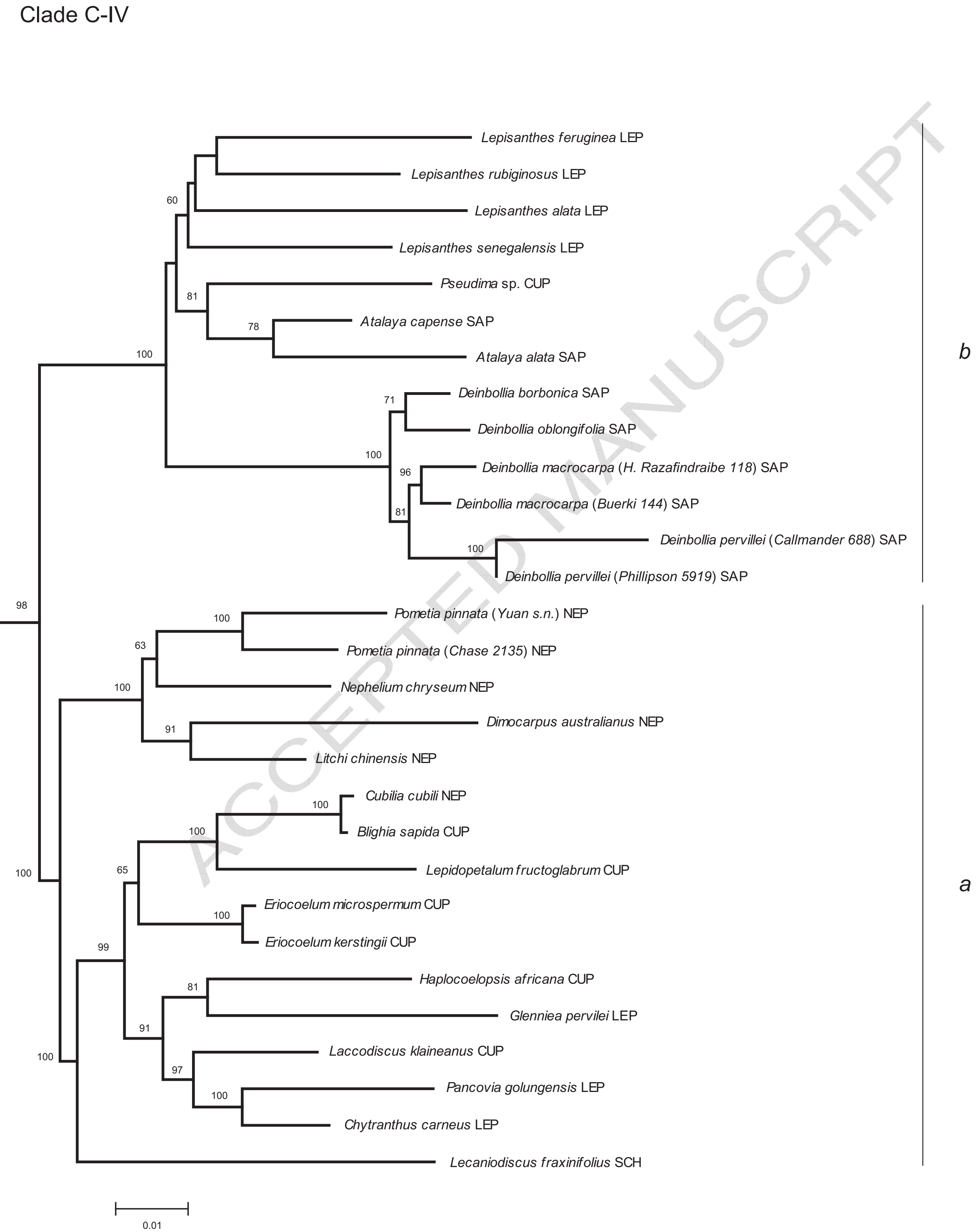

Clade C
Serjania communis PAU
Serjania glabrata PAU
Serjania altissima PAU
Paullinia subauriculata PAU
Paullinia pinnata PAU
Cardiospermum sp. (Yuan s.n.) PAU
Urvillea ulmaceae PAU
Bridgesia incisifolia THO
Thouinia acuminata THO
Sapindus oligophyllus SAP
Diatenopteryx sorbifolia THO
Athyana weinmannifolia THO
X Paullinia group
Melicoccus bijugatus MEL
Melicoccus lepidopetalus MEL
Talisia nervosa MEL
Talisia angustifolia MEL
Talisia obovata MEL
IX Melicoccus group
Blomia prisca CUP
Haplocoelum foliosum subsp. foliosum SCH
Haplocoelum foliosum SCH
VIII Blomia group
Tristiropsis canarioides MEL
Dictyoneura obtusa CUP
VII Tristiropsis group
VI Cupania group
Plagioscyphus unijugatus SCH
Plagioscyphus aff. louvelii SCH
Pappea capensis NEP
Haplocoelum perrieri SCH
Conchopetalum brachysepalum HAR
Beguea apetala (Buerki 149) SCH
Beguea apetala (Vary 40) SCH
Macphersonia gracilis SCH
Macphersonia chapelieri SCH
V Macphersonia group
IV Litchi group
Paranephelium macrophyllum CUP
Paranephelium xestophyllum CUP
Schleichera oleosa SCH
Amesiodendron chinensis CUP
III Schleichera group
Koelreuteria paniculata (Harder 5668) KOE
Koelreuteria sp. (Harder 5724) KOE
Koelreuteria paniculata (Yuan CN2006-3) KOE
II Koelreuteria group
Delavaya yunnanensis
I Delavaya group
83
99
55
100
100
100
99
100
100
61
100
100
79
92
100
61
85
87
100
82
75
97
100
100
87
100
100
100
100
45
100
87
98
100
100
87
100
100
87
100
0.02

Figure 5



Clade C-IV
Lepisanthes feruginea LEP
Lepisanthes rubiginosus LEP
Lepisanthes alata LEP
Lepisanthes senegalensis LEP
Pseudima sp. CUP
Atalaya capense SAP
Atalaya alata SAP
Deinbollia borbonica SAP
Deinbollia oblongifolia SAP
Deinbollia macrocarpa (H. Razafindraibe 118) SAP
Deinbollia macrocarpa (Buerki 144) SAP
Deinbollia pervillei (Callmander 688) SAP
Deinbollia pervillei (Phillipson 5919) SAP
Pometia pinnata (Yuan s.n.) NEP
Pometia pinnata (Chase 2135) NEP
Nephelium chryseum NEP
Dimocarpus australianus NEP
Litchi chinensis NEP
Cubilia cubili NEP
Blighia sapida CUP
Lepidopetalum fructoglabrum CUP
Eriocoelum microspermum CUP
Eriocoelum kerstingii CUP
Haplocoelopsis africana CUP
Glenniea pervilei LEP
Laccodiscus klaineanus CUP
Pancovia golungensis LEP
Chytranthus carneus LEP
Lecaniodiscus fraxinifolius SCH
60
81
78
100
71
100
96
81
100
98
100
63
100
91
100
100
65
100
99
81
91
97
100
100
100
b
a
0.01

Figure 6

Clade C-VI