Blackwell Publishing

Cladistics 21 (2005) 101–130

Cladistics

www.blackwell-synergy.com

Phylogeny, character evolution, and classification of Sapotaceae (Ericales)

Ulf Swenson* and Arne A. Anderberg

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

Accepted 1 February 2005

Abstract

We present the first cladistic study of the largely tropical family Sapotaceae based on both morphological and molecular data. The data were analyzed with standard parsimony and parsimony jackknife algorithms using equally and successive weighted characters. Sapotaceae are confirmed to constitute two main evolutionary lineages corresponding to the tribes Isonandreae-Mimusopeae-Sideroxyleae and Chrysophylleae-Omphalocarpeae. The Sideroxyleae are monophyletic, Isonandreae are polyphyletic as presently circumscribed, and as suggested by the analyses, the subtribe Mimusopeae-Mimusopinae has evolved within the Mimusopeae-Manilkarinae, which hence is also paraphyletic. Generic limits must be altered within Sideroxyleae with the current members *Argania*, *Nesoluma* and *Sideroxylon*. *Argania* cannot be maintained at a generic level unless a narrower generic concept is adopted for *Sideroxylon*. *Nesoluma* cannot be upheld in a narrow or broad generic concept of *Sideroxylon*. The large tribe Chrysophylleae circumscribes genera such as *Chrysophyllum*, *Pouteria*, *Synsepalum*, and *Xantolis*, but the tribe is monophyletic only if the taxa from Omphalocarpeae are also included. Neither *Chrysophyllum* nor *Pouteria* are monophyletic in their current definitions. The results indicate that the African taxa of *Pouteria* are monophyletic and distinguishable from the South American taxa. Resurrection of *Planchonella*, corresponding to *Pouteria* section *Oligotheca*, is proposed. The African genera *Synsepalum* and *Englerophytum* form a monophyletic group, but their generic limits are uncertain. Classification of the Asian genus *Xantolis* is particularly interesting. Morphology alone is indecisive regarding *Xantolis* relationships, the combined unweighted data of molecules and morphology indicates a sister position to Isonandreae-Mimusopeae-Sideroxyleae, whereas molecular data alone, as well as successive weighted combined data point to a sister position to Chrysophylleae-Omphalocarpeae. An amended subfamily classification is proposed corresponding to the monophyletic groups: Sarcospermatoideae (*Sarcosperma*), Sapotoideae (Isonandreae-Mimusopeae-Sideroxyleae) and Chrysophylloideae (Chrysophylleae-Omphalocarpeae), where Sapotoideae circumscribes the tribes Sapoteae and Sideroxyleae as well as two or three as yet unnamed lineages. Morphological characters are often highly homoplasious and unambiguous synapomorphies cannot be identified for subfamilies or tribes, which we believe are the reason for the variations seen between different classifications of Sapotaceae.

Sapotaceae are one member of the Ericales, a clade of morphologically variable angiosperm families where the relationships are not yet fully understood (Anderberg et al., 2002). The Sapotaceae are currently subdivided into five tribes with 53 genera and approximately 1250 species (Pennington, 1991; Govaerts et al., 2001). It consists of trees or shrubs with a world-wide distribution, although the highest species diversity is found in the tropical and subtropical regions of Asia and South America. Sticky and often white latex is found in cuts of bark, branches, leaves and fruits, although it often appears slowly in species growing in dry conditions. Leaves are generally alternate, simple, and entire, but exceptions of opposite leaves are present in *Leptostylis*, *Pichonia*, and *Sarcosperma*. Flower structure provides many useful characters for the circumscription of tribes and genera, and can be either simple or complex. Stamens are always opposite the corolla lobes, but many genera have staminodes in the corolla lobe sinuses.

---

*Corresponding author.
E-mail address: Ulf.Swenson@nrm.se

Another family characteristic is the malpighian hairs on different organs, a structure of a small stalk and two branches, often of different length.

Lam (1939), Aubréville (1964), Baehni (1965), and most recently Pennington (1991), proposed systems of classification based on morphological investigations that exhibited more or less contradictory results. Pennington accepted five tribes: Chrysophylleae, Isonandreae, Mimusopeae, Omphalocarpeae and Sideroxyleae, with Mimusopeae subdivided into three subtribes. Chrysophylleae are the largest tribe and traditionally recognized on 4–5-merous flowers with a calyx in a single whorl, imbricate or quincuncial sepals, and with corolla lobes, stamens and staminodes (if present) in the same number as sepals. Each corolla lobe is always entire. Sideroxyleae is characterized by features similar to those of Chrysophylleae, but differ in the usually divided corolla lobes. Mimusopeae and Isonandreae have a calyx in two whorls of 2–4 valvate sepals in each, though, never two in Mimusopeae or four in Isonandreae. Both tribes show exceptions to this pattern and rather have a calyx in a single whorl of 4–5 sepals. Corolla lobes, stamens and staminodes are usually of the same number as sepals in Mimusopeae whereas they are 2–3 times in Isonandreae, the latter consistently lacking staminodes. The two tribes also differ in having subdivided (Mimusopeae) or entire (Isonandreae) corolla lobes. In other words, both tribes have complex combinations of flower structures. The Omphalocarpeae are an assemblage of genera that are difficult to place. All members have several stamens opposite the corolla lobe, which is not unique to this tribe, but stamen position is not precise.

As a result of the often complex distribution of characters and character states, Pennington (1991) circumscribed several genera in a wide sense. He revised and reduced Aubréville's (1964) 122 accepted genera to 53. For instance, *Bumelia*, *Dipholis*, *Monotheca* and *Mastichodendron*, that had been recognized by earlier workers are now included in *Sideroxylon*. Another example is *Aningeria*, *Calocarpum*, *Lucuma*, *Malacantha* and *Planchonella*, which are all presently included in *Pouteria*. As a result, *Chrysophyllum* (81 spp.), *Pouteria* (304 spp.) and *Sideroxylon* (76 spp.) are now large groups that are difficult to recognize and lack apparent synapomorphies.

A phylogenetic hypothesis based on cpDNA sequence data from the plastid gene *ndh*F was presented by Anderberg and Swenson (2003) in which three main evolutionary lineages in the family were identified, *Sarcosperma* (referred to as Clade 1), Isonandreae, Mimusopeae, and Sideroxyleae (Clade 2), and Chrysophylleae-Omphalocarpeae (Clade 3). While providing a robust basis for future work, this molecular study did not provide sufficient data to resolve the generic relationships within each of the two larger clades, and did not shed light on the distribution of stated morphological features to diagnose genera and tribes. Adding morphological data to our previous molecular data set will help to better understand the diagnostic value of the morphological features on which earlier taxonomists relied to form their opinions of generic relationships.

In our previous paper (Anderberg and Swenson, 2003) we proposed that one problem with morphological characters was that character states combined in various ways, and that some states, such as a simple calyx, entire corolla lobes, and stamens equalling the sepals in number, may be symplesiomorphies, as they are present in *Sarcosperma*, the sister to all other Sapotaceae (Anderberg et al., 2002; Anderberg and Swenson, 2003). Other more complex features, such as stamens in one or two whorls, double calyx, and divided corolla lobes, were proposed to be synapomorphic. For example, the simple, often 5-merous flowers of *Chrysophyllum* could represent a symplesiomorphic type of flower, whereas the complex flowers of *Manilkara* could indicate a more derived relationship in the family.

The present analysis increases the sampling of taxa and combines the DNA sequence data from the *ndh*F gene with a newly developed morphological data set. The primary goal is to investigate the monophyly of the two main evolutionary lineages and the large genera of Sapotaceae. We also attempt to evaluate the diagnostic value of the more important and often used morphological characters.

Materials and methods

Taxon sampling

A total of 99 taxa were selected for this study (Table 1). They represent all the tribes and subtribes recognized by Pennington (1991), and we follow his classification and generic concepts if no alternative is stated. An initial aim was to erect a combined data set of morphology and *ndh*F sequences for all recognized genera, including a minimum of two species from non-monotypic genera. It proved difficult to obtain useful material from the following genera (species number in parentheses): *Aulandra* H.J. Lam (3), *Baillonella* Pierre (1), *Chromolucuma* Ducke (2), *Eberhardtia* Lecomte (3), *Gluema* Aubrév. & Pellegr. (1), *Isonandra* Wight (10), *Labourdonnaisia* Bojer (7), *Letestua* Lecomte (1), *Neohemsleya* T.D. Penn. (1), *Sarcaulus* Radlk. (5), *Tridesmostemon* Engl. (2), *Tsebona* Capuron (1), and *Vitellaria* C.F. Gaertn. (1). These genera were not included in the present analysis.

An outgroup external to Sapotaceae is difficult to choose, given the uncertainty regarding its sister group within Ericales (Anderberg et al., 2002; Bremer et al.,

Table 1
Taxa included in this phylogenetic study of Sapotaceae based on morphology and *ndh*F sequences data. Tribal and subtribal classification is from Pennington (1991) with author, voucher information, GenBank accession, and standard reference for the included taxa. Country of origin of cultivated species is given in parentheses. Taxonomy follows Govaerts et al. (2001). New sequences published here are marked with an asterix (*)

Taxon	Origin and voucher	*ndh*F GenBank acc. no.	Standard Reference
CHRYSOPHYLLEAE			
Aubregrinia taiensis (Aubrév. & Pellegr.) Heine	Ghana: *Enti 6871* (P)	AY230665	Pennington (1991)
Breviea sericea Aubrév. & Pellegr.	Cameroon: *Letouzey 8319* (P)	AY230667	Pennington (1991)
Capurodendron mandrarense Aubrév.	Madagascar: *Schatz & Miller 2477* (S)	AY230668	Aubréville (1974)
Chrysophyllum boivinianum (Pierre) Baehni	Madagascar: *McPherson 14426* (WAG)	AY230671	Aubréville (1974)
Chrysophyllum section *Aneuchrysophyllum*			
C. bangweolense R. E. Fr.	Zaire: *Malaisse 9600* (WAG)	AY230670	Hemsley (1968)
C. gonocarpum (Mart. & Eichler ex Miq.) Engl.	Paraguay: *Zardini & Guerrero 44475* (S)	AY230672	Pennington (1990)
C. perpulchrum Mildbr. ex Hutch. & Dalziel	Ghana: *Jongkind, Schmidt & Abbiw 1814* (MO)	AY230675	Hemsley (1968)
C. venezuelanense (Pierre) T. D. Penn.	Ecuador: *Ståhl, Knudsen & Lindström 5001* (S)	AY230678	Pennington (1990)
Chrysophyllum section *Donella*			
C. ogowense A. Chev.	Gabon: *McPherson 17910a* (S)	AY230673	Aubréville (1961)
C. pruniforme Engl.	Ghana: *Jongkind 3762* (WAG)	AY230676	Aubréville (1961)
C. roxburghii G. Don	Madagascar: *Solo & Randrianasolo 33* (WAG)	AY230677	Aubréville (1974)
Chrysophyllum section *Chrysophyllum*			
C. argenteum Jacq.	Suriname: *Miller & Hauk 9238* (S)	AY230669	Pennington (1990)
C. cainito L.	(Thailand): *Chantaranothai 2304* (Khon Kaen University Herb.)	*AY603774	Pennington (1990)
C. oliviforme L.	Cuba: *Gutiérrez & Nilsson 1* (S)	AY230674	Pennington (1990)
Delpydora gracilis A. Chev.	Ivory Coast: *Jongkind 5074* (WAG)	AY230679	Breteler and Nzabi (1995)
Delpydora macrophylla Pierre	Cameroon: *Wieringa & Haegens 2092* (WAG)	AY230680	Breteler and Nzabi (1995)
Ecclinusa guianensis Eyma	Brazil: Ducke Res. 05-906 (aliquot, Jodrell Laboratory, Kew)	*AY603776	Pennington (1990)
Ecclinusa ramiflora Mart.	Surinam: *Irwing* et al. *55081* (S)	AY230683	Pennington (1990)
Elaeoluma schomburgkiana (Miq.) Baill.	Brazil: *Keel & Coelho 243* (S)	AY230684	Pennington (1990)
Englerophytum magalismontanum (Sond.) T. D. Penn.	South Africa: *Devenish 1368* (S)	AY230685	Pennington (1991)
Englerophytum natalense (Sond.) T. D. Penn.	Tanzania: *Kayombo 3483* (S)	AY230686	Pennington (1991)
Leptostylis filipes Benth.	New Caledonia: *Webster & Hildreth 14665* (P)	AY230692	Aubréville (1967)
Leptostylis petiolata Vink	New Caledonia: *Mackee 12746* (P)	AY230693	Aubréville (1967)
Micropholis egensis (A. DC.) Pierre	Brazil: *Dionizia, Coêlho & Ernesto 73* (U)	AY230697	Pennington (1990)
Micropholis guyanensis (A. DC.) Pierre	Puerto Rico: *Taylor 11691* (MO)	AY230698	Pennington (1990)
Micropholis venulosa (Mart. & Eichler ex Miq.) Pierre	Brazil: *Assunção 122* (U)	AY230699	Pennington (1990)
Niemeyera francei (Guillaumin & Dubard) T. D. Penn.	New Caledonia: *Munzinger 965* (P)	*AY603778	Aubréville (1967)
Niemeyera whitei (Aubrév.) L. W. Jessup	Australia: *Floyd s.n.* (S)	AY230705	L. Jessup, pers. comm.
Pichonia novocaledonica (Engl.) T. D. Penn.	New Caledonia: *Veillon 377* (P)	AY230710	Aubréville (1967)
Pouteria section *Antholucuma*			
P. dominigensis (C. F. Gaertn.) Baehni	Cuba: *Gutiérrez & Nilsson 13* (S)	AY230718	Pennington (1990)
Pouteria section *Oligotheca*			
P. australis (R. Br.) Baehni	Australia: *Floyd s.n.* (S)	AY230713	van Royen (1957a)
P. baillonii (Zahlbr.) Baehni	New Caledonia: *Mackee 9914* (P)	AY230714	Aubréville (1967)
P. baueri (Montrouz.) Baehni	New Caledonia: *Munzinger 340* (P)	AY230715	Aubréville (1967)
P. cinerea (Pancher ex Baill.) Baehni	New Caledonia: *Veillon 7878* (P)	AY230717	Aubréville (1967)
P. eerwah (F. M. Bailey) Baehni	Australia: *Floyd s.n.* (S)	AY230719	van Royen (1957)
P. linggensis (Burck) Baehni	New Caledonia: Unknown collector (P 208707)	AY230721	Aubréville (1967)
P. myrsinifolia (F. Muell.) Jessup	Australia: *Floyd s.n.* (S)	AY230722	Royen (1957a)
P. obovata (R. Br.) Baehni	Taiwan: *Chung & Anderberg 1166* (HAST)	AF421071	Royen (1957a)
P. sandwicensis (A. Gray) Baehni & O. Deg.	Hawaii: *Koolan 119d* (GB)	AY230723	Royen (1957a)

Table 1
Continued

Taxon	Origin and voucher	*ndh*F GenBank acc. no.	Standard Reference
Pouteria section *Pouteria*			
P. gardneriana (A. DC.) Radlk.	Argentina: *Schworz 8216* (UPS)	*AY603780	Pennington (1990)
P. juruana K. Krause	Panama: *McPherson 15957* (MO)	AY230720	Pennington (1990)
P. lucuma (Ruiz & Pavón) Kuntze	Ecuador: Freire & Freire 815 (GB)	AY230725	Pennington (1990)
Pouteria section *Rivicoa*			
P. adolfi-friedericii (Engl.) A. Meeuse	Ethiopia: *Friis, Gilbert & Vollesen 3502* (UPS)	AY230711	Hemsley (1968)
P. alnifolia (Baker) Roberty	Ghana: *Jongkind & Noyes 1322* (MO)	AY230712	Hemsley (1968)
P. campechiana (Kunth) Baehni	(Taiwan): *Wang W00798* (HAST)	AY230716	Pennington (1990)
Pradosia brevipes (Pierre) T. D. Penn.	Brazil: *Lindeman 6743* (U)	AY230727	Pennington (1990)
Pradosia schomburgkiana (A. DC.) Cronquist	Brazil: Ducke Reserve 05– 1829 (aliquot, Jodrell Laboratory, Kew)	*AY603781	Pennington (1990)
Pradosia surinamensis (Eyma) T. D. Penn.	Guyana: *Harris 1076* (U)	AY230728	Pennington (1990)
Pycnandra sp.	New Caledonia: *McPherson & Munzinger 18106* (S)	*AY603782	Pennington (1990)
Synsepalum afzelii (Engl.) T. D. Penn.	Ivory Coast: *Jongkind 4770* (WAG)	AY230737	Pennington (1991)
Synsepalum brevipes (Baker) T. D. Penn.	Kenya: *Brunt 1543* (S)	AY230738	Hemsley (1968)
Synsepalum dulcificum (Schumach. & Thonn.) Daniell	Ghana: *Welsing, Merello & Schmidt 24* (WAG)	AY230739	Aubréville (1961)
Synsepalum fleuryanum A. Chev.	Gabon: *McPherson 17904* (S)	AY230740	Pennington (1991)
Synsepalum passargei (Engl.) T. D. Penn.	Tanzania: *Magogo 2452* (UPS)	AY230741	Hemsley (1968)
Trouettia sp.	New Caledonia: *Munzinger 696* (S)	*AY603784	Vink (1958)
Xantolis cambodiana (Pierre ex Dubard) P. Royen	Thailand: *Chantaranothai 2507* (Khon Kaen University Herb.)	*AY603787	van Royen (1957b)
Xantolis siamensis (Fletcher) P. Royen	Thailand: *Smitairi 1* (L)	AY230744	van Royen (1957b)
ISONANDREAE			
Burckella macropoda (K. Krause) H. J. Lam	(Indonesia, Bogor): *Chase 1359* (K)	*AY603773	Lam and Royen (1952a)
Diploknema butyracea (Roxb.) H. J. Lam	Nepal: *Polunin, Sykes & Williams 3975* (UPS)	AY230681	van Royen (1958)
Diploknema oligomera H. J. Lam	(Indonesia, Bogor): *Chase 1360* (K)	*AY603775	van Royen (1958)
Madhuca microphylla (Hook.) Alston	Sir Lanka: *Fagerlind 4790* (S)	AF421064	Pennington (1991)
Palaquium formosanum Hayata	Taiwan: *Chung & Anderberg 1421* (HAST)	AF421068	Yang (1998)
Payena acuminata (Blume) Pierre	(Indonesia, Bogor): *Chase 1368* (K)	*AY603779	van Bruggen (1958)
Payena lucida A. DC.	Borneo: *Ambri* et al. *AA1604* (L)	AY230709	van Bruggen (1958)
MIMUSOPEAE			
Glueminae			
Inhambanella henriquezii (Engl. & Warb.) Dubard	South Africa: *de Winter & Vahrmeijer 8536* (S)	AY230688	Pennington (1991)
Lecomtedoxa klaineana (Pierre ex Engl.) Pierre ex Dubard	(Holland): *Veldhuizen 1509* (WAG)	AY230691	Aubréville (1961)
Neolemonniera clitandrifolia (A. Chev.) Heine	Ghana: *Jongkind, Schmidt & Abbiw* 1777 (MO)	AY230703	Pennington (1991)
Manilkarinae			
Faucherea parvifolia Lecomte	Madagascar: *Birkinshaw* et al. 357 (P)	AY230687	Aubréville (1974)
Labramia costata (Hartog ex Baill.) Aubrév.	Madagascar: *Randriamanalinarivo 577* (UPS)	AY230689	Aubréville (1974)
Labramia mayottensis Labat, Pignal & Pascal	Comoro Islands, Mayotte: *Labat* et al. *3309* (P)	AY230690	Labat et al. (1997)
Manilkara hexandra (Roxb.) Dubard	Thailand: *Chantaranothai 2340* (Khon Kaen Univ. Herb.)	AY230694	Pennington (1991)
Manilkara kauki (L) Dubard	Thailand: *Chantaranothai 2341* (Khon Kaen Univ. Herb.)	AY230695	Pennington (1991)
Manilkara zapota (L) P. Royen	(Thailand): *Chantaranothai 2378* (Khon Kaen Univ. Herb.)	AY230696	Pennington (1990)
Northia seychellana Hook.f.	Seychelles: *Chong-Seng s. n.* (S)	AY230706	Pennington (1991)
Mimusopinae			
Autranella congolensis (De Wild.) A. Chev.	Congo: *Bokdam 4401* (WAG)	AY230666	Pennington (1991)
Mimusops comorensis Engl.	Comoro Islands: *Pignal & Ginguette 1065* (P)	AY230700	Aubréville (1974)
Mimusops elengi L.	Thailand: *Chantaranothai 2305* (Khon Kaen Univ. Herb.)	AY230701	Pennington (1991)
Mimusops zeyheri Sond.	South Africa: *Dahlstrand 6386* (GB)	AY230702	Pennington (1991)

Tieghemella heckelii (A. Chev.) Pierre ex Dubard	Ghana: *Jongkind 3936* (WAG)	AY230742	Pennington (1991)
Vitellariopsis cuneata (Engl.) Aubrév.	Tanzania: *Thomas 3662* (WAG)	AY230743	Hemsley (1968)
Vitellariopsis dispar (N. E. Br.) Aubrév.	South Africa: *Pentz 2* (P)	*AY603785	Pennington (1991)
Vitellariopsis kirkii (Baker) Dubard	Kenya: *Robertson 4085* (WAG)	*AY603786	Hemsley (1968)
OMPHALOCARPEAE			
Magodendron mennyae (C. T. White & W. D. Francis) Vink	New Guinea: *Takeuchi, Ama & Siga 16570* (S)	*AY603777	Vink (1995)
Omphalocarpum pachysteloides Mildbr. ex Hutch. & Dalziel	Ghana: *Jongkind 2351* (WAG)	AY230707	Pennington (1991)
Omphalocarpum strombocarpum Y. B. Harv. & J. C. Lovett	Tanzania: *Frimodt-Møller, Joker & Ndangalasi TZ538* (C)	AY230708	Harvey and Lovett (1999)
SIDEROXYLEAE			
Argania spinosa (L.) Skeels	Morocco: *Nordenstam 9325* (S)	AY230664	Pennington (1991)
Diploon cuspidatum (Hoehne) Cronquist	French Guiana: *Pennington* et al. *13843* (U)	AY230682	Pennington (1990)
Nesoluma polynesicum (Hillebr.) Baill.	Hawaii: *Degener 20770* (S)	AY230704	Pennington (1991)
Sarcosperma laurinum (Benth.) Hook. f.	Hong Kong: *Saunders s.n.* (S)	AF421080	Lam and Royen (1952b)
Sideroxylon betsimisarakum Lecomte	Madagascar: *Schönenberger* et al. *A-102* (UPS)	AY230729	Aubréville (1974)
Sideroxylon foetidissimum Jacq.	Hispaniola: *Lundin 638* (S)	AY230730	Pennington (1990)
Sideroxylon horridum (Griseb.) T. D. Penn.	Cuba: *Gutiérrez & Nilsson 5* (S)	AY230731	Pennington (1990)
Sideroxylon inerme L.	(Denmark, Aarhus): *Nielsen s.n.* (S)	AY230732	Pennington (1991)
Sideroxylon lanuginosum Michx.	USA: *Correll & Ogden 28456* (S)	AY230733	Pennington (1990)
Sideroxylon marmulano Banks ex Lowe	Canary Islands: *Swenson & Fernandez 581* (S)	*AY603783	Pennington (1991)
Sideroxylon mascatense (A. DC.) T. D. Penn.	Yemen: *Thulin, Beier & Hussein 9774* (UPS)	AF421066	Pennington (1991)
Sideroxylon reclinatum Michx.	USA: *Traverse 592* (GB)	AY230734	Pennington (1990)
Sideroxylon salicifolium (L.) Lam	Cuba: *Gutiérrez & Nilsson 14* (S)	AY230735	Pennington (1990)
Sideroxylon saxorum Lecomte	Madagascar: *Jongkind 3500* (WAG)	AY230736	Aubréville (1974)

2002). Sapotaceae are traditionally closely associated with Ebenaceae, Styracaceae and Symplocaceae (Pennington, 1991), but the above mentioned studies indicate possible affinities to Lecythidaceae or Maesaceae. Since one of the primary goals of this study was to investigate the monophyly and internal relationships among the genera using combined data of *ndh*F sequences and morphology, an outgroup that was not likely to cause large coding problems (homology, inapplicable states, character interpretation, etc.) was sought. For our analysis *Sarcosperma* was therefore selected as the outgroup, based on the results of Anderberg et al. (2002) and Anderberg and Swenson (2003). Both studies place *Sarcosperma* with maximum support (100%) as the sister group of all other Sapotaceae. *Sarcosperma* is most often included in Sapotaceae, but it differs from all other members of the family and it has been recognized in its own family, Sarcospermataceae (Lam, 1925).

Morphological data

Characters and character states were studied from herbarium material from the herbaria C, GB, L, MO, P, S, U, UPS and WAG. Characters were also checked against the literature, e.g., Aubréville (1967), Hemsley (1968), Pennington (1990, 1991), or other treatments (Table 1). Flowers and fruits were boiled in Copenhagen mixture (70 mL ethanol, 29 mL distilled water, 1 mL glycerol, two drops of methanol) in a microwave oven and/or soaked overnight for later examination under stereo and light microscope. A total of 78 morphological characters were assembled and presented in Appendix A. The matrix for the morphological data is presented in Appendix B.

Coding information into characters and character states is the most critical stage in a cladistic analysis of morphological data. Too often methodological explanations for choosing characters and character states are not discussed (Wiens, 2001). Discussions on general problems regarding missing information, quantitative data, independent characters, and ordered versus unordered characters are found in Stevens (1991), Wilkinson (1992, 1995), Maddison (1993), Slowinski (1993), Pleijel (1995), Wiens (1995), Wiens and Servedio (1997), Strong and Lipscomb (1999), and Simmons and Freudenstein (2002). Therefore, we include a discussion in Appendix A for some of the character codings. Characters of special importance, like characters for generic recognition, are mentioned in the discussion. A selection, given in italics, is also plotted on the jackknife tree (Figs 3 and 4).

Polymorphic characters are common in the Sapotaceae and can be treated in different ways. Kornet and Turner (1999) reviewed seven methods of coding polymorphic characters and finally recommended that they should be coded as plesiomorphic, in favor of the observed intraspecific variation, unless the ancestral state is unknown. Assessment of the ancestral state in Sapotaceae is generally missing. Since we agree with Wiens and Servedio (1997), that polymorphic characters do provide a phylogenetic signal, polymorphic characters are scored with the observed states.

Multistate characters are generally treated as unordered in phylogenetic reconstruction, but if there is a reason to believe a character state is a subset of another character state, such characters may be ordered (Wilkinson, 1992, 1995). In the present study, eight characters (30, 42, 51, 60, 61, 65, 73, and 78) were identified as having subsets of character states, and were consequently treated as ordered. Another feature of multistate characters is when discernible states from possible different characters form what often is called composite characters. Simmons and Freudenstein (2002) pointed out two related problems with composite character coding: (1) loss of hierarchic information if unordered states are used, and (2) the risk that linking separate characters could create synapomorphies that are not present with a reductive character coding. Reductive coding, such as present/absent characters, are often straightforward and simple, but may also create problems such as character redundancy, inapplicable states, and over-weighted characters (Pleijel, 1995; Wilkinson, 1995; Strong and Lipscomb, 1999). These issues became pertinent in coding leaves arrangement (chars. 1, 4), leaf pubescence (chars. 11–12), inflorescence (chars. 16–17), anther pubescence (chars. 49–50), and exserted or included anthers (char. 30, 42).

Molecular data

Most *ndh*F sequences used in this study were published by Anderberg and Swenson (2003). In the present paper, using the same molecular methods and primers, sequences from an additional 15 species were added to the original molecular data set (Table 1).

Phylogenetic analyses

The morphological data set in Appendix B, containing 99 taxa and 78 characters, was combined with the *ndh*F molecular data set and analyzed using PAUP ver. 4.0b10 (Swofford, 2002). An initial Heuristic search was performed using the settings: 1000 replicates, random step-wise addition, TBR branch swapping, collapse of branches if minimum length was zero, and steepest descent not in effect. The assumption of Fitch parsimony (Fitch, 1971) was used on all except eight morphological characters which we found reasonable to treat as ordered (see above). Taxa with variable character states were interpreted as polymorphic in PAUP. *A. priori* weighting of sequence data (especially the third position) is not advisable (Källersjö et al., 1999; Sennblad and Bremer, 2000), and not pursued here.

Subsequent to the search for the most parsimonious (MP) solutions, a jackknife analysis (Farris et al., 1996), as implemented in PAUP (Swofford, 2002), was undertaken using the Heuristic search option as mentioned above. However, to reduce the number of trees to swap and the computing time, the settings needed to be adjusted to 10 000 replicates, each with 100 random replicates, with the MULTREES option not in effect. Jackknifing investigates the structure in a matrix without permutation, but excludes an assigned fraction of characters, here set to 37% in order to resemble the proper Jac algorithm (Farris et al., 1996). Groups with support frequencies below 50% are not recognized. Support values of 50–69% are recognized as weak, 70–89% as moderate, and 90% or more as strong.

To evaluate characters with the strongest phylogenetic signal, a successive weighting analysis (Farris, 1969) was undertaken subsequent on the result of the search after the MP solutions, using the same Heuristic search settings. In weighting the characters we used the rescaled consistency index (RC) (Farris, 1989) and a base weight of 1000 to avoid fractions. The process was reiterated until the same tree length was obtained twice. This consensus was saved and a new jackknife analysis was performed on the character weights recovered from the successive weighting procedure.

MacClade 4.0 (Maddison and Maddison, 2000) was used to trace character transformation on the resulting trees. Some morphological characters are addressed in the following discussion. A selection of molecular and morphological characters (italics) are optimized in Figs 3 and 4.

Results

The morphological data matrix contains 78 characters, of which seven are uninformative. In total, there are 7722 data entries of which 117 (1.5%) are scored as unknown, 274 (3.5%) as inapplicable (lack of a structure), and 300 (3.9%) as polymorphic. The aligned *ndh*F sequences resulted in a matrix of 1980 base pairs (bp), of which 151 are informative. Thus, in total, there are 222 informative characters in the combined analysis.

The heuristic search for the most parsimonious (MP) solution recovered 19 396 trees (after 225 computing hours) with a length of 1683 steps (Fig. 1). These trees have a retention index (RI) of 0.686, a rescaled consistency index (RC) of 0.381, and a consistency index (CI) of 0.453 without uninformative characters. If polymorphic character states are instead interpreted as uncertainties, the trees are 1360 steps long, an indication of extensive homoplasy in the morphological data set. The resulting tree is much collapsed and considering the

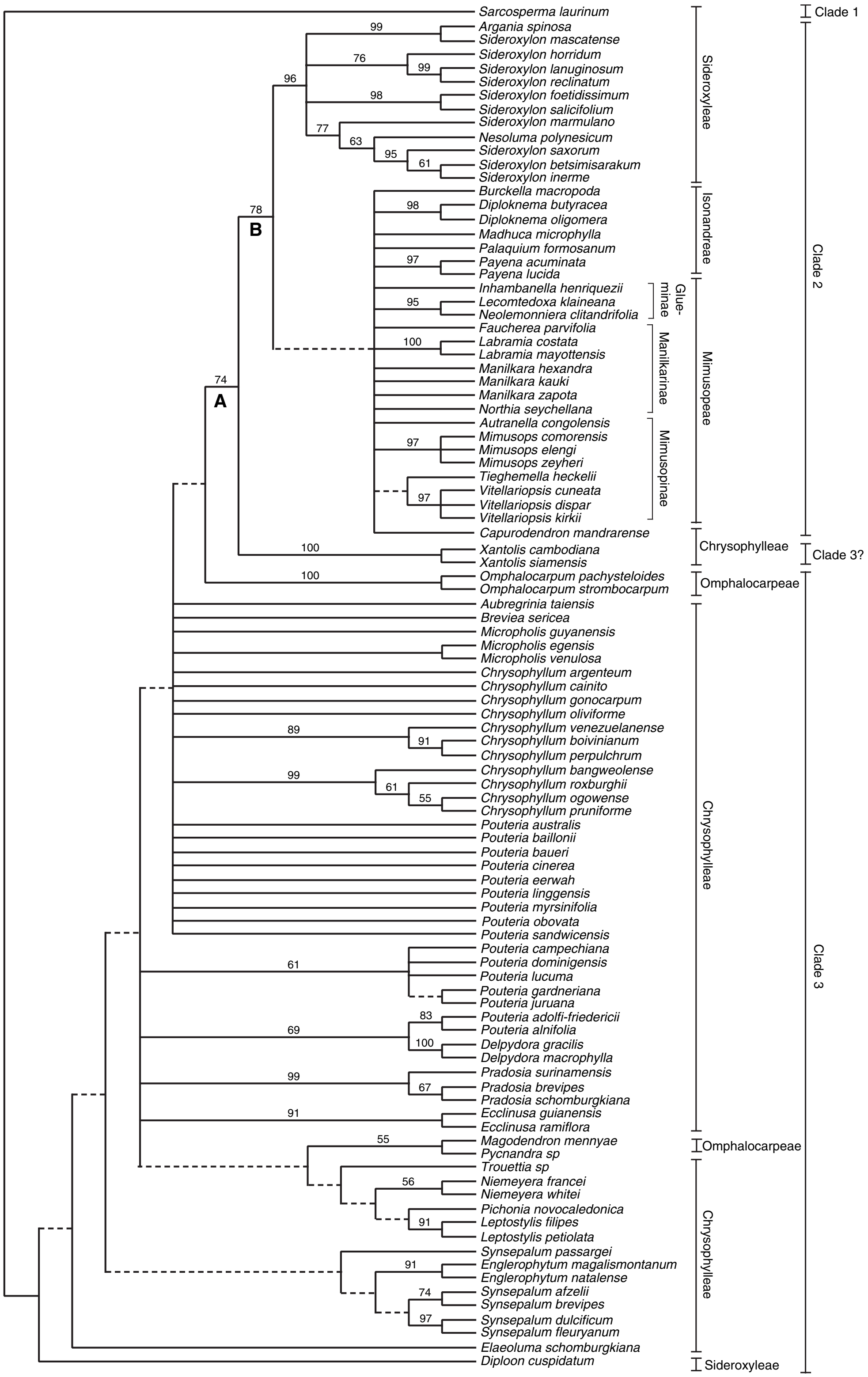


Fig. 1. Strict consensus of 19 396 MP trees of Sapotaceae from a parsimony analysis based on equal weights of a data set of morphology and cpDNA (*ndh*F) sequences. *Sarcosperma* was used as the outgroup. Branches with jackknife values of 50% or more are given above the branches. Dashed branches lack jackknife support. Bold letters indicate branches discussed in the text. The tribal and subtribal classification is from Pennington (1991). Clade numbers (far right) are from Anderberg and Swenson (2003).

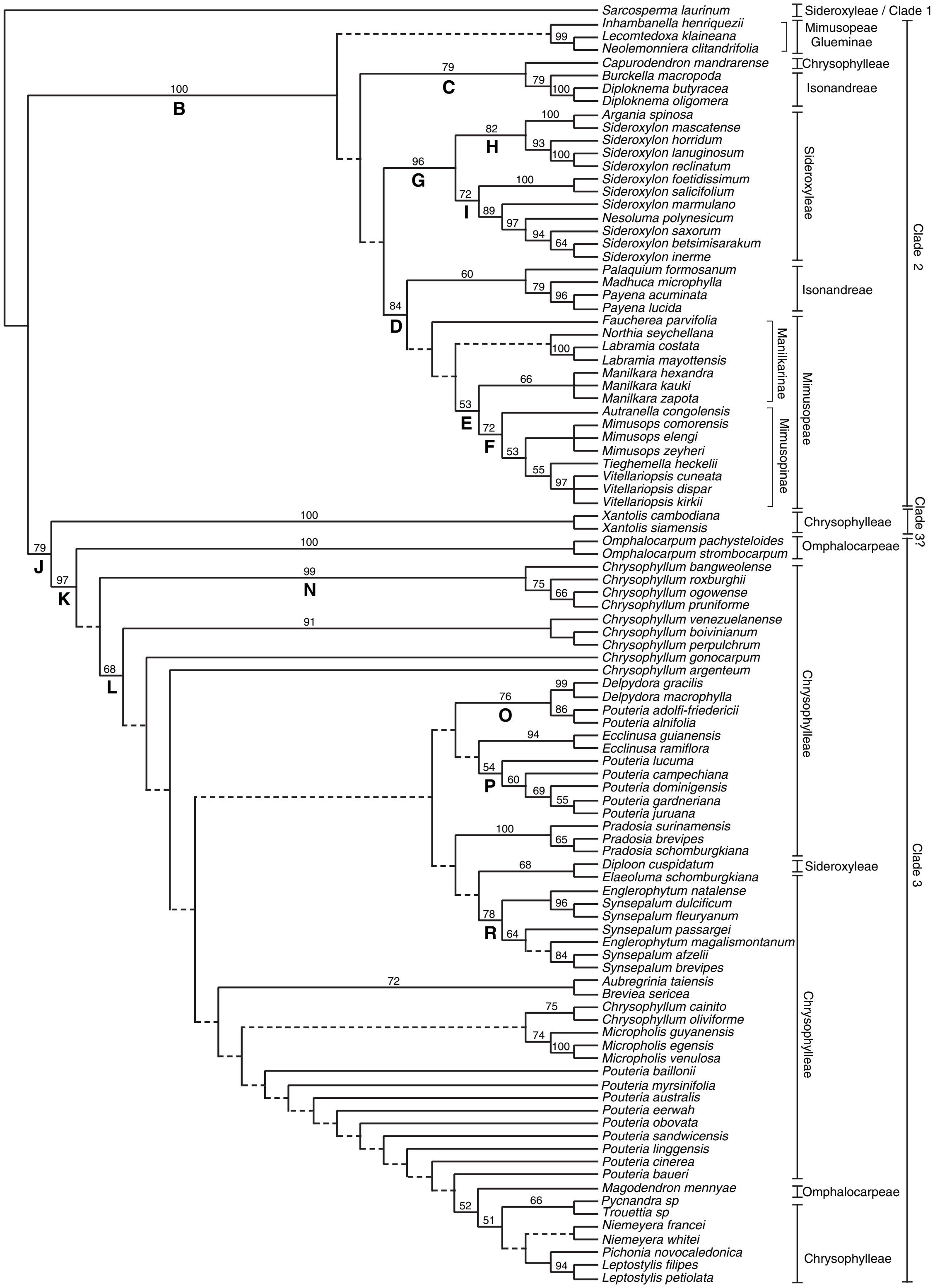


Fig. 2. Strict consensus of three MP trees of Sapotaceae from a parsimony successive weighting analysis of combined data of morphology and *ndh*F sequences. *Sarcosperma* was used as outgroup. Jackknife support values of 50% or more are given above the branches. Dashed branches have no jackknife support. Branches discussed in the text are marked with bold letters. Tribal and subtribal classification is from Pennington (1991). Clade numbers (far right) are from Anderberg and Swenson (2003).

support values, as estimated with the jackknife analysis and group frequencies of 50%, most relationships must be considered uncertain (Fig. 1). The analysis lends no support for the tribal circumscription *sensu* Pennington (1991) except for Sideroxyleae. This tribe includes *Argania*, *Nesoluma* and *Sideroxylon*. *Sideroxylon* is only monophyletic if *Argania* and *Nesoluma* are considered part of *Sideroxylon*. However, monophyly is recovered for several genera such as *Diploknema*, *Labramia*, *Payena* and *Vitellariopsis*. The jackknife analysis could not find support for a basal dichotomous pattern in the family. Moderate support (74%) is presently found for Clade A, placing the Asian genus *Xantolis* as sister to a large clade corresponding to Isonandreae-Mimusopeae-Sideroxyleae (Clade B), or Clade 2 *sensu* Anderberg and Swenson (2003). Hence, referring to the jackknife tree, Clade A is monophyletic, but with a large polytomy of taxa assigned to the tribes Isonandreae and Mimusopeae. Monophyly of Chrysophylleae, or Clade 3, is not supported and it is collapsed into a large polytomy. Genera such as *Delpydora*, *Ecclinusa*, *Leptostylis* and *Pradosia* are strongly supported, but *Chrysophyllum* and *Pouteria* have their members occurring in different clades.

Successive weighting of the characters recovered a stable result of three trees after four iterations with a score of 453804 steps (RI: 0.941, RC: 0.855, CI: 0.711). Searches for optimal trees found a much more resolved MP consensus, but several nodes are yet not supported in the jackknife analysis. Compared to the result based on equal weights, several internal nodes gain support (Fig. 2). Moreover, it is fully congruent with the analysis based on chloroplast DNA sequences (Anderberg and Swenson, 2003), and has improved support for several groups. Clade 2, here Clade B, is recovered with maximum support (100%). Sideroxyleae (Clade G) are again found with the same taxa and jackknife support (96%), but higher internal resolution. Within Clade B, one group (Clade D) corresponds to Mimusopeae and parts of Isonandreae, a group with moderate jackknife support (84%). The remaining representatives of Isonandreae form another moderately supported clade with uncertain affinity in Clade B (Clade C, 79%). Monophyly of the Mimusopeae subtribes, Manilkarinae and Mimusopinae, are not supported where the latter seem to be an ingroup of the former. In agreement with the results by Anderberg and Swenson (2003), in this analysis *Xantolis* is the sister group to the large clade of Chrysophylleae-Omphalocarpeae (Clade J). Support for this position has increased in this study from 66% to 79%, but it still differs from the present results from the analysis based on equal weights (see above). The next branch (Clade K) is strongly supported (97%) and includes all genera of Chrysophylleae and Omphalocarpeae. The included genera of the latter tribe, *Omphalocarpum* and *Magodendron*, do not form a monophyletic group. *Delpydora*, *Ecclinusa*, *Leptostylis* and *Pradosia* were monophyletic based on equal weights, a result corroborated by successive weights. However, *Chrysophyllum* and *Pouteria* are again non-monophyletic in their present circumscription.

Jackknife-supported intergeneric relationships in Chrysophylleae are few. The African genus *Delpydora* is consistently sister to two African *Pouteria* species (Clade O, 76%). *Diploon* has an affinity to *Elaeoluma* (68%), both of South American origin. *Aubregrinia* and *Breviea* are two monotypic genera from Africa and form a moderately supported clade (72%). A monophyletic species complex of *Englerophytum* and *Synsepalum* (Clade R, 78%) is also from Africa. The position of this complex is uncertain, because the weighted jackknife analysis lends a weak support for it as sister to a large polytomy of most other Chrysophylleae genera, whereas the MP solution nests within the very same clade (cf. Clade L, Figs 2 and 5). The jackknife analysis also weakly supports a group consisting of all sampled Chrysophylleae taxa (Clade Q, 51%) confined to Australia, New Caledonia and neighboring areas. This relationship, albeit weak, is corroborated by a phylogenetic study based on sequences of the nuclear ribosomal DNA (ITS) of the Australasian *Pouteria* complex (Bartish et al., 2005). Other intergeneric relationships in Clade K must be recognized as weak or uncertain.

Discussion

Selecting a tree topology

The MP analyses of equal and successive weighted characters differ in tree topology. In order to discuss monophyly, classification and character evolution, it is necessary to justify which of the two hypotheses we believe is the better supported phylogeny of Sapotaceae. More or less recent objections to successive weighting have been put forward in the literature (Swofford and Olsen, 1991; Swofford et al., 1996; Källersjö et al., 1999), but the critique has also been rebutted (Farris, 2001). Successive weighting is a technique invented and improved by Farris (1969, 1989). The basic idea is that characters with no or little homoplasy are more reliable for phylogenetic inference. In other words, characters which best contribute to a hierarchic pattern are given higher weights in an *a posteriori* procedure. In practice, as implemented in PAUP (Swofford, 2002), after the search for most parsimonious trees, the characters are reweighted based on the rescaled consistency index (RC), as preferred by Farris (1989). A new parsimony analysis is undertaken, and the procedure is then reiterated until the same result appears twice.

Much of the objection to successive weighting rests in the debate of silent substitutions, saturation in protein coding genes, and the different *a priori* weighting schemes of nucleotide sites. This applies especially to the third position, which has been suggested to be excluded or downweighted in phylogenetic analyses (see Wenzel and Siddall, 1999; references therein). Källersjö et al. (1999) performed a jackknife analysis on a large molecular data set representing cyanobacteria as well as angiosperms. They concluded that homoplasy increases (not decreases) phylogenetic structure and that the third position conveys much more phylogenetic information than was previously believed. We are convinced that their critique against successive weighting does not necessarily apply here for several reasons. First, our analysis deals with a single family (not deep rooted phylogenies), where saturation in the *ndh*F gene is unlikely. The problem here is opposite: one of too little variation. Second, the critique applies, as far as we know, to molecular data (see Källersjö et al., 1999; Farris, 2001), not necessarily to morphology. Third, the weighted analysis fully agrees with our previous results obtained from separately analyzed *ndh*F sequence data (Anderberg and Swenson, 2003). Fourth, phylogenetic reconstruction in Chrysophylleae, using equal weights of ITS sequence (Bartish et al., 2005), also corroborate the successive weighting results.

There are two different possible outcomes of a successive weighting analysis: a subset of the initial MP trees (same tree length) versus a different tree topology (suboptimal tree length). The first situation can be referred to as a method for choosing among multiple equally parsimonious trees and the second may occur when multiple characters have very low consistency indices (Carpenter, 1988). Our combined data set of morphology and *ndh*F sequences is an example of the latter. Counting informative characters, 14% and 63%, respectively, of the molecular and morphological characters have a consistency index lower than 0.5 in the equally weighted analysis. This certainly indicates a high ratio of morphological homoplasy, which has also caused trouble for earlier taxonomists in agreeing on a stable classification (cf. Lam, 1939; Aubréville, 1964; Baehni, 1965; Pennington, 1991).

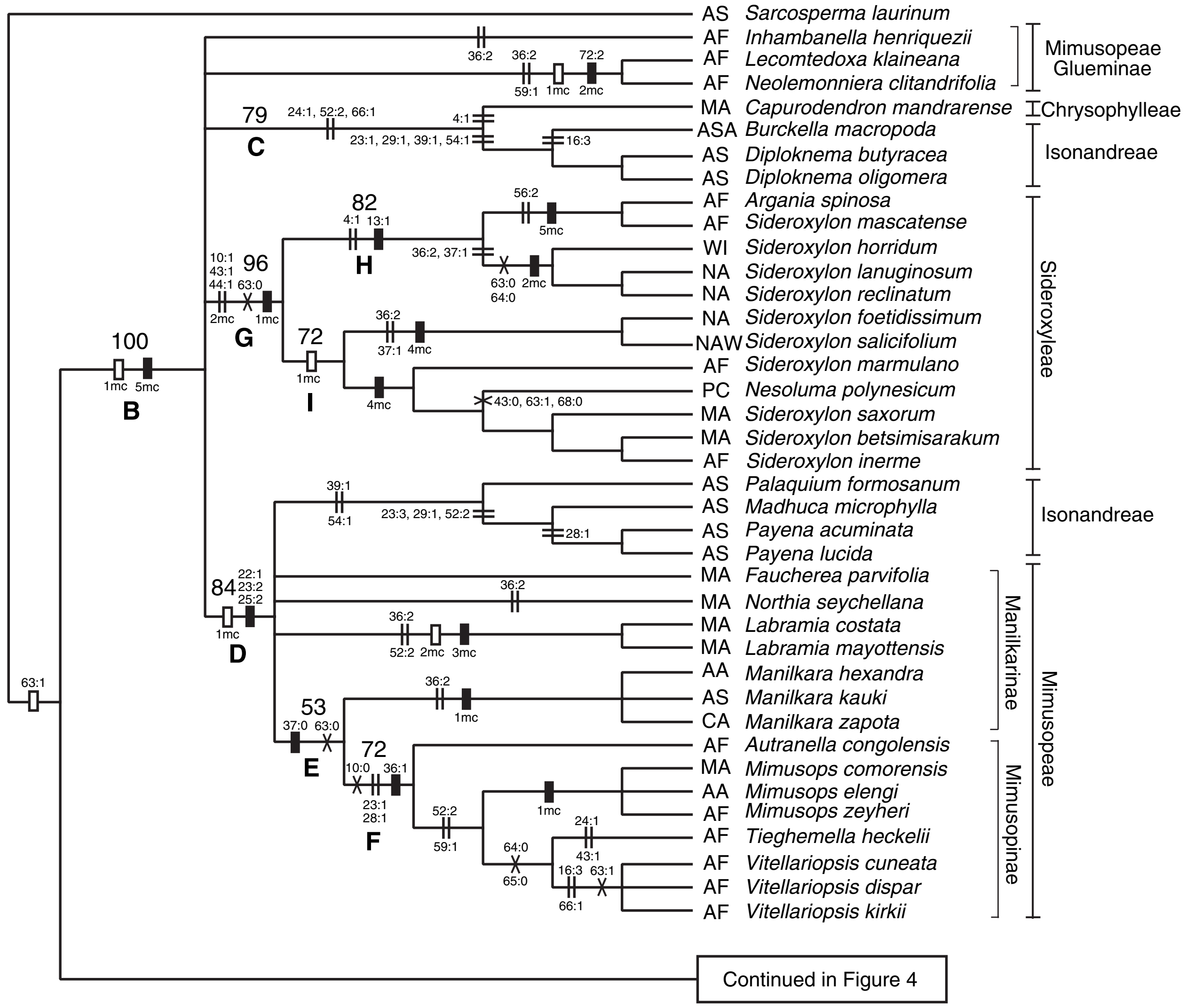
Returning to the consensus trees from the two analyses. Provided characters are assigned the same weight, the trees are 1683 (Fig. 1) and 1703 (Fig. 2) steps long. Hence, the successive weighting procedure leads to a suboptimal, less parsimonious phylogenetic reconstruction for the observed data, but Farris (1983) meant that all characters do not necessarily deserve the same weight or provide equally strong evidence for phylogenetic inference, a statement we agree with. If all branches without support in Figs 1 and 2 are collapsed, two similar "jackknife" trees are gained, where only one supported relationship in the equally weighted analysis disagrees with the successive weighting analysis. This is the position of the genus *Xantolis* (cf. Fig. 1, Clade A and Fig. 2, Clade J). Jackknife support for the unweighted relationship *Xantolis* being sister to Isonandreae-Mimusopeae-Sideroxyleae is 74%, versus a sister position to Chrysophylleae-Omphalocarpeae is 79% in the weighted analysis. Considering *Xantolis* possible sister groups, Clade B gain 78% jackknife in the unweighted analysis (Fig. 1), but in the weighted analysis the same clade is recovered with 100% and Clade K with 97% (Fig. 2). This shift in topology and increase in support for sister relationships are evidence that the weighted analysis eliminates the conflict caused by extensive homoplasy in the morphological data.

Based on these reasons we believe that the tree topology from the successive weighting analysis is at present the best hypothesis of Sapotaceae phylogeny and, is therefore used as the basis for the following discussion (Fig. 2). In addition, in order to discuss the evolution and optimization of some molecular and morphological characters, an MP tree is often selected. However, until a more robust and resolved phylogeny of the family is presented, we prefer to take a more conservative stand point and optimize the characters on the jackknife tree based on successive weights (Figs 3 and 4).

Xantolis and Diploon

Apart from *Sarcosperma*, which is designated as the outgroup, our analysis resolves Sapotaceae into two main lineages of evolution conforming to the tribes Isonandreae-Mimusopeae-Sideroxyleae and Chrysophylleae-Omphalocarpeae, respectively. This agrees well with the molecular phylogeny presented by Anderberg and Swenson (2003), where the two clades were referred to as Clade 2 and Clade 3, respectively. They also found *Xantolis* to possibly be a member of the large Clade 3, but its position was uncertain due to weak support. *Xantolis* is a genus sharing morphological similarities with several tribes and is sister to Chrysophylleae-Omphalocarpeae. This relationship gain in this study is supported slightly more strongly (79% jackknife), and is supported by a single molecular synapomorphy and several homoplasious morphological characters such as conspicuous intersecondary leaf veins (char. 7, ci: 0.095), lack of stipules (*char. 10:1*, ci: 0.125), foliaceous cotyledons (*char. 64:2*, ci: 0.273), and an embryo with copious endosperm (*char. 65:2*, ci: 0.174). In other words, no unambiguous morphological characters diagnose Clade J.

The monophyly of *Xantolis* is supported by three molecular synapomorphies. Furthermore, the flowers have five sepals, five corolla lobes, five stamens, and five staminodes, thus resembling the alleged plesiomorphic condition seen in *Sarcosperma*. *Xantolis* was described in



Continued in Figure 4

Fig. 3. Some selected morphological and molecular (mc) characters optimized on the branches of the jackknife tree of Clade B, the Isonandreae-Mimusopeae-Sideroxyleae (Sapotaceae) using combined data of morphology and *ndh*F sequences and successive weighting. Boxes represent synapomorphies without (filled) or with (open) homoplasy, crosses are reversals, and parallel lines are parallelisms. Letters in bold and jackknife in larger font refer to branches discussed in the text. Tribal and subtribal classification is from Pennington (1991). Distribution areas are given between tree topology and terminals: AF = Africa, AS = Asia, ASA = Asia to Australia, CA = Central America (including Mexico), MA = Madagascar (including Mayotte Island, Seychelles), NA = North America, NAW = North America and West Indies, PC = Pacific, WI = West Indies.

1838 but since then it has been merged with *Sideroxylon* or *Planchonella* (= *Pouteria*). van Royen (1957b) revised *Xantolis* and considered it close to, but distinguished from, *Planchonella* by its short corolla tube, sagittate (calcarate) anthers, and a connective extending above the theca (appendage, *char. 52*). Royen also mentioned a hairy corolla throat, a character we find inconsistent in the genus. Pennington (1991) reduced *Planchonella* to a synonym of *Pouteria*, but hesitated whether to place *Xantolis* in Isonandreae, Sideroxyleae or Chrysophylleae. Characters supporting a placement in Isonandreae, perhaps near *Palaquium*, include a short corolla tube with spreading corolla lobes, and a long exserted style. On the other hand, *Xantolis* differs by having staminodes and a uniseriate calyx, characters shared with many members of Chrysophylleae like *Micropholis*, *Pouteria* and *Synsepalum*. Pennington also pointed to the similar 5-merous flower of *Sideroxylon*, a type of flower indeed predominant in Chrysophylleae. He also suggested that the similarities with Isonandreae may not reflect a common ancestry but rather are based on non-homologous or homoplasious characters. Our previous analysis of Sapotaceae supported Pennington's conclusion, placing *Xantolis* as the sister taxon of Chrysophylleae-Omphalocarpeae. However, this position received weak support from the *ndh*F gene alone, so a close relationship to Chrysophylleae was considered tentative. Character optimization on the tree obtained in this analysis reveals affinities to both groups. Calcarate anthers are found in many members of Clade J, including *Omphalocarpum, Planchonella* and *Pouteria*. A short corolla tube, on the other hand, is more common in Clade B, but is also found in several members of Clade J such as *Ecclinusa*, *Pichonia* and *Synsepalum*. However, useful characters to distinguish *Xantolis* from all other members of Chrysophylleae-Omphalocarpeae include acute anther appendages, lanceolate calyx and corolla lobes, and fimbriate and aristate staminodes. Since most of those characters were not used in this analysis, it is not known if they are actually synapomorphies for *Xantolis*. Otherwise, *Xantolis* has many plesiomorphic character states that occur throughout the family.

Diploon is a monotypic genus and has been associated with either *Chrysophyllum* and *Pradosia* (Cronquist, 1946; Aubréville, 1964), or with *Sideroxylon* (Pennington, 1991). A short corolla tube, widely spreading corolla lobes, and exserted stamens, were put forward by Pennington as similarities to *Sideroxylon*, but he also mentioned the absence of staminodes and its flavonoid chemistry as being more similar to *Chrysophyllum*. The chemical profile of *Diploon* includes high concentrations of myricetin and gallic acid and the absence of quercetin, a common substance present in most investigated species of *Chrysophyllum, Pouteria and Pradosia* (Waterman and Mahmoud, 1991). *Diploon* has a similar chemical profile to *Synsepalum*, also a taxon of the Chrysophylleae. In accordance with our molecular study, successive weighting embedded *Diploon* with strong support in Chrysophylleae. It further suggests a possible affinity to the South American genus *Elaeoluma*. The morphological differences between the two genera are striking, and chemically *Elaeoluma* has high concentration of quercetin, which is absent in *Diploon* (Waterman and Mahmoud, 1991). In addition, *Diploon* possesses several traits that are rare in Chrysophylleae-Omphalocarpeae, such as a glabrous ovary with 1–2 loculi and a basi-ventral seed scar. These features are present in the basal *Sarcosperma*. Many of the characters that Chrysophylleae and *Diploon* do share, are plesiomorphic, since they are shared with *Sarcosperma*. A position within Clade K can be considered likely, but a close relationship to *Elaeoluma* is less likely.

Isonandreae-Mimusopeae-Sideroxyleae (Clade B)

This clade corresponds to one of the two major evolutionary lineages in Sapotaceae that were identified by Anderberg and Swenson (2003), and comprises, with few exceptions, all taxa of the previously defined tribes Isonandreae, Mimusopeae and Sideroxyleae (Clade B, Figs 2 and 3). The clade is supported by 100% jackknife support and conforms to Clade 2 of our earlier study. The lineage is diagnosed by five non-homoplasious and one homoplasious molecular synapomorphy, but no included unreversed morphological characters occur for the group. Of the three tribes, Sideroxyleae *s. str.* can be considered monophyletic, but Isonandreae and Mimusopeae seem to be paraphyletic based on the successively weighted combined data.

Isonandreae. Based on the tribal characters mentioned by Pennington (1991), Anderberg and Swenson (2003) suggested that the monophyly of Isonandreae could be supported by some floral characters such as two or three times as many corolla lobes as sepals (*char. 29*), and two or three times as many stamens as corolla lobes (*char 39 : 1*). Numerous corolla lobes are present in Isonandreae, but missing in *Palaquium*, *Aulandra* and *Isonandra* (the latter two are not included in the present study). Likewise, the presence of numerous stamens opposite each corolla lobe is a homoplasious feature present in *Pycnandra* and *Omphalocarpum* of Chrysophylleae-Omphalocarpeae (Clade J). Hence, neither of these characters unambiguously supports a group corresponding to the traditional circumscription of Isonandreae.

In our present analysis, Isonandreae were represented by the genera *Burckella*, *Diploknema*, *Madhuca*, *Palaquium* and *Payena*. Analysis of combined morphological and molecular data places the genera in two clades: one is part of the basal polytomy (Clade C, Figs 2 and 3), and the other is sister to (or part of) the tribe Mimusopeae (Clade D, Figs 2 and 3). Both groups have moderate support. Some of the morphological characters Pennington (1991) assigned to Isonandreae were a calyx of usually two whorls with two or three sepals in each (*chars 22:1, 23:2–3*), the outer sepals being valvate (*char. 25:2*), undivided corolla lobes (char. 36:0), and absence of staminodes (*char. 54:1*). The latter two characters do conform to taxa of this tribe, but the other three do not support Isonandreae in particular. For instance, undivided corolla lobes is a symplesiomorphy that is widely distributed in the family, present in genera such as *Sarcosperma*, *Sideroxylon* and the entire Clade J. In addition, a calyx of two whorls (*char. 22:1*) with valvate outer sepals (*char. 25:2*) divides the tribe in two groups because it is a synapomorphy for the clade comprising some Isonandreae, Manilkarinae and Mimusopinae (Clade D, Fig. 3), but is not found in the other parts of Isonandreae (Clade C) or the subtribe Glueminae of Mimusopeae. In Clade C, *Capurodendron* and *Diploknema,* but not *Burckella*, have an uniseriate and quincuncial calyx, a plesiomorphic condition, whereas *Palaquium, Madhuca* and *Payena* have a calyx of two series with valvate sepals. Both latter characters are found also in some African and South American *Pouteria* species, but are nevertheless homoplasious synapomorphies diagnosing the Isonandreae-Mimusopeae clade. Finally, the presence or absence of staminodes (*char. 54*) does not seem to be a reliable character for relationships across the family. Staminodes are plesiomorphic structures which are suggested from our analysis to have been lost at least twice in Clade B, and multiple times in Clade J, where they have also reappeared (see "Morphological homoplasy" below). Thus, the monophyly of Isonandreae is not supported and characters used to distinguish the tribe do not fit all members. The tribe is better split into two groups: one with the subtribes Manilkarinae and Mimusopinae, the other with uncertain affinity in Clade B.

Capurodendron is a genus of 23 species endemic to Madagascar, recognized by its uniformly 5-merous flowers with triangular, woolly hairy staminodes. Aubréville (1974) considered *Capurodendron* to be closely related to *Sideroxylon* (Sideroxyleae) or *Tsebona* (Omphalocarpeae). Pennington later considered it to be near *Synsepalum* in Chrysophylleae, a notion based on its overall similarity with two notable exceptions, hairy staminodes and contorted corolla lobes in *Capurodendron*, as opposed to glabrous staminodes and imbricate or valvate corolla lobes in *Synsepalum*. Our previous analysis based on *ndh*F sequence data showed that *Capurodendron* was part of the Isonandreae-Mimusopeae-Sideroxyleae lineage, not close to *Synsepalum* or other genera of Chrysophylleae-Omphalocarpeae. Our present analysis varifies this and further suggests, with moderate support, a sister relationship of *Capurodendron* with *Burckella* and *Diploknema* (Clade C, Figs 2 and 3). Morphological support for this position comes partly from fused sepals (*char. 24:1*) and lanceolate, appendaged anthers (*char. 52:2*).

Mimusopeae. Mimusopeae usually comprise the subtribes Glueminae, Manilkarinae and Mimusopinae which are represented here by three, four and four genera, respectively (Table 1). Diagnostic features separating the subtribes include the number of floral parts such as sepal whorls, sepals per whorl and stamens. Support for a monophyletic Mimusopeae was not recovered. Glueminae are not part of the alliance of Manilkarinae and Mimusopinae. The latter two were recovered as monophyletic in the MP solution but without sufficient jackknife support. Weak support (53%) was recovered for *Manilkara-Vitellariopsis* (Clade E, Figs 2 and 3), but this excludes *Faucherea*, *Labramia* and *Northia* of the Mimusopeae, which fall back into a polytomy (Clade D). As to the morphology, not a single unambiguous character diagnoses the tribe. Moreover, some characters form transitions between Mimusopeae and Isonandreae. For example, the sampled species of *Faucherea* has an intermediate number of ovary loculi (5–6), as well as the same type of tertiary veins as in *Palaquium* and *Payena*.

Glueminae are characterized by a single calyx whorl versus two whorls in Manilkarinae and Mimusopinae. In our present analysis, *Inhambanella*, *Lecomtedoxa* and *Neolemonniera* represent the subtribe. Material from *Eberhardtia* and *Gluema* did not yield useful DNA. Based on our previous *ndh*F sequence study, *Lecomtedoxa* and *Neolemonniera* formed a group with moderately strong support, but *Inhambanella* did not group with these taxa. The present analysis pulls all three genera together, but jackknife support for the position of *Inhambanella*, as sister to the other strongly supported group (99%), is absent (Figs 2 and 3). Corolla lobes subdivided into three segments (*char. 36:2*) unites all genera of Glueminae, but the character is also present in some taxa of Mimusopeae and Sideroxyleae. A single morphological character unites most genera of Glueminae (but not *Inhambanella*), i.e., a loculicidal capsule (*char. 72:2*) which is a unique fruit type in the family and a possible synapomorphy for the group. *Inhambanella* has the plesiomorphic condition, a one-seeded berry which is the most common fruit type in Sapotaceae. The only other taxon with a capsular-like fruit is *Omphalocarpum* of Chrysophylleae, but the fruits of *Omphalocarpum* are usually large, globose and many-seeded, and have a woody, indehiscent pericarp (Pennington, 1991). *Lecomtedoxa* and *Neolemonniera* also share three molecular synapomorphies, one of which is homoplasious. Hence, Glueminae is not part of Mimusopeae, but forms one or two clades with uncertain affinity in Clade B. Future phylogenetic studies will have to test the sister relationships and whether *Inhambanella* belong in this group of taxa.

Manilkarinae and Mimusopinae are traditionally recognized by multiples of three and four floral parts, respectively. Limits between the subtribes were not resolved in our previous study based only on *ndh*F data. All included taxa were then found in the large well-supported group (Clade 2), but internal resolution was low. The present analysis conforms well to the previous results, but resolves some parts of the clade better. Mimusopinae, Clade F (72% jackknife), is an integral part of Manilkarinae, which was not found to be monophyletic (Figs 2 and 3). Mimusopinae could be diagnosed by presence of stipules (*char. 10:0*), a calyx of four sepals in each series (*char. 23:1*), a corolla with two pseudoseries (*char. 28:1*), and lobes subdivided into two segments (*char. 36:1*). Traditionally diagnostic features shared by members of both subtribes include a calyx of two series (*char. 22:1*) and valvate outer sepals (*char. 25:2*). Neither character diagnoses the currently accepted subtribes. Instead, both characters optimize as parallelisms present also in Chrysophylleae. Within Clade B, however, they could be perceived as synapomorphies. Similarly, a calyx of three or four sepals is not a diagnostic feature, as was proposed by Pennington (1991). Three sepals (*char. 23:2*) is a synapomorphy for Clade D with a loss of one sepal in *Madhuca-Payena* (*char. 23:3*), and a gain in Clade F (*char. 23:1*). Loss and gains of sepals are also found in some species of South American *Pouteria*, and in *Leptostylis*.

We suggested earlier (Anderberg and Swenson, 2003) that additional support for relationships within Mimusopeae might come from overlooked traits such as leaf venation, stipules, and entire versus subdivided corolla lobes. Most genera in Clade B have a brochidodromous leaf venation, a character also common in Chrysophylleae-Omphalocarpeae. Caducous stipules is likely to be a symplesiomorphic feature occurring in *Sarcosperma* and in clades B and J with several losses. Subdivided corolla lobes (*char. 36*) are only found in Clade B, and can either appear with two or three corolla lobe segments. This character is partly correlated with what we here term a corolla with a pseudoseries, i.e., having two segments attached on the dorsal surface of the median segment (*char. 28:1, Payena*, Clade F). In addition, several characters are distributed in parts of the representatives from both subtribes Mimusopinae and Manilkarinae. For instance, segments of the corolla lobe being of unequal length (*char. 37*) and placentation (*char. 63*) are found in Clade E, not in the members *Faucherea*, *Labramia*, and *Northia* of Manilkarinae. Hence, we conclude that maintaining Mimusopinae would render Manilkarinae paraphyletic. Moreover, separating one part of Isonandreae from Mimusopeae cannot be defended, because the common node D is supported by both molecular and morphological characters.

In Mimusopeae, our study based on *ndh*F data found only support for the monophyly of *Labramia* and *Mimusops*, not *Manilkara*. Morphological data and additional sampling enhanced the support for *Labramia*, *Mimusops* and *Vitellariopsis*, and also gave weak support for a monophyletic *Manilkara*. The genus *Labramia* is endemic to Madagascar and its monophyly is supported by five molecular characters and a combination of a calyx in two series, each with three sepals, and a corolla lobe subdivided into three segments. These features are also all present in *Northia*, *Manilkara* and *Letestua* (not included). *Manilkara* and *Mimusops* are each supported by a single molecular synapomorphy, but lack unreversed morphological characters. Both genera are recognized by floral parts in multiples of three (*Manilkara*) or four (*Mimusops*), but in fact, representing various character combinations with other genera such as *Faucherea, Labramia*, *Tieghemella* or *Vitellariopsis*. It is possible that the present generic limits do not reflect monophyletic groups, a view supported by preliminary studies of the nuclear genome (ITS; T. Pennington, *in. litt.*). *Vitellariopsis* is a small genus of five African species and can be diagnosed by its inflorescence and subsessile anthers. According to Pennington (1991), the inflorescence is axillary, but we find the flowers clustered at the shoot apex in a way we refer to as pseudo-terminal (*char 16:3*), a type also found in *Burckella* (Isonandreae) and *Baillonella* (Mimusopeae, not included). We consider *Vitellariopsis* to be close to *Tieghemella* and *Mimusops*.

Sideroxyleae. Pennington (1991) included *Argania*, *Diploon*, *Neohemsleya*, *Nesoluma*, *Sarcosperma* and *Sideroxylon* in this tribe. *Sarcosperma* has been shown to be the sister taxon to all other Sapotaceae (Anderberg et al., 2002; Anderberg and Swenson, 2003), and has been excluded from this group. *Diploon* was shown earlier to be nested within Chrysophylleae, a result verified by the present study using successive weighting (see above). *Neohemsleya* was not included because the material for DNA analysis has thus far not been available. The monophyly of the remaining three genera as monophyletic gets strong jackknife support of 96% (Figs 2 and 3, Clade G). Three molecular characters support the modified Sideroxyleae, but it is difficult to characterize in morphological terms. Sideroxyleae *sensu* Pennington (1991) is recognized by having floral parts in multiples of five, calyx in a single whorl with imbricate or quincuncial sepals, corolla lobes frequently subdivided into three segments, staminodes often present, and a seed scar being basal or basi-ventral. Most of those characters are symplesiomorphies present in *Sarcosperma*, and provide Pennington's reason to include it in Sideroxyleae. However, most characters also occur in many other Sapotaceae of Chrysophylleae-Omphalocarpeae. Corolla lobes subdivided into two or three segments (*char. 36*) is restricted to Clade B, where the segments are either of equal size, or the lateral are shorter than the median (*char. 37*). Corolla segments of different lengths are found in two of the three clades of *Sideroxylon*, a trait not unique for the genus, also occurring in *Inhambanella* (Glueminae), *Labramia* and *Northia* (Manilkarinae). Therefore, the character is neither diagnostic of the Sideroxyleae tribe, nor of *Sideroxylon* s. lat., but may prove useful if a narrow generic concept once again should be adopted for *Sideroxylon*. Potentially, the most important character for tribal recognition is a basi-ventral placentation (*char. 63:0*). This is a plesiomorphic trait present in *Sarcosperma*, which thereafter transforms into an axil placentation but reverses in Sideroxyleae (and reverses back in *Nesoluma*). Dorsifixed anthers (*char. 43:1*) with a flexible connective (*char. 44:1*) may prove useful as a diagnostic feature for the tribe. These two characters give the anthers a horizontal position in the flower, exposing the pollen to the pollinator.

Generic concepts within Sideroxyleae have varied over the years. Pennington (1991) adopted a wide circumscription of *Sideroxylon* and reduced several genera accepted by his predecessors (Baehni, 1938, 1965; Aubréville, 1964). However, he maintained the monotypic North African genus *Argania* and the Pacific genus *Nesoluma*. Among the reduced genera are the New World *Bumelia*, *Dipholis* and *Mastichodendron*, and the African *Spiniluma* and *Monotheca*. The decision to reduce these into *Sideroxylon* was based on presumed close relationships and the complexity of diagnostic character combinations which he considered to be artificial. Our previous study included nine species of *Sideroxylon* representing *Bumelia* (*S. horridum*, *S. lanuginosum*, *S. reclinatum*), *Dipholis* (*S. salicifolium*), *Mastichodendron* (*S. foetidissimum*), *Monotheca* (*S. mascatense*) and *Sideroxylon* (*S. betsimisarakum*, *S. inerme*, *S. saxorum*). A ruminate endosperm is one of the salient characters of *Monotheca* and in order to test if this type of endosperm is a possible generic character, we added the Macaronesian species *Sideroxylon marmulano*, which also possess this feature.

Our analysis improved the resolution and support for generic concepts within Sideroxyleae. *Argania* and *Sideroxylon* (*Monotheca*) *mascatense* attach as sister to each other, defined by five molecular synapomorphies, and could be recognized by having subulate staminodes (*char. 56:2*) and exserted styles. Subulate staminodes are unique in Clade B, but frequent in Clade J. Exserted styles, however, are common in Sapotaceae, especially in Clade B, and not a phylogenetically strong character. *Argania* is diagnosed by its fused seeds, which form a single woody stone, but this is an autapomorphy that does not merit the recognition of *Argania* as a monotypic genus (Anderberg and Swenson, 2003). *Monotheca* was characterized by a ruminate endosperm, which also occurs in *Sideroxylon marmulano*. These two species are not closely related but belong in two completely different clades of Sideroxyleae and, hence, our results support Pennington's view that a ruminate endosperm does not define the genus.

Sister to *Argania* and *Sideroxylon mascatense* are three species of *Sideroxylon* representing the former genus *Bumelia* (*S. horridum*, *S. lanuginosum*, *S. reclinatum*), a relationship supported by leaves fascicled on brachyblasts (*char. 4:1*) and plants often armed with axillary spines (*char. 13:1*) (Clade H, Figs 2 and 3). *Bumelia* is supported by both molecular and morphological characters, but the latter show parallel evolution in other clades and are the cause of a long controversy. Two morphological traits, subdivided corolla lobes (*char. 36:2*) with lateral segments shorter than the median (*char. 37:1*), diagnose this clade, a situation which is also true for *S. foetidissimum* (vestigial lateral segments) and *S. salicifolium*. These characters have been used to recognize *Bumelia*, *Dipholis* and *Mastichodendron* (Cronquist, 1945; Aubréville, 1964). Pennington (1991) believed these features were artificial, and not reflective of the phylogeny. Our results suggest that both characters evolved in parallel in two separate New World clades and do not constitute a synapomorphy, thus corroborating Pennington's view. Nevertheless, if *Bumelia* is recognized under a more narrow generic concept, an embryo with plano-convex cotyledons (*char. 64:0*) could be useful.

A core of *Sideroxylon* species (*S. marmulano*-*S. inerme*) with *Nesoluma* included therein is supported by four molecular synapomorphies. However, not a single morphological character maps on this node, which makes *Sideroxylon* s. str. hard to delineate. An alternative generic circumscription would be to include *S. salicifolium* and *S. foetidissimum*, i.e., the once-recognized genera *Dipholis* and *Mastichodendron*, in *Sideroxylon* (Clade I, Figs 2 and 3). This is feasible on two plesiomorphic and one highly homoplasious characters, plus several widely distributed characters, including absence of thorns, leaves not fascicled on brachyblasts, and visible intersecondary veins. At this point, it is not known if these characters break down with a more ample sampling.

Nesoluma is diagnosed by the high but variable number of corolla lobes, stamens and staminodes, which were earlier hypothesized to represent autapomorphies (Anderberg and Swenson, 2003), and this is confirmed here, except that presence of staminodes is a plesiomorphy, which is also present in the outgroup. Staminodes in *Nesoluma* are often much reduced, often to only one or two examples. Since *Sideroxylon inerme*, the generic type, is a member of this clade, we suggest that *Nesoluma* is reduced to a synonym of *Sideroxylon*, independent of whether a broad or narrow generic concept will be adopted in the future.

Chrysophylleae-Omphalocarpeae (Clade J)

The second major evolutionary lineage found in the previous study (Anderberg and Swenson, 2003) was Chrysophylleae-Omphalocarpeae (Clade 3). Here, this result is confirmed with a jackknife support of 97% (Clade K, Figs 2 and 4). This large group is supported by five molecular synapomorphies, but morphology does not contribute much to this basal bifurcation, illustrating the difficulty in recognizing groups corresponding to tribes or subfamilies on morphological features. Including *Xantolis*, the least homoplasious characters are the absence of stipules (*char. 10:1*), foliaceous cotyledons (*char. 64:2*) and seeds with copious endosperm (*char. 65:2*), three characters which reverse within Clade J, but are also present in Isonandreae-Mimusopeae-Sideroxyleae. The most parsimonious solution suggests numerous relationships within the tribe, but taking a more conservative view, jackknife support values for many relationships are weak. However, it seems clear that the Omphalocarpeae are not monophyletic, but form subgroups within Chrysophylleae, and that several genera such as *Chrysophyllum*, *Englerophytum*, *Pouteria* and *Synsepalum* are not monophyletic. Other taxa such as *Ecclinusa*, *Leptostylis* and *Pradosia* seem less problematic and conform to current generic delimitations.

Omphalocarpeae. This former tribe included *Magodendron*, *Omphalocarpum*, *Tridesmostemon* and *Tsebona* (Pennington, 1991). *Omphalocarpum* is said to be close to *Tridesmostemon* (not included), and comprises 6–27 African species. This genus is currently under revision (Govaerts et al., 2001). The monophyly of *Omphalocarpum* is well supported by one molecular and several morphological characters such as flowers with several bracts at the base of the pedicel (*char. 20:1*), 3–6 stamens opposite each corolla lobe (*char. 39:1*), hairy and inflexed staminodes (*chars 58:1, 59:1*), and an indehiscent fruit with woody pericarp (*char. 72:3*) (Fig. 4). Both staminode characters are rare in Clade J, but are found in *Magodendron*, and are possibly the reason why *Magodendron* and *Omphalocarpum* were associated with one another by Aubréville (1964) and Pennington (1991). *Magodendron* is a genus with two species in New Guinea, characterized by an odd character, namely pluriloculate anther thecae and two stamens opposite each corolla lobe (Vink, 1995). *Pycnandra* of New Caledonia is the other member of Chrysophylleae-Omphalocarpeae with polymerous stamens. This character could contribute to such a relationship, but there is no indication of a close relationship in our analyses. Hence, Omphalocarpeae can be eliminated since their representatives do not form a monophyletic group and all members are embedded within Chrysophylleae-Omphalocarpeae (Figs 2 and 4). The affinity of the non-sampled members of Omphalocarpeae (*Tridesmostemon* and *Tsebona*) remain unknown.

Chrysophylleae. This large group of taxa, as mentioned above, corresponds to the second major evolutionary lineages in Sapotaceae, and circumscribes approximately 600 species. Except for *Diploon* and *Capurodendron*, the analysis verifies much of Pennington's (1991) view of the tribe. In general, it is recognized on its simple flowers with a calyx in a single whorl, 4–5 sepals, and usually with corolla lobes and stamens of the same number as sepals. Staminodes are sometimes present, as in *Micropholis* and *Pouteria*, but often fixed outside or above the anthers, a condition which is never found in Isonandreae-Mimusopeae-Sideroxyleae. Here, staminodes are fixed in a single whorl with the anthers. This may indicate that staminodes in the two major lineages are not homologous structures, but of different origin.

A polyphyletic Chrysophyllum. The generic delimitation of *Chrysophyllum* is a well known systematic problem in Sapotaceae. Aubréville (1964) presented a narrow generic concept and gave generic rank to a number of groups. Opposed to this view, Baehni (1965) had a very broad generic concept of *Chrysophyllum* and also included the South American genera *Ecclinusa* and *Pradosia*, as well as the African genus *Delpydora* in that genus. Pennington (1991) adopted an intermediate view, which was based on a broad suite of correlated characters, including the absence of stipules, five floral parts, a shortly tubular or cyathiform corolla, absence of staminodes, foliaceous cotyledons, endospermous seeds and an exserted radicle. He recognized six sections: *Aneuchrysophyllum* found in Africa, Madagascar and South America, *Donella* in Africa and Asia, the New World nominal section *Chrysophyllum*, and the strictly South American sections *Ragala*, *Prieurella* and *Villocuspis*.

Our previous study based on *ndh*F data could neither find conclusive evidence for nor against the monophyly of *Chrysophyllum*, but did indicate that the sections were not monophyletic. The present study supports this, with additional evidence that *Chrysophyllum* is polyphyletic in its present circumscription. Clade N (Figs 2 and 4) are all from Africa and is strongly supported by at least four molecular synapomorphies. It is positioned outside the core group of Chrysophylleae and represents the sections *Aneuchrysophyllum* (*C. bangweolense*) and *Donella* (*C. roxburghii*, *C. ogowense*, *C. pruniforme*). The other species of *Chrysophyllum* are found in several clades or as part of a large polytomy. One moderately supported clade is formed by *C. cainito* and *C. oliviforme*, members of section *Chrysophyllum*, and is diagnosed by more or less round seeds. Interestingly, both species are poor in secondary compounds and rich only in gallic acid, a rare chemical profile in *Chrysophyllum* (Waterman and Mahmoud, 1991). Another clade includes *C. boivinianum*, *C. perpulchrum* and *C. venezuelanense*, the latter two from the section *Aneuchrysophyllum*. Each clade finds support from a single molecular synapomorphy. One observation from the MP solution is that one of these clades is sister to a large portion of Chrysophylleae, where the other has an affinity to *Micropholis*. However, there is no jackknife support for these relationships.

The above diagnostic suite of characters for *Chrysophyllum* can, with a single exception, be rejected as diagnostic of the genus. All features, except the absence of staminodes, are shared with at least *Pouteria* section *Oligotheca*. This section is distributed in Australasia, formally known as *Planchonella*, and is particularly similar to *Chrysophyllum* in all seed characters. Seed characters and the absence of staminodes are, however, useful to distinguish *Chrysophyllum* from New World *Pouteria* where both genera occur. Likewise, the alleged 5-merous flowers are a plesiomorphic character state almost ubiquitous in the Chrysophylleae (exception *Leptostylis*) and not at all unique to *Chrysophyllum*. Thus, the monophyly of *Chrysophyllum* in its current circumscription is not supported. Future phylogenetic studies focused on the phylogeny and character evolution of *Chrysophyllum*, and allied genera will help to indicate if a more narrow generic concept is to be preferred.

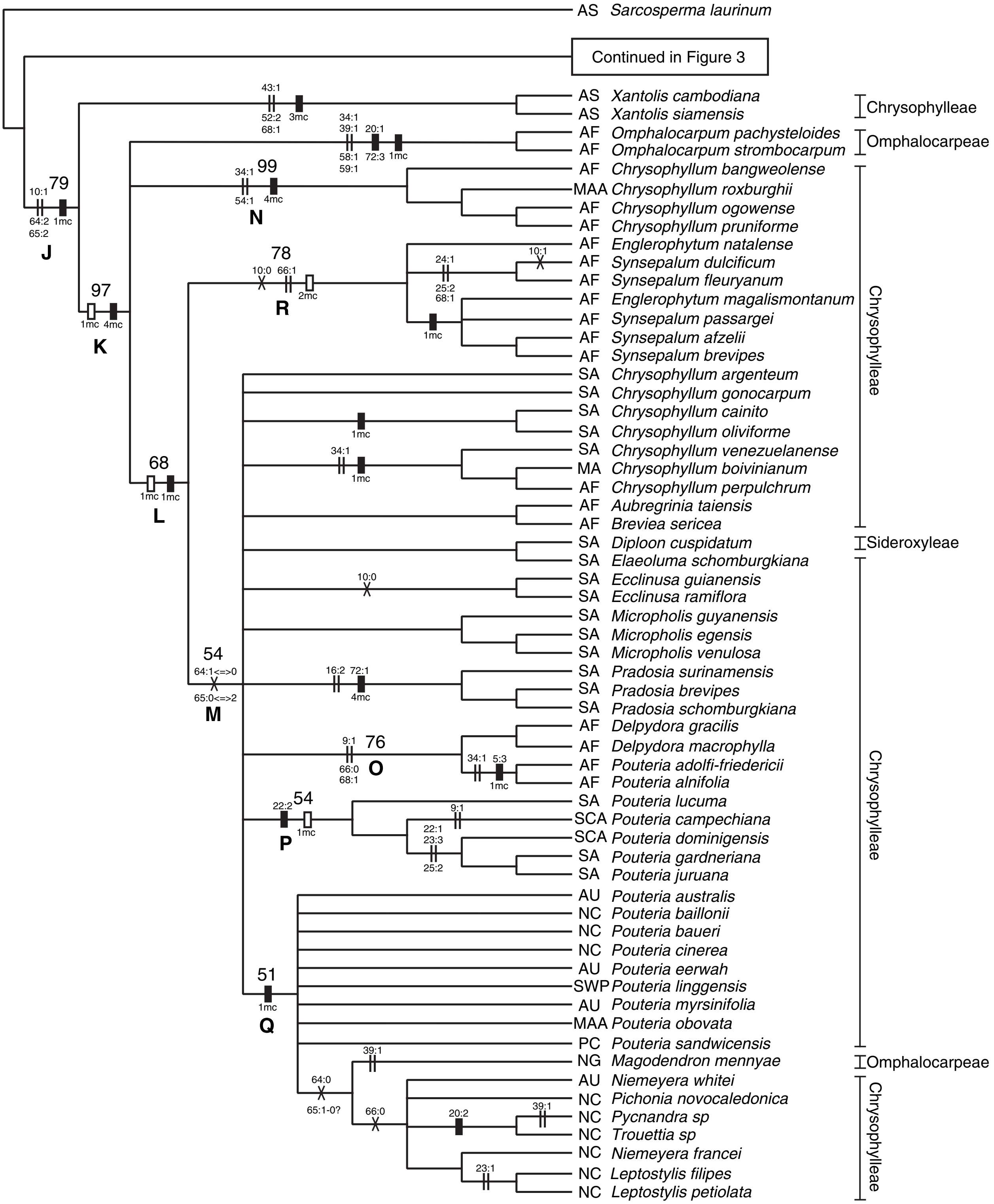


Fig. 4. As Fig. 3, but Clade J, the Chrysophylleae-Omphalocarpeae (Sapotaceae). Additional distribution areas: AU = Australia, MAA = Madagascar to Australia, NC = New Caledonia, NG = New Guinea, SA = South America, SCA = South and Central America, SWP = Southwest Pacific.

A polyphyletic Pouteria. This is the largest genus of Sapotaceae, with some 300 species distributed in the New World, Australasia, and a few in Africa (Govaerts et al., 2001). This group includes many satellite genera which were once accepted in a more narrow generic concept (Aubréville, 1964; Baehni, 1965) compared to that of Pennington (1991). Pennington recognized nine sections of *Pouteria*, six restricted to the New World, two to Australasia and the Pacific, and one section disjunct between South America and Africa. Four sections and a sample of species from all continents are represented in this study: *Oligotheca* (Australasia, Pacific). *Antholucuma* and *Pouteria* (South America) and *Rivicoa* (South America and Africa) (Table 1). Our analyses clearly indicate a polyphyletic origin of *Pouteria* in its present circumscription. Jackknife support for interrelationships is mostly below 50% for most *Pouteria* relationships. Molecular variation (*ndh*F) pertinent to *Pouteria* and its satellite genera is low and their relationships are unresolved. However, the jackknife support analysis of combined data and successively weighted characters recovered a consensus tree in which members of *Pouteria* group together in three distinguished clades. Interestingly they are fully congruent with geographic distribution (Fig. 4). First, two African species group with the African genus *Delpydora* (Clade O). Second, all New World species form a clade (Clade P). Third, all Australasian and Pacific species form another group recovered in the jackknife analysis (Clade Q).

In our analysis *Pouteria adolfi-friedericii* and *P. alnifolia* represent the African taxa formerly known as *Aningeria adolfi-friedericii* and *Malacantha alnifolia*. In a narrow generic concept, *Aningeria* circumscribes 3–4 species, whereas *Malacantha* is monotypic. According to Hemsley (1968), the former is distinguished from the latter by having subsessile flowers, lacking staminodes, and having seeds with a narrow seed scar. Pennington rejected the first two characters because pedicel length and absence of staminodes are not consistent, and he merged the two genera with *Pouteria,* giving them a systematic position in a group together with similar South American taxa in the section *Rivicoa*. This section is also represented here by *P. campechiana* and *P. lucuma*. Our analysis found *P. adolfi-friedericii* and *P. alnifolia* as sister to *Delpydora*, another African genus, not to the other species of *Pouteria* section *Rivicoa*. At present, this relationship has moderate jackknife support (76%), but morphologically the group can be recognized by their leaves, which have small translucent dots (*char. 9:1*), an embryo with included radicle (*char. 66:0*), and slender styles (*char. 68:1*). African *Pouteria* are further distinguished from South American ones in having marginocamptodromous leaf venation (*char. 5:3*), ciliate corolla lobes (*char. 34:1*), and by having a non-homoplasious molecular synapomorphy. For the time being it seems better to look upon the African *Pouteria* as a monophyletic group, perhaps referenced by a single generic name if a more narrow concept is adopted. An earlier proposed relationship to *P. sapota*, section *Aneulucuma* (see Anderberg and Swenson, 2003) was due to the misidentification of a cultivated specimen of *P. alnifolia*.

Pennington (1991) decided to reduce the Australasian species formally known as *Planchonella* into section *Oligotheca*. In a traditional view, species with endospermous seeds and foliaceous cotyledons have been placed in *Planchonella*. In contrast, those species with non-endospermous seeds and plano-convex cotyledons have been placed in *Pouteria* (Lam, 1925; van Royen, 1957a; Aubréville, 1964). Although Baehni (1965) did not recognize these characters as taxonomically important, a majority of the New World species have seeds without an endosperm, in contrast to the Australasian and Pacific species, which have endospermous seeds (Pennington, 1991). According to Pennington, species with an intermediate condition are known, and he concluded that no clear distinction could be made. Despite this, our jackknife analysis with successive weighting places all the species of *Pouteria* section *Oligotheca* together with all other Chrysophylleae taxa from Australasia and neighboring areas, in a single clade (Clade Q). Support for this group, however, is weak (51%). The type of cotyledons, presence of endosperm, and an exserted radicle are useful features to distinguish the section *Oligotheca* from *Pouteria* s. str. and other genera from Australasia, including *Leptostylis*, *Niemeyera* and *Pycnandra*. Several other characters can be added to these. The secondary veins are bow-shaped near the leaf margin in *Pouteria* s. str., but are straight until they meet a vein or loop of the tertiary order in section *Oligotheca*. A calyx with sepals in one or two series, or spirally arranged, and the type of aestivation, are also of interest. All investigated taxa of section *Oligotheca* have a symplesiomorphic type of calyx with quincuncial sepals in a single series, whereas the South American species have imbricate or valvate sepals which are frequently arranged in two series or spirals. The length ratio between the corolla lobes and tube is also important. In most species of the section *Oligotheca*, the lobes are about as long as the tube, whereas in other *Pouteria* the tube is longer than the lobes. All these characters conform well to the tree and Chrysophylleae, but it is not currently known if these characters remain diagnostic with a larger sample of species from *Pouteria*. Recent molecular evidence from the ribosomal genome, however, has indicated that most Australasian *Pouteria* species form several strongly supported groups with uncertain sister relationships (Bartish et al., 2005). One of these groups corresponds to section *Oligotheca*, that is *Planchonella*, which is not closely related to *Pouteria* species from South America. We therefore believe that

Planchonella merits recognition as a separate genus, but its circumscription has not yet been disclosed.

Other genera of Chrysophylleae. *Micropholis* is a genus of 38 species found in the New World. Its relationship to other genera are not clear. Pennington (1991) excluded most South American *Pouteria* from being related to *Micropholis* because they differ in leaf arrangement and lack an endosperm. Based on leaf venation he suggested an affinity to *Chrysophyllum* section *Donella*. Our analysis recovered moderate support (74%) for a monophyletic *Micropholis* and, using the MP solution (Fig. 2), weak support for a relationship with *Chrysophyllum* section *Chrysophyllum*. *Micropholis* differ from other *Chrysophyllum* in having a corolla without ciliate margin, sessile to subsessile anthers, and a style with round stigmatic areas.

Evidence for the monophyly of *Synsepalum* and *Englerophytum* was not obtained in our earlier molecular study of *ndh*F, but the combined data indeed recovered moderate support for it (Clade R, Figs 2 and 4). Two molecular synapomorphies support this group and another gives support to internal relationships. Both genera are African and together comprise approximately 50 species, but neither seems to be monophyletic. This group of taxa lacks a modern taxonomic treatment, and thus the generic boundaries are still unclear. For example, *Synsepalum* includes what was earlier recognized as *Afroseralisia*, *Bakerisideroxylon*, *Bakeriella*, *Pachystela* and *Vincentella* (Govaerts et al., 2001). A close relationship between *Synsepalum* and *Englerophytum* was suggested by Pennington (1991), who emphasized the frequent presence of stipules, 5-merous flowers, and similar seed and embryo characters. Stipules are very rare in Chrysophylleae and restricted to *Ecclinusa*, *Synsepalum and Englerophytum*. However, the presence of stipules is not consistent and they are missing in some species, e.g. *S. dulcificum*. The 5-merous flower is a plesiomorphy and therefore not diagnostic at this level. Furthermore, the seed and embryo characters are common in many members of Chrysophylleae. *Synsepalum* is diagnosed by long spreading corolla lobes, which separate them from the more erect ones in *Englerophytum*. This intricate group of taxa is in need of future research, preferably that which includes a taxonomic revision.

Pradosia is strongly supported, being defined by four molecular synapomorphies and at least one unique morphological character, namely its fruit. This is a one-seeded drupe with a thinly cartilaginous endocarp (*char. 72:1*), something that is not found elsewhere in the family. A ramiflorous inflorescence is also consistent in the sampled taxa (*char. 16:2*). *Pradosia* is distinguished from *Pouteria* by its absence of staminodes and from all *Chrysophyllum* and possible satellite genera by plano-convex cotyledons and the absence of an endosperm. With a few exceptions, a rotate corolla with spreading corolla lobes is also characteristic. However, exserted stamens are not a reliable character because when stamen and corolla lobe lengths are compared, stamens are shorter than the lobes in both *Pradosia* and *Chrysophyllum s. lat. The curvature of the corolla lobes gives the flower an appearance of having exserted stamens. Pennington (1991) discussed a possible relationship of Pradosia* to the Australasian genus *Niemeyera*. Our analyses nest *Niemeyera* within a clade with other Australasian taxa, not with *Pradosia*. This affinity is likely to be superficial, because *Niemeyera* is highly polyphyletic (Bartish et al., 2005). *Pradosia* is, or was, restricted to South and Central America, but *P. spinosa* was recently described in Africa (Ewango and Breteler, 2001). It is similar in many respects to *Pradosia*, for instance, in lacking staminodes and having a ramiflorous inflorescence. However, with a partly fused calyx and variation in the presence or absence of stipules, staminodes, and a 5-merous flower, it cannot be excluded that *P. spinosa* belongs to the *Synsepalum–Englerophytum* complex.

Morphological homoplasy

To find homologous morphological states for the cladistic analysis of Sapotaceae is difficult. Our results show that the morphological data set is homoplasious, and that few unambiguous diagnostic characters can be found for major evolutionary lineages. Doubtless, molecular data provide the necessary structure for a reconstruction of the basal cladogenesis, whereas morphology introduces noise at that level but some structure at the higher levels of the phylogeny. It may therefore be of interest to briefly discuss the reason for the observed extensive homoplasy.

The homoplasy in Sapotaceae is probably due to both a miscoding of non-homologous features of the same character, or due to convergence or reversals. Misinterpreted homologies will inevitably introduce errors into a matrix, and are due to the investigator (see Grant and Kluge, 2003). The presence and origin of staminodes is a candidate for such a character in Sapotaceae. Staminodes are distributed in the entire family and are believed to be derived from stamens by the loss of the anther (Pennington, 1991). This must be tested initially, and we therefore coded the presence or absence of staminodes as one character. We reveal here that the presence of staminodes may in fact be a non-homologous feature since they may be arranged in two different ways. Staminodes are either situated outside (above) the stamens, or in a single whorl together with the stamens. The former feature is confined to a crown group of Chrysophylleae-Omphalocarpeae (Clade M, Fig. 5). A reinterpretation of staminode presence into two different characters would affect other pertinent characters such as staminode form, position, and pubescence, which consequently would be superficial similarities, not homologous character states. One interesting observation is the point where anther filaments are fused with the corolla (char. 41). The plesiomorphic state is with stamens inserted at the top of the tube, but the stamens are also inserted within the tube, or near its base. These characters are, including African *Chrysophyllum*, restricted to the crown group mention above (Clade M). A different coding approach would therefore decrease homoplasy, recover a "clean" phylogenetic signal, and possibly form unambiguous synapomorphies.

Plasticity or variation can also cause extensive homoplasy in our data set. For instance, there is no doubt that the corolla is a homologous structure. Sapotaceae has a sympetalous corolla but the corolla tube varies from being very short (rotate flower) to long (tubular flower). A rotate flower predominates in Isonandreae-Mimusopeae-Sideroxyleae, whereas a cyathiform/tubular corolla is common in Chrysophylleae-Omphalocarpeae. Variation is even found within genera, e.g. *Payena*, *Pouteria* s. lat. and *Sideroxylon*. If the ratio of corolla tube/lobes was consistent, it would need two character state transformations, but on the current tree topology it changes in 30 steps. Other homoplasious characters, which most probably represent variation, include stamen filament length (char. 42) with 22 instead of two steps, ovary form (char. 62) with 28 instead of a single step, and the amount of endosperm (char. 65) with 23 instead of two steps as a minimum.

A new classification of the Sapotaceae

The results of our present analyses, as well as those from our earlier study (Anderberg and Swenson, 2003) show that the current suprageneric classification of Sapotaceae needs revision. Previous classifications have only been based on morphological information, but we have demonstrated that morphology alone does not contribute enough information for a stable classification, due to a high degree of homoplasy. Morphology combined with molecular data gives a more robust result, and identifies the same two main evolutionary lineages in Sapotaceae as the molecular data alone. We have earlier, and in this paper, referred to these lineages including the outgroup, by the informal names Clade 1 (*Sarcosperma*), Clade 2 or B (Isonandreae-Sideroxyleae-Mimusopeae), and Clade 3 or J (Omphalocarpeae-Chrysophylleae), but for communication purposes we believe that formal names for these clades would be useful.

We propose that three subfamilies are recognized, corresponding to the three main clades. Hence, as summarized in Fig. 5: Sarcospermatoideae (Clade 1: *Sarcosperma*), Sapotoideae (Clade 2 or B: Isonandreae-Mimusopeae-Sideroxyleae), and Chrysophylloideae (Clade 3 or J: Omphalocarpeae-Chrysophylleae). Within Sapotoideae, two supported clades, one with the well known name Sideroxyleae, should be maintained as tribes.

Subfamily **Sarcospermatoideae** (Lam) Swenson and Anderb., **stat. nov.**

Basionym: Sarcospermaceae Lam, Bull. Jard. Bot. Buitenzorg, Ser. 3, 7: 248 (1925).

Type: *Sarcosperma* Hook. f.

Nomenclatural note. Lam (1925) described "Sarcospermaceae" and later corrected the spelling to Sarcospermataceae (Lam and Varossieau, 1938). *Sarcosperma* could either be treated as a family or a subfamily of its own without violating the primary principle of monophyly in classification of plant families (Backlund and Bremer, 1998). The second order of principles include maximum stability and identification. *Sarcosperma* has most often been included in Sapotaceae (see, Pennington, 1991). Support for *Sarcosperma* as sister to the remaining Sapotaceae is very strong (100%), and as we agree with Pennington that *Sarcosperma* should be maintained in the Sapotaceae, we propose that this lineage is given the rank of subfamily.

Subfamily **Sapotoideae** Eaton, Bot. Dict., ed. 4: 35 (1836).

Type: *Sapota* Mill., nom. illegit. (≡ *Achras* L., ≡ *Manilkara* Adans., nom. cons.).

This subfamily has two recognized tribes, i.e., Sapoteae (see below) and Sideroxyleae. There are also several genera with uncertain affinity and left for the time being as tribus insertae sedis: *Burckella*, *Capurodendron*, *Diploknema*, *Inhambanella*, *Lecomtedoxa*, *Neolemonniera* (Fig. 5). Genera probably belonging to Sapotoideae but not sampled in this analysis are *Aulandra*, *Baillonella*, *Eberhardtia*, *Gluema*, *Isonandra*, *Labourdonnaisia*, *Letestua*, *Neohemsleya*, and *Vitellaria*.

Nomenclatural note. Clade 2 of Anderberg and Swenson (2003), here recovered as Clade B, corresponds to the large and well supported lineage of the Sapotoideae, including members of formerly recognized tribes Isonandreae, Mimusopeae and Sideroxyleae. Hartog (1878) was the first to introduce a formal classification of Sapotaceae and both Isonandreae and Mimusopeae were described by him. Lam (1939) elaborated on this system and was first to introduce subfamilies with the suffix -oideae, thus describing Madhucoideae, Mimusopoideae and Sideroxyloideae. All three subfamilies belong to Clade B, which includes also *Sapota* Mill., which is an illegitimate name for *Achras* L., now *Manilkara* Adans, nom. cons.

The International Code of Botanical Nomenclature (Greuter, 2000) states under Article 19.4 that: "The name of any subdivision of a family that includes the type of the adopted, legitimate name of the family to which it is assigned is to be based on the generic name equivalent to that type." Furthermore, in Article 19.5, it

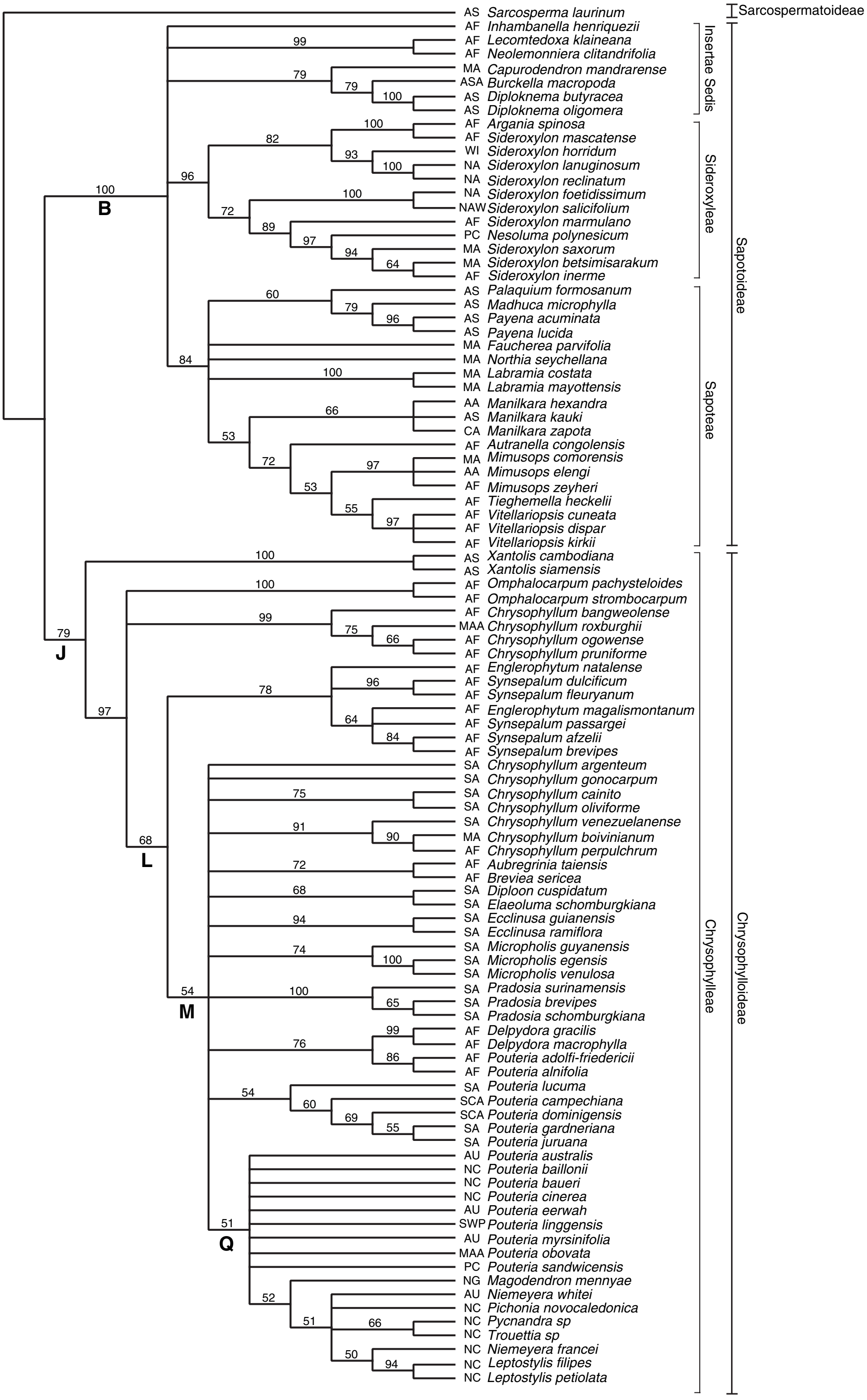


Fig. 5. Summary of a new proposed system of Sapotaceae classification to subfamilies and tribes (right) based on successive weighting analyses. Jackknife support for groups are above branches, distribution areas as abbreviated in Figs 3 and 4. Letters B and J correspond to the two major evolutionary lineages, Sapotoideae and Chrysophylloideae, L a sister relationship between *Englerophytum-Synsepalum* and a crown group of Chrysophylloideae, and M and Q are two branches with jackknife support not present in the MP solution.

is stated that such a subdivided name is not valid if it is based on an illegitimate generic name *if the family name is not conserved* [our italics]. The last sentence is crucial, since Sapotaceae is a conserved name (Greuter, 2000, p. 140), and therefore the use of the illegitimate name *Sapota* becomes necessary. Hence, the subfamily name has to be Sapotoideae, and the tribe in which *Manilkara zapota* (*Achras zapota*) belongs must be named Sapoteae.

Classification within Sapotoideae is non-problematic for tribes Sapoteae and Sideroxyleae. Circumscription of Sideroxyleae is maintained, except that *Sarcosperma* and *Diploon* are excluded. Sapoteae include the formerly recognized subtribes Manilkarinae and Mimusopinae, but exclude Glueminae (Mimusopeae), and one part of the tribe Isonandreae. Isonandreae is polyphyletic and it is uncertain to which of the two clades the nominal genus *Isonandra* (not included) will prove to belong. The heretofore Isonandreae genera (*Madhuca*, *Palaquium*, *Payena*) are included in the Sapoteae, because they are morphologically recognized and cladistically supported as part of this tribe. The other part of Isonandreae (*Capurodendron*, *Burckella*, *Diploknema*) forms a monophyletic group and could be recognized as Isonandreae nom. cons. if the nominal genus is placed within this clade. Subtribal classification of Sapoteae ("Mimusopeae") as proposed by Pennington (1991) does not correspond to monophyletic groups. Subtribe Manilkarinae forms a grade at the base of subtribe Mimusopinae, and subtribe Glueminae seems to form one or possibly two separate lineages, not related to any of other two subtribes.

Tribe Sapoteae Rchb., Handb. Nat. Pfl.-Syst. 214 (1837).

Type: *Sapota* Mill., nom. illegit. (≡ *Achras* L., ≡ Manilkara Adans.).

Included genera: *Autranella*, *Faucherea*, *Labramia*, *Madhuca*, *Manilkara*, *Mimusops*, *Northia*, *Palaquium*, *Payena, Tieghemella*, and *Vitellariopsis*.

Tribe Sideroxyleae H.J. Lam, Occas. Pap. Bernice Pauahi Bishop Mus. 14(9): 139 (1938).

Type: *Sideroxylon* L.

Included genus: *Sideroxylon* s. lat. This genus is in need of phylogenetic study in order to investigate if a more narrow generic concept is applicable. Presently, neither *Argania* nor *Nesoluma* can be upheld if *Sideroxylon* is to be maintained with its present circumscription.

Subfamily **Chrysophylloideae** Luerss., Handb. Syst. Bot. 2: 946 (1882).

Type: *Chrysophyllum* L.

Approximately 25 genera: *Aubregrinia*, *Breviea*, *Delpydora*, *Diploon*, *Ecclinusa*, *Elaeoluma*, *Chrysophyllum* s. lat., *Englerophytum*, *Leptostylis*, *Magodendron*, *Micropholis*, *Niemeyera*, *Omphalocarpum*, *Pichonia*, *Planchonella*, *Pouteria*, *Pradosia*, *Pycnandra*, *Synsepalum* and *Xantolis*. Genera not investigated but probably belonging to this subfamily are *Chromolucuma*, *Sarcaulus*, *Tridesmostemon* and *Tsebona*.

Nomenclatural note. Clade 3 of Anderberg and Swenson (2003), herein referred as Clade J, corresponds to the tribes Chrysophylleae and Omphalocarpeae. The circumscription of Chrysophylloideae discovered here corresponds exactly with the tribe Chrysophylleae, and Omphalocarpeae is eliminated from recognition as a taxonomic entity. Hartog (1878) described this tribe and called it a "division". Luerssen (1882, p. 945) followed Hartog (1878) but called the group "Unterfamilien" (subfamily). Under Article 19.6, improper use of termination is to be corrected without change of author and publication. Thus Luerssen's (1882) "Chrysophylleae" must be used after correction to Chrysophylloideae.

Acknowledgments

The directors at the herbaria C, GB, L, MO, P, U, UPS and WAG are acknowledged for permission to examine and to take samples from herbarium material. Angel Fernandez (La Gomera), Jérome Munzinger (Noumea), Laurie Jessup (Brisbane), Pranom Chantaranothai (Khon Kaen) and Wayne Takeuchi (Lae) are thanked for help in obtaining additional plant material. Mats Thulin, Mats Hjertson, James Reveal, Steve Farris and Mari Källersjö are thanked for help with discussions on nomenclature and successive weighting. We are also grateful to Bodil Cronholm for laboratory assistance. This study is part of an ongoing research project on the phylogeny of the Ericales. Financial support has been received from the Swedish Research Council to Ulf Swenson (grant 621-2002-330) and to Arne A. Anderberg (grant 621-2002-3740).

Department of Phanerogamic Botany, Swedish Museum of Natural History, PO Box 50007, SE-104 05 Stockholm, Sweden. Ulf.Swenson@nrm.se (Ulf Swenson); Arne.Anderberg@nrm.se (Arne A. Anderberg).

References

Anderberg, A.A., Rydin, C., Källersjö, M., 2002. Phylogenetic relationships in the order Ericales s.l.: analyses of molecular data from five genes from the plastid and mitochondrial genomes. Am. J. Bot. 89, 677–687.

Anderberg, A.A., Swenson, U., 2003. Evolutionary lineages in Sapotaceae (Ericales): a cladistic analysis based on *ndh*F sequence data. Int. J. Plant Sci. 164, 763–773.

Aubréville, A., 1961. Flore du Gabon. 1. Sapotacées. Muséum National d'Histoire Naturelle, Paris.

Aubréville, A., 1964. Les Sapotacées: taxonomie et phytogéographie. Adansonia, 1, 1–157.

Aubréville, A., 1967. Flore de la Nouvelle Calédonie et Dépendances. 1. Sapotacées. Muséum National d'Histoire Naturelle, Paris.

Aubréville, A., 1974. Flore de Madagascar et des Comores. 164e famille. Sapotacées. Muséum National d'Histoire Naturelle, Paris.

Backlund, A., Bremer, K., 1998. To be or not to be – principles of classification and monotypic plant families. Taxon, 47, 391–400.

Baehni, C., 1938. Mémoires sur les Sapotaceés. Candollea, 7, 394–508.

Baehni, C., 1965. Mémoires sur les Sapotacées 3. Inventaire des Genres. Boissiera, 11, 1–262.

Bartish, I.V., Swenson, U., Munzinger, J., Anderberg, A.A., 2005. Phylogenetic relationships among New Caledonian Sapotaceae (Ericales): molecular evidence for polyphyly and repeated dispersal. Am. J. Bot. 92, 667–673.

Bremer, B., Bremer, K., Heidari, N., Erixon, P., Olmstead, R.G., Anderberg, A.A., Källersjö, M., Barkhordarian, E., 2002. Phylogenetics of asterids based on 3 coding and 3 non-coding chloroplast DNA markers and the utility of non-coding DNA at higher taxonomic levels. Mol. Phyl. Evol. 24, 274–301.

Breteler, F.J., Nzabi, T., 1995. *Delpydora* (Sapotaceae), its taxonomy and myrmecophily. Bull. Mus. Nat. Hist. Nat., Paris, 4 sér., Section B, Adansonia, 17, 29–38.

van Bruggen, A.C., 1958. Revision of the Sapotaceae of the Malaysian area in a wider sense. XV: *Payena* A. De Candolle. Blumea, 9, 89–138.

Carpenter, J.M., 1988. Choosing among multiple equally parsimonious cladograms. Cladistics, 4, 291–296.

Cronquist, A., 1945. Studies in the Sapotaceae 3, *Dipholis* and *Bumelia*. J. Arnold Arbor. 26, 435–471.

Cronquist, A., 1946. Studies in the Sapotaceae 6. Bull. Torrey Bot. Club. 73, 465–471.

Ewango, C.E.N., Breteler, F.J., 2001. Présence du genre *Pradosia* (Sapotaceae) en Afrique: description d'une nouvelle espèce, *P. spinosa*. Adansonia, 23, 147–150.

Farris, J.S., 1969. A successive approximations approach to character weighting. Syst. Zool. 18, 374–385.

Farris, J.S., 1983. The logical basis of phylogenetic analysis. In: Platnick, N.I., Funk, V.A. (Eds.), Advances in Cladistics. Columbia University Press, New York, pp. 7–36.

Farris, J.S., 1989. The retention index and the rescaled consistency index. Cladistics, 5, 417–419.

Farris, J.S., 2001. Support weighting. Cladistics, 17, 389–394.

Farris, J.S., Albert, V.A., Källersjö, M., Lipscomb, D., Kluge, A.G., 1996. Parsimony jackknifing outperforms neighbor-joining. Cladistics, 11, 99–124.

Fitch, W.M., 1971. Towards defining the course of evolution: minimum changes for a specific tree topology. Syst. Zool. 20, 406–416.

Govaerts, R., Frodin, D.G., Pennington, T.D., 2001. World Checklist and Bibliography of Sapotaceae. Royal Botanic Gardens, Kew, UK.

Grant, T., Kluge, A.G., 2003. Data exploration in phylogenetic inference: scientific, heuristic, or neither. Cladistics 19, 379–418.

Greuter, W. (Ed.), 2000. International Code of Botanical Nomenclature. IAPT. Koeltz Scientific Books, Germany.

Hartog, M., 1878. On the floral structure and affinities of Sapotaceae. J. Bot. 16, 65–72.

Harvey, Y.B., Lovett, J.C., 1999. A new species of *Omphalocarpum* (Sapotaceae) from eastern Tanzania. Kew Bull. 54, 197–202.

Hemsley, J.H., 1968. Sapotaceae. In: Milne-Redhead, E., Polhill, R.M. (Eds.), Flora of Tropical East Africa. Whitefriars Press, London.

Hickey, L.J., 1973. Classification of the architecture of dicotyledonous leaves. Am. J. Bot. 60, 17–33.

Källersjö, M., Albert, V.A., Farris, J.S., 1999. Homoplasy increases phylogenetic structure. Cladistics, 15, 91–93.

Kornet, D.J., Turner, H., 1999. Coding polymorphism for phylogenetic reconstruction. Syst. Biol. 48, 365–379.

Labat, J.-N., Pignal, M., Pascal, O., 1997. Une nouvelle espèce de *Labramia* (Sapotaceae) de I'lle de Mayotte dans I'Archipel des Comores. Adansonia, 19, 213–216.

Lam, H.J., 1925. The Sapotaceae, Sarcospermaceae and Boerlagellaceae of the Dutch East Indies and surrounding countries, V des 'Contributions à l'étude de la flore des Indes Néerlandaises'. Bull. Jard. Bot. Buitenzorg. Series 3, 7, 1–289.

Lam, H.J., 1939. On the system of the Sapotaceae, with some remarks on taxonomical methods. Recueil Trav. Bot. Neerl. 36, 509–525.

Lam, H.J., van Royen, P., 1952a. Revision of the Sapotaceae of the Malaysian area in a wider sense. II: *Burckella* Pierre. Blumea, 6, 580–593.

Lam, H.J., van Royen, P., 1952b. Concise revision of the Sarcospermataceae. Blumea, 7, 148–153.

Lam, H.J., Varossieau, W.W., 1938. Revision of the Sarcospermataceae. Blumea, 3, 183–200.

Luerssen, C., 1882. Handbuch der Systematischen Botanik. Verlag von H. Haessel, Leipzig.

Maddison, W.P., 1993. Missing data versus missing characters in phylogenetic analysis. Syst. Biol. 42, 576–581.

Maddison, D.R., Maddison, W.P., 2000. MacClade 4: Analysis of Phylogeny and Character Evolution, Version 4.0. Sinauer Associates, Sunderland, MA.

Pennington, T.D., 1990. Flora Neotropica Monograph 52: Sapotaceae. New York Botanical Garden, New York.

Pennington, T.D., 1991. The Genera of the Sapotaceae. Royal Botanic Gardens, Kew, UK.

Pleijel, F., 1995. On character coding for phylogeny reconstruction. Cladistics, 11, 309–315.

van Royen, P., 1957a. Revision of the Sapotaceae of the Malaysian area in a wider sense. VII. *Planchonella* Pierre. Blumea, 8, 235–445.

van Royen, P., 1957b. Revision of the Sapotaceae of the Malaysian area in a wider sense. VI: *Xantolis* Raf. Blumea, 8, 207–234.

van Royen, P., 1958. Revision of the Sapotaceae of the Malaysian area in a wider sense. XIV: *Diploknema* Pierre. Blumea, 9, 75–88.

Sennblad, B., Bremer, B., 2000. Is there a justification for differential *a priori* weighting in coding sequences? A case study from *rbc*L and Apocynaceae s.l. Syst. Biol. 49, 101–113.

Simmons, M.P., Freudenstein, J.V., 2002. Artifacts of coding amino acids and other composite characters for phylogenetic analysis. Cladistics, 18, 354–365.

Slowinski, J.B., 1993. 'Unordered' versus 'ordered' characters. Syst. Biol. 42, 155–165.

Stevens, P.F., 1991. Character states, morphological variation, and phylogenetic analysis: a review. Syst. Bot. 16, 553–583.

Strong, E.E., Lipscomb, D., 1999. Character coding and inapplicable data. Cladistics, 15, 363–371.

Swofford, D.L., 2002. PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods), Version 4. Sinauer Associates, Sunderland, MA.

Swofford, D.L., Olsen, G.L., 1991. Phylogeny reconstruction. In: Hillis, D.M., Moritz, C. (Eds.), Molecular Systematics. Sinauer Associates, Sunderland, MA, pp. 411–501.

Swofford, D.L., Olsen, G.L., Waddell, P.J., Hillis, D.M., 1996. Phylogenetic Inference. In: Hillis, D.M., Moritz, C., Mable, B.K. (Eds.), Molecular Systematics. Sinauer Associates, Sunderland, MA, pp. 407–514.

Vink, W., 1958. Revision of the Sapotaceae of the Malaysian area in a wider sense. XIII: *Chrysophyllum* L. Blumea, 9, 21–74.

Vink, W., 1995. Revision of *Magodendron* (Sapotaceae) with observations on floral development and morphology. Blumea, 40, 91–107.

Waterman, P.G., Mahmoud, E.N., 1991. Chemical taxonomy of the Sapotaceae: Patterns in the distribution of some simple phenolic compounds. In: Pennington, T.D. (Ed.), The Genera of the Sapotaceae. Royal Botanic Gardens, Kew, UK, pp. 23–50.

Wenzel, J.W., Siddall, M.E., 1999. Noise. Cladistics, 15, 51–64.

Wiens, J.J., 1995. Polymorphic characters in phylogenetic systematics. Syst. Biol. 44, 482–500.

Wiens, J.J., 2001. Character analysis in morphological phylogenetics: problems and solutions. Syst. Biol. 50, 689–699.

Wiens, J.J., Servedio, M.R., 1997. Accuracy of phylogenetic analysis including and excluding polymorphic characters. Syst. Biol. 46, 332–345.

Wilkinson, M., 1992. Ordered versus unordered characters. Cladistics, 8, 375–385.

Wilkinson, M., 1995. A comparison of two methods of character construction. Cladistics, 11, 297–308.

Yang, T.Y.A., 1998. The family Sapotaceae in Taiwan. Taiwania 43, 193–202.

Appendix A

Morphological characters and character states used in the present cladistic analysis of Sapotaceae. Characters treated as ordered in the analyses are marked with an asterisk (*).

Leaves, shoots and hairs

1. Leaves alternate (0); opposite (1).

2. Leaves along the branches spirally arranged (0); distichous (1).

3. Leaves spaced along branches (0); clustered at apex (1).

4. Leaves not fascicled (0); fascicled on brachyblast (1). Frequently, leaves are clustered at branch apices or fascicled on brachyblasts, i.e., short, condensed spur shoots. Composite coding of a these characters is feasible, but any combination of the discernible states will create problems of homology or the problems pointed out by Simmons and Freudenstein (2002). For example, some species have the majority of leaves condensed on brachyblasts, which could be perceived as a third state to alternate/opposite. However, those species also have young or long shoots with alternate leaves, and, hence, would have to be coded as polymorphic. Thus, reductive, binary characters were used for the leaves.

5. Leaf venation eucamptodromous (0); craspedodromous (1); brochidodromous (2); marginocamptodromous (3). Pennington (1990, 1991) emphasized that leaf venation provides several useful characters on both a generic, sectional and species level. Depending on how secondary veins behave at the margin, Pennington identified leaves as being craspedodromous, eucamptodromous or brochidodromous (Hickey, 1973). Secondary veins in eucamptodromous leaves gradually disappear and never reach the leaf margin, whereas secondaries terminating in the leaf margin are called craspedodromous. However, in several species the secondaries bend at the margin, and anastomose with the uppermost approximate secondary vein to form a marginal vein. Hickey (1973) did not mention this type, which is here termed *marginocamptodromous*.

6. Secondary veins towards the margin arcuate (0); straight (1). Pennington (1990, 1991) described many taxa with brochidodromous leaves as having arcuate secondaries. We find such leaves having the secondaries joined by a series of submarginal loops, either by the secondaries or by veins, possibly better referred to the tertiary order. This character refers to the secondary vein between the midrib and to the point where it meets the next vein, either next secondary or a loop of tertiary order to form the brochidodromous pattern.

7. Intersecondary veins inconspicuous or absent (0); conspicuous (1). These veins originate at the midrib and most often run parallel to the secondaries, but fade away approximately in the middle between the midrib and the margin.

8. Tertiary veins crossing between secondaries (0); reticulate (1); parallel to secondaries (2); indistinguishable (3). Tertiary veins are the next order of veins often visible to the naked eye. Pennington (1990) noted three main patterns and a few subpatterns. We found the main patterns possible to track, but not the subpatterns. One frequent pattern is when tertiary veins run across to rejoin with the next adjacent secondary. This state includes both horizontal and oblique tertiaries (see Pennington, 1990). The reticulate pattern is formed when tertiaries anastomose with secondary or other tertiary veins. Tertiaries parallel to the secondaries run from the leaf margin toward the leaf axis and often decrease in size.

9. Leaves without transparent structures (0); with transparent lines or dots (1).

10. Stipules present (0); absent (1). Stipules, when present, are most frequently caducous, but persistent in a few species. Our sampling includes only one species with persistent stipules, *Neolemonniera clitandrifolia*, a character not used (autapomorphy). If this character is to be used in future analyses where it could become informative, it ought to be composite coded and ordered to avoid inapplicable states and losses of hierarchic information. For example, binary coding of this information will introduce missing data in taxa without stipules.

11. Leaves glabrous (0); hairs present (1). Indumentum of Sapotaceae on leaves, petioles, petals, sepals and ovary are characterized by malpighian hairs. *Delpydora* is the only genus covered by simple hairs (Pennington, 1991), but malpighian hairs are also sparsely intermixed (Breteler and Nzabi, 1995). Moreover, many taxa loose their hairs on fully developed leaves. Considering this variation and the possibility that simple hairs could be malpighian hairs with one reduced branch, a simple presence/absence character was used to score hairs on the lower surface of leaves, especially on young leaves.

12. Hairs brown (0); white (1); yellow (2).

13. Plants unarmed (0); armed by unbranched shoots (1).

14. Hollow pits in the axile of secondary veins present (0); hollow pits absent (1).

15. At the base of leaf lamina, no pouches for myrmecophily (0); pouches for myrmecophily (1).

Inflorescence, calyx and corolla

16. Inflorescences axillary (0); in the axils of fallen leaves (1); ramiflorous (2); pseudo-terminal (3); cauliflorous (4). Inflorescences of the Sapotaceae are to a large extent constant throughout the family. Most species have a simple fascicle, which is sometimes reduced to a single flower. Apart from *Sarcosperma*, which has flowers arranged in racemes, the position and development of the fascicles vary. Most frequently, the fascicle is in the axil of a persistent or fallen leaf. There is a clear transition between these two states, which is why many taxa are coded as polymorphic in this character. Ramiflory is easily confused with fascicles in the axils of fallen leaves, occurs in various genera of the Sapotaceae, and is believed to be of little systematic value (Pennington, 1991). In scoring ramiflory and to distinguish it from fascicles in the axils of fallen leaves, fascicles of flowers must reappear on old wood below the leaves and without any adjacent leaf scar. Cauliflory is scored only for taxa where flowers appear on the main trunk. Another type of inflorescence is called pseudo-terminal. Here, the flowers are clustered in the axils of small scaly leaves at the shoot apex, which after anthesis continues to grow into a leafy vegetative shoot (Pennington, 1991).

17. Flowers not on brachyblasts (0); brachyblasts (1). In some cases, the fascicle produces flowers time after time and develops into a short, stout spur shoot, a brachyblast. This is a modification of the fascicle (Pennington, 1991) and, hence, a separate character, not composite coded with inflorescences.

18. Flowers in racemes (0); solitary or in fascicles (1).

19. Flowers pedicellate (0); subsessile or sessile (1).

20. Flowers subtended by bract (0); bracts several, spirally arranged, and inserted at the base of the pedicel (1); bracts several, spirally arranged, and distributed along the pedicel (2). In some cases, several bracts are distributed along the pedicel. The uppermost bract is then often large, of the same size as the sepals, and must be distinguished from the calyx.

21. Flowers not carnose (0); carnose (1).

22. Calyx in one series (0); two series (1); spirally arranged (2).

23. Sepals in each series ≥ 5 (0); 4 (1); 3 (2); 2 (3). Taxa with two calyx series have rarely series of 5 sepals, but 4, 3, or 2. Provided the option that the biseriate calyx evolved through a duplication, characters 22 and 23 cannot be ordered.

24. Sepals free (0); partly fused to at least a third of its length (1). Sepals are most often free, but can be partially fused. It was necessary to score the character when at least a third of the sepal length was fused.

25. Sepals quincuncial (0); imbricate (1); valvate (2).

26. Sepals (sub)glabrous (0); hairy (1); woolly (2).

27. Calyx pubescence not differentiated (0); differentiated (1). Calyx indumentum is variable, especially on taxa having biseriate calyx. In order to capture this variation, two characters were constructed.

28. Corolla in 1 series (0); 2 pseudoseries with strongly overlapping petals (1). In some taxa, the corolla lobes are overlapping. This arrangement is particularly obvious in some taxa with divided corolla lobes (char. 36), where two lateral segments are attached on the dorsal surface of the median segment. We refer to this as a corolla with two pseudoseries, a type found in members of the subtribe Mimusopinae, *sensu* Pennington (1991). The ontogeny of this corolla type is presently unknown.

29. Corolla with sepals isomerous (0); dimerous (1); trimerous (2); anisomerous (3). Corolla structure provides some of the most important characters for Sapotaceae systematics. Flowers are always actinomorphic, sympetalous, and having at least one stamen opposite each corolla lobe, except in strictly female flowers. Most often, the number of sepals and petals correspond and can be coded as iso-, di-, or trimerous. Deviation from this pattern, e.g., five sepals and eight petals, is found in several taxa, a state termed anisomerous. It could be argued that this coding loses potential synapomorphies and it would better to code the number of petals present (4, 5, 6, 7, etc.). However, this also confers a problem, i.e., too many character states become necessary, most taxa would be polymorphic, and unnecessary noise would be introduced.

*30. Corolla tube shorter than lobes (0); tube and lobes more or less equal (1); tube longer than lobes (2). The fused proportion of the corolla tube varies in length and influences the corolla shape to be rotate (often short tubes), cyathiform (tube and lobes of ± equal length), or tubular (lobes short). Pennington (1991) coded anthers as exserted or included, but we find this trait to be a complex character partly dependent on the degree that the lobes bend outwards and the length of the stamens. For instance, a taxon may have exserted anthers, but only after the corolla lobes are curved backwards. To avoid ambiguous situations, we therefore scored corolla lobe inclination in full bloom as ≥90 degrees (spreading), 25–90 degrees (reflexed), and more or less erect (char. 31), together with stamen length (char. 42).

31. Corolla lobes spreading (0), reflexed (1); erect or infolded (2).

32. Corolla aestivation imbricate (0); valvate (1); contorted (2); quincuncial (3).

33. Corolla glabrous (0); hairy (1).

34. Corolla margin eciliate (0); ciliate (1).

35. Corolla creamish (0); white (1); greenish (2); pale yellow (3); red (4).

36. Corolla lobes entire (0); subdivided into 2 segments (1); subdivided into 3 segments (2).

37. Corolla segments ± equal size (0); lateral shorter than median (1).

Androecium

38. Stamens opposite corolla lobes (0); alternate (submarginal) and opposite (1). As mentioned above, essentially all flowers have at least one anther opposite each corolla lobe, but an increased number of stamens is found in the tribes Isonandreae and Omphalocarpeae. Some taxa with more stamens than petals, provided the staminodes are missing, may in fact have stamens alternating with the corolla lobes.

39. Stamen isomerous to each corolla lobe (0); di- or polymerous (1). Stamens are sometimes more than one opposite each corolla lobe, often 2–3, but may vary even more. When variation is present, it often occurs in the same taxon, which is the reason why clear discrete states were impossible to formulate. Stamen number is therefore scored as isomerous or di- to polymerous to the corolla lobes.

40. Stamen inserted in 1 whorl (0); 2 whorls (1). Most flowers have the stamens inserted in one whorl, but in few cases, 2–3 whorls are present below each other.

41. Stamen fixed at the top of the tube (0); in the tube (1); near the base (2); on the lobes (3).

*42. Stamen subsessile, anther filament very short (0); shorter than corolla (1); longer than corolla (2). In some cases, the anther filaments are fused to the corolla for a small portion. The length of stamen was then scored from the point where the filament becomes free. As mentioned above, stamen versus corolla length was used in order to describe exserted or included stamens. Stamen length was determined by the total length of the filament and the anther. When filament length is half of the anther, it was coded as (sub)sessile. In cases where half of the anther is as long as, or the entire stamen is longer than the corolla, it was, respectively, scored as shorter or longer than the corolla. This character is a transition series and ordered.

43. Anthers basifixed (0); dorsifixed (1). Most anthers are basifixed, but some are dorsifixed, a state not to be confused with taxa having spurred anthers (char. 45). The two spurs are often so closely adjacent to each other that the connective superficially seems to be attached to the dorsal surface of the anther.

44. Anther connective fixed (0); flexible (1).

45. Anthers ecalcarate (0); calcarate (spurred) (1). See also char. 43.

46. Anthers oblong (0); ovate (1); lanceolate (2).

47. Anthers latrorse (0); extrorse (1).

48. Anther filament geniculate (0); not geniculate (1). A filament may be geniculate in different ways, sometimes contorted, but we have not been able to classify these into separate character states.

49. Anther filament glabrous (0); hairy (1).

50. Anther thecae glabrous (0); hairy (1). Pennington (1991) stated that the stamens are hairy or glabrous, without distinction between the filament and the thecae. We found this to be a simplification, because hairs are confined to either the filament, thecae, or both. Separate characters were found to be appropriate

*51. Anther filaments free (0); partly fused in groups (1); completely fused in groups (2). Filaments are generally free but in some cases become partly to entirely fused in groups. This character with three states is perceived as a transition series and ordered.

52. Anther without appendage (0); appendage minute beaked (1); appendage triangular and acute (2); appendage irregular (3).

53. Anther thecae uniloculate (0); pluriloculate (1).

Staminodes

54. Staminodes well developed (0); absent (1).

55. Staminodes fixed with anthers in a single whorl (0); outside or above anthers (1). Staminodes are present in all tribes except Isonandreae (Pennington, 1991). Except for *Gluema* (not included), staminodes alternate with the fertile stamens and are normally positioned in the lobe sinuses.

56. Staminodes short, flat, often rounded apex (0); petaloid (1); filamentous or subulate (2); carnose, stamen-like (3). Staminodes are classified here into four morphological types. Petaloid is the most common type and similar to a petal, i.e., often ovate and flat. Small, short, and flat staminodes with a round apex are distinguished from filamentous or subulate, which are terete rather than flat. Staminode morphology was even coded for taxa with usually vestigial staminodes.

57. Staminodes margin entire (0); lobed (1); dentate to fimbriate (2); aristate at apex (3).

58. Staminodes erect (0); inflexed against style and ovary (1).

59. Staminodes glabrous (0); hairy (1).

Gynoecium and embryo

*60. Ovary glabrous (0); subglabrous (1); hairy (2).

*61. Ovary 1–2-locular (0); 3–5-locular (1); 6–8-locular (2); ≥ 9-locular (3). The Sapotaceae ovary has 1-many locules, each being uniovulate. A 3–5-locular ovary is frequent, but deviations from this number are also common. Few species have an ovary of 1–2 locules, others with 6–8 locules, and still others with more than 9 locules. Intermediate numbers of locules are also present and in order not to infer a state for each locus, we perceived the character states in conjunction with the number of sepals and petals. However, there are polymorphic taxa and we believe that the best way to treat this character is as ordered

62. Ovary broadly ovoid to subglobose (0); conical (1).

63. Placentation basal to basi-ventral (0); axile (1). Placentation describes the position of the ovule to the placenta. This point of attachment will eventually

develop to a seed scar in the mature fruit. In order to not duplicate this character, placentation was coded but not the seed scar.

64. Cotyledons plano-convex (0); thick and flat (1); foliaceous (2).

*65. Endosperm absent (0); present and scarce (1); present and copious (2). An embryo with plano-convex cotyledons usually lacks endosperm, whereas an embryo with foliaceous cotyledons occurs with a developed endosperm (Pennington, 1991). However, this is not always true, because there are a small number of taxa having intermediate type of cotyledons and a small amount of endosperm. In reference to the endosperm, three character states were therefore used and treated as ordered.

66. Radicle included (0); extended to surface (1); exserted (2).

67. Style included in flower (0); clearly exserted out of the flower (1). A style can either be included or exserted. Like scoring the length of the anther, style length was compared to the length of the corolla.

68. Style stout (0); slender (1). The difference between a stout and a slender style could be described as the ratio between length and width. A stout style is 2–3-fold longer than its width, whereas a slender style is ≥ 3-fold longer than its width.

69. Style with round stigmatic areas (0); minutely lobed (1); simple (2). A style is at one point attached to the ovary and at the other end has stigmatic areas. Most frequently, the stigmatic areas are smooth and cannot be distinguished from the sterile tissue. However, especially in Chrysophylleae, several species have round stigmatic areas at the style apex. It is easy to envision a further development of those areas into minute lobes, found in *Magodendron*.

70. Style glabrous (0); glabrous in the upper part and hairy in the lower part (1); hairy (2).

71. Styles not jointed (0); jointed (1).

Fruits and seeds

72. Fruit a berry (0); drupe (1); capsule (2); drupe-like with a woody pericarp (3). The Sapotaceae fruit is a fleshy berry except in a few cases, where the outer pericarp is leathery with an endocarp being cartilaginous (drupe) or hard (capsule)

*73. Fruits glabrous (0); hairy (1); hispid (2). Hairy ovaries are not correlated with hairy fruits, which is why this is used as a separate character. In addition, a hispid hairy fruit is hairy with a special long type of hairs, and is thus an ordered character.

74. Fruits 1-seeded (0); 2–3-seeded (1); ≥ 5-seeded (2).

75. Testa smooth (0); roughened (1); papyraceous (2).

76. Seeds not fused (0); fused along the adaxial surface (1).

77. Seeds globose (0); ellipsoid (1); ovoid (2); obovoid (3).

*78. Seeds not laterally compressed (0); laterally compressed (1); strongly laterally compressed (2). Pennington (1991) addressed the importance of seed shape, the position of the seed scar, and surface portion between the seed scar and the smooth testa, used for generic classification. Discrete states have been possible to score for seed form and seed scar position, but the third character forms a continuum, which we have not included.

Appendix B

Data matrix of 78 morphological characters of Sapotaceae. *Sarcosperma laurinum* was used as outgroup. In the matrix, inapplicable states are coded with a dash (-) and polymorphic taxa with letters: a = 0/1; b = 1/2; c = 0/2; d = 0/3; e = 1/3; f = 2/3; g = 0/1/2; h = 1/2/3; i = 0/2/3; ? = missing data

Taxon	Character number							
	0000000001	1111111112	2222222223	3333333334	4444444445	5555555556	6666666667	77777777
	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	12345678
Argania spinosa	0001c10101	0-11000110	0000000000	000000-000	0211011100	0000020002	a102221120	000a0100
Aubregrinia taiensis	0000000101	0-01000100	0000000002	2?00?0-000	1100111100	0000120002	2012220020	00?200?1
Autranella congolensis	0010210200	0-01000100	0110221102	1?10010000	0000101000	3000001002	2102220020	000a0031
Breviea sericea	0100011?01	1a01000100	0000020002	200120-000	1101120100	0000122002	f012220020	00?20032
Burckella macropoda	00102??100	0-01030100	0011100010	2010?0-01?	0100?21110	0?01-----0	1?10011120	00000010
Capurodendron mandrarense	00111a0100	1101001100	0001020001	020010-000	0101120100	0200010112	1011011120	000000?0
Chrysophyllum argenteum	0100000101	1001000100	0000010002	2110c0-000	0001011000	0001-----2	b012220020	001g0002
Chrysophyllum bangweolense	0000201101	1001000100	0000000001	2?0110-000	b101011100	0001-----2	1112220020	000a0002
Chrysophyllum boivinianum	0000000001	1a010a0110	0000010001	200100-000	2101011100	0001-----2	1112220020	001g0012
Chrysophyllum cainito	0100211101	1001000100	00b0110001	c010f0-000	0001011100	0001-----2	3012220000	000b001a
Chrysophyllum gonocarpum	0000001a01	11010a0100	0000010001	200020-000	1101011100	0001-----2	1012220020	000g0012
Chrysophyllum ogowense	01002a1101	10010a0100	0000000001	200120-000	2101011000	0001-----2	1112120020	00020012
Chrysophyllum oliviforme	0100211101	1001000100	00a0110002	00a020-000	0001011100	0001-----2	b012220000	001000b0
Chrysophyllum perpulchrum	0000000001	10010a0110	0000010001	230120-000	b101011100	000a000002	1012220020	00120012
Chrysophyllum pruniforme	0100211101	10010a0100	0000000001	200120-000	1101011100	0001-----2	1012120020	00020012
Chrysophyllum roxburghii	0100211101	10010a0100	0000000001	2?01?0-000	11???11???	0?01-----2	1012220020	00120032
Chrysophyllum venezuelanense	0000000001	10010a0100	00000a0001	200120-000	210101?100	000a000002	1012220020	00020022
Delpydora gracilis	0000201011	1001aa0100	0000010002	200010-000	2100121100	0001-----2	1010000100	00220011
Delpydora macrophylla	0000201011	1001aa0100	0000010002	200010-000	2100121100	0001-----2	1010000100	00220011
Diploknema butyracea	0010000000	11010a0100	00a1a20011	c20030-010	0100121000	0201-----2	f110011121	001a0010
Diploknema oligomera	00100a0000	1101000100	00a1a20011	a200?0-011	01??1?1000	?201-----2	f110011101	00100010
Diploon cuspidatum	0100211001	0-010c0100	00a0100000	1?0010-000	0100010100	0001-----0	0000010020	00000010
Ecclinusa guianensis	00a0c00000	a001010110	0000010000	200000-000	0101111100	0001-----2	1010010000	001g0011
Ecclinusa ramiflora	0010000000	1001010110	0000010000	201020-000	0101111100	0001-----2	1a10010000	00110011
Elaeoluma schomburgkiana	00a0200101	0-01000100	0000a00000	0?0010-000	0100000100	000a000002	1a10120120	00000000
Englerophytum magalismontanum	0000211100	10010c0100	0000010001	2000a0-000	0100110000	010a020002	1110010021	00?0001a
Englerophytum natalense	0000211100	1001000110	0000010001	200020-000	01001b0100	010a020002	1110010020	00100010
Faucherea parvifolia	0010210201	0-01000100	0120201000	1?0020-000	0100011100	0000002002	b012221120	00100021
Inhambanella henriquezii	0010000100	0-01000100	0000010000	0?00d21000	0101121000	0000010002	1010020020	0000a032
Labramia costata	001021110?	0-01010100	0120201001	0?00121000	0100121100	0200002000	2112221120	00000010
Labramia mayottensis	0010211101	0-01010100	0120201001	0?00121000	0100121100	0200001000	3012220120	00000010
Lecomtedoxa klaineana	0010200200	0-01000100	0000010000	0?a0?21000	0100111101	0c00010012	1011120020	02000032
Leptostylis filipes	1010210101	0-010c0100	0010100032	100040-000	1201111000	0001-----2	11100a1120	10000010
Leptostylis petiolata	1010210101	0-01020110	0010100030	100010-000	0201111000	0001-----2	11100a1120	10000010
Madhuca microphylla	0000210100	0-01000100	01302010b1	0000?0-111	0000?2??10	?201-----0	1?1aa21120	0000001b
Magodendron mennyae	0010000101	0-01040100	000001000b	030030-010	020101?001	00101121a0	ba10a20010	00000030
Manilkara hexandra	001021a101	0-01000100	0120201000	1000020000	0101111000	0100012001	f002221120	00a00031
Manilkara kauki	0010210101	0-01000100	0120211000	0000520000	0101121100	0000011002	f002221120	000?0031
Manilkara zapota	0010210101	0-01000100	01hab2100b	00001c0000	0101111100	00000100a2	f002221120	000b00e1
Micropholis egensis	0100210301	0-01000100	00a0a10002	b?1010-000	0a00111100	0000120002	1012220000	000000e2
Micropholis guyanensis	0a00210101	10010a0100	0000a10002	b?00a0-000	0000111100	0000120002	1112220000	001000e2
Micropholis venulosa	0100210301	10010c0100	00a0a10002	b010f0-000	0a00111100	0000120002	1012220000	000000e2

Mimusops comorensis	0000210200	0-01000100	0110211100	2?00110000	0101021110	0200010112	2002221121	00000011
Mimusops elengi	0000210200	0-01000100	0110221100	2?00010000	0101021110	0200010112	2002220120	00100011
Mimusops zeyheri	0000210200	1001000100	0110221100	2?10010000	0101021110	0200010112	2002221120	00010011
Neolemonniera clitandrifolia	1010001100	0-01000100	0000010001	2?00121000	0100121110	0000010112	1111120022	02000012
Nesoluma polynesicum	0000211101	10010a1100	00a0a000e0	c000c0-000	02011a1000	00000b0002	b112220120	000000e0
Niemeyera francei	0010201101	0-010c0100	0000a200d1	100110-000	0211111000	0101-----2	1110001120	10010010
Niemeyera whitei	0000200a01	10010c0110	0000120000	1000?0-000	0211111000	0101-----2	1010001120	00000000
Northia seychellana	0010111101	1001000100	1120211001	2210221000	0110011000	0000000002	2010011120	00100010
Omphalocarpum pachysteloides	0000001101	0-01040101	0000020000	2d0100-010	010111110a	0000010112	1012220020	03120012
Omphalocarpum strombocarpum	0000001101	0-01040101	0000010000	2d0110-010	0101121100	1000012112	1012220020	03020012
Palaquium formosanum	0010210200	0-010a0100	0120211000	1200f0-010	3100001100	0001-----0	1012111120	00000011
Payena acuminata	0100210200	1001000100	0130221110	c00010-110	000000a100	2201-----2	b012221120	001a0012
Payena lucida	0100210200	a101000100	013a211111	2000e0-110	000000a000	2201-----2	h012221120	00100012
Pichonia novocaledonica	1000210101	0-01000100	00a0010000	100040-000	320111a000	0000010002	1110101121	00000010
Pouteria adolfi-friedericii	0000300011	10010c0100	01f0210002	2001c0-000	a101110100	0000120002	1010000100	001000b0
Pouteria alnifolia	0000300011	11010a0110	0000010002	200130-000	11011a0100	000a?????2	1010000101	00100011
Pouteria australis	0000211101	0-01000100	0000a10001	200000-000	2001110000	0000120002	1012220000	001g0011
Pouteria baillonii	0010210101	1001000100	00aa020002	230040-000	0001111100	000012c002	1112221101	00100011
Pouteria baueri	0010210101	0-010a0100	0000010001	230010-000	2001111100	0000100002	1112220021	00000011
Pouteria campechiana	0010000011	0-010a0100	02001110d2	200020-000	0a010b1000	000002c002	ba10010000	000g0010
Pouteria cinerea	0010211101	1101000100	0000010001	2000?0-000	1a01110100	0000100002	1012120021	00ab0011
Pouteria dominigensis	00102a1b01	a0010a0100	01302110d2	200020-000	0001121000	0000120002	b010010000	000g0010
Pouteria eerwah	0000211101	0-01000100	0000000001	2000?0-000	000111?000	0000100002	1012220000	00000031
Pouteria gardneriana	0000001101	0-010a0100	0130211002	2?0020-000	10011b1100	0000100002	a010000001	00000010
Pouteria juruana	00a0000001	a0010a0100	0130201002	2?1030-000	1????11?00	0?00?20002	1110000020	00?0001?
Pouteria linggensis	00002a1101	0-010a0100	0000010001	200020-000	1001110000	0000100002	1012120020	000g0011
Pouteria lucuma	0010c00001	1a010a0100	0200110002	2000c0-000	0001111000	0000020002	b010010020	000a0011
Pouteria myrsinifolia	0000211101	11010a0100	00a0010001	230020-000	0a01110100	0000120002	1012220000	001a00e2
Pouteria obovata	0000211101	10010a0100	0000000001	2300b0-000	0101110100	0000120001	1012220020	000a0031
Pouteria sandwicensis	0000211101	1a01000100	00a0010001	2?0020-000	1001110000	0000120002	1112120020	0010000a
Pradosia brevipes	0010300001	1001020100	0000010001	001020-000	0101011000	0001-----2	1a100b0020	01100011
Pradosia schomburgkiana	1000c01101	1a01020100	00a??10001	0?10h0-000	01?1?a?000	0?01-----2	b010020020	011000a1
Pradosia surinamensis	00a0000101	0-01020110	0000010001	001020-000	010111?000	0001-----2	1010010020	011000b1
Pycnandra sp	0010000101	0-010g0102	00000100d1	100010-010	0201121101	0101-----2	b010010020	0000001?
Sarcosperma laurinum	0000000000	0-00000000	0000010000	0d00c0-000	000001c000	0000000000	00000a0000	000a0000
Sideroxylon betsimisarakum	00102a1101	0-010a0100	0000010001	030020-000	01110a1a00	0000010102	1002120020	00000000
Sideroxylon foetidissimum	00a0c01101	0-010a0100	0000000000	0d003c1000	01110c1000	0000002000	b1022200c0	00000000
Sideroxylon horridum	0001210?01	a011000110	00000a0001	0000?21000	0111001000	000001210a	1000120120	00000010
Sideroxylon inerme	00102a1101	0-010a01a0	0000010000	0300c0-000	0b110a1a00	0000010102	1002120a20	000a0000
Sideroxylon lanuginosum	0001c01101	1a11001100	00a0020000	0d00021000	01111a1100	0000010102	b000120120	000a0000
Sideroxylon marmulano	0000211101	0-01011100	0000000000	000010-000	02111a1000	0000010102	1102220120	000a0030
Sideroxylon mascatense	0001210?01	1111001110	0000010000	000010-000	0211001100	0000020002	1002121121	00000000
Sideroxylon reclinatum	0001ca0101	aa11001100	0000000000	0100121000	01111a1100	000001010c	1100120120	00000000
Sideroxylon salicifolium	0000211201	0-010a0100	0000010001	0000a21000	0111001000	0000012100	1002220120	000000aa
Sideroxylon saxorum	00102a1101	0-01010100	000000000a	0310?0-000	0211011100	0000010102	100??20020	00?00000
Synsepalum afzelii	0000000100	aa010a0110	00a0010001	000020-000	0b00111100	000a002002	11100b0021	00000011
Synsepalum brevipes	0010000a00	11010a01a0	0000010000	0?10i0-000	0200111100	000a003002	11100b0021	00100010
Synsepalum dulcificum	0010200101	11010c0110	0001220001	000010-000	0000111100	0000012002	1110011120	00100010
Synsepalum fleuryanum	0010200100	a1010c0110	0001210001	000010-000	0100011100	0000012002	10100b0120	00?00010

Appendix B
Continued

Taxon	Character number							
	0000000001 1234567890	1111111112 1234567890	2222222223 1234567890	3333333334 1234567890	4444444445 1234567890	5555555556 1234567890	6666666667 1234567890	77777777 12345678
Synsepalum passargei	0000200000	0-010c0100	0000010000	0000b0-000	0100111100	000002c002	10100b1120	00100010
Tieghemella heckelii	0010211101	0-01000100	0111201101	0000010000	0111021100	0000030012	2000020021	00?a0011
Trouettia sp	0010200101	1001020112	0000a10031	100010-000	1201111000	0001-----2	1110000020	00???0??
Vitellariopsis cuneata	0010201100	0-01030100	0110221100	1?00310000	0001121111	0200010112	2010011121	00???0??
Vitellariopsis dispar	0010211100	a001030100	0110211100	??00d10000	0001121?00	0200010112	2010011121	001a0011
Vitellariopsis kirkii	0010211100	a001030100	0110211100	1?00d10000	0001121100	0200010112	2010011121	001a0011
Xantolis cambodiana	000021?101	1-0100010?	0000010000	0010?0-000	011??21100	0200010012	1112221120	001a00b2
Xantolis siamensis	0000211101	0-01000100	0000010000	000000-000	0111121100	0200010002	1112221120	00?a00?b