Studies in Malesian Gentianaceae I: *Fagraea* sensu lato—complex genus or several genera? A molecular phylogenetic study

M. Sugumaran[1] and K.M. Wong[2]

[1]Rimba Ilmu Botanic Garden,
Institute of Biological Sciences, University of Malaya,
50603 Kuala Lumpur, Malaysia
maran@um.edu.my
[2]Singapore Botanic Gardens,
1 Cluny Road, Singapore 259569
wkm2000@gmail.com

ABSTRACT. Phylogenetic studies of *Fagraea* s.l. based on maximum parsimony and Bayesian analyses of gene sequences for the nuclear ITS region and a number of chloroplast regions (*trn*L intron, *trn*L–F spacer and two partial sequence regions of *ndh*F) were carried out. Separate experiments with an ingroup of 29 taxa of *Fagraea* s.l. (8 from section *Cyrtophyllum*, 16 from section *Fagraea* and 5 from section *Racemosae*; all new sequences) were made with individual gene-region and combined data sets; and with 43 taxa using only an ITS data set that included published gene sequences of other recently revised, well-established genera of the same tribe (Potalieae). Reasonably consistent clade composition was obtained with all analyses: two clades could be equated to sections *Fagraea* and *Racemosae*, another two (Elliptica and Gigantea clades) are different portions of the section *Cyrtophyllum*, and the solitary *F. crenulata* resolved basal to the Fagraea clade in the chloroplast gene analyses but was a distinct lineage in a polytomy with the Fagraea, Racemosa and Gigantea clades in the ITS analyses. The equivalence of these clades and the *F. crenulata* lineage to other monophyletic groups represented by established genera in the expanded-ITS analysis, as well as considerations of potential morphological synapomorphies for these individual entities, suggest that *Fagraea* s.l. is too morphologically and phylogenetically divergent to be considered a single genus.

Keywords. *Cyrtophyllum*, *Fagraea*, generic circumscription, Gentianaceae, Malesia, molecular phylogenetic analyses, morphology, Potalieae, *Racemosae*, synapomorphic characters

Introduction

Fagraea Thunb. and s.l. are paleotropical, with a distribution from Sri Lanka and India, through tropical South East Asia, reaching east to Polynesia (Struwe et al. 2002). The genus (or group) is centred in Malesia, where over 70 species are present (Struwe et al. 2002), and around 50 species are distributed within the Malay Peninsula and Borneo (Leenhouts 1962; Wong & Sugau 1996). Species of *Fagraea* s.l. represent a variety of life forms. They are tall canopy trees, smaller understory trees reaching only a few meters in height, shrubs, or are epiphytes and hemi-epiphytes. They occur from sea level to about 3000 m in very moist montane conditions, mostly in forest gaps, forest edges, rocky outcrops, along stream beds in wet tropical forests, and less commonly in

mesic forests, mangrove swamps and savannas (Motley 2004).

With a large diversity in habit and form, *Fagraea* s.l. includes species that are both conspicuous and ecologically important in natural landscapes. The widespread *F. fragrans* Roxb. is a common pioneer on sandy sites, a frequent secondary forest species in the lowlands, and can persist in mature forest. In the Malay Peninsula, *F. racemosa* Jack ex Wall. is also a common secondary forest species and the large-leafed *F. auriculata* Jack with long-tubed flowers is often conspicuous in coastal sandy sites and on quartz ridges in the lowlands. Throughout the lowland and lower montane forests of the Malay Peninsula and Borneo, the frequent presence of *Fagraea* epiphytes or hemi-epiphytes is detected by fallen corollas on the ground at different times during the year. Burkill (1936), Macmillan (1991), Motley (2004), Perry (1980), Quisumbing (1978) and Watson (1935) describe how a number of species are used mainly for wood, medicinal purposes and horticulture.

In the past, *Fagraea* s.l. was placed within the Loganiaceae (e.g., Leeuwenberg & Leenhouts 1980; Leenhouts 1962) but this classification is now controversial. Subsequent analyses of morphological data (Struwe et al. 1994; Struwe & Albert 1997) and molecular characteristics (Struwe et al. 2002; Downie & Palmer 1992; Olmstead et al. 1993) have demonstrated that *Fagraea* s.l. and its tribe Potalieae are better placed in the Gentianaceae. There is also support for this from some phytochemical evidence (Jensen 1992; Jensen & Schripsema 2002). Potalieae members have diverse flower merosity (e.g., 3-merous in the genus *Pycnosphaera* Gilg, up to 16-merous in the genus *Anthocleista* Afzel. ex R.Br.) and habit (trees, lianas, shrubs, scramblers and herbs) compared to most other tribes within Gentianaceae. The distribution of Potalieae is strictly tropical, mostly found around the equatorial regions (Struwe et al. 2002).

Molecular analyses have consistently shown that *Anthocleista*, *Fagraea* and *Potalia* Aublet form a monophyletic group that justifies inclusion in the same subtribe (Potaliinae) within Gentianaceae (Struwe et al. 2002; Molina & Struwe 2009). Whereas *Fagraea* s.l. is a large Indo-Pacific group, *Potalia* is found in the Central and South American region with nine recorded species and *Anthocleista* is only found in tropical Africa and Madagascar with 14 known species (Struwe et al. 2002). The only two clear morphological (non-synapomorphous) characters that make the subtribe Potaliinae aberrant in Gentianaceae are the occurrence of fleshy berries and occasional large tree habit in some taxa (Molina & Struwe 2009). Other members of the Gentianaceae have dry and capsular fruits and relatively smaller size; and there are trees in other gentian genera, but typically much less than 10 m high (Struwe et al. 2002).

Within *Fagraea*, three subgeneric groups have been recognised, considered as the sections *Cyrtophyllum* (Reinw.) Blume, *Fagraea* and *Racemosae* Benth. (Leenhouts 1962). This classification was adopted by Wong & Sugau (1996), who delimited the sections by inflorescence form and branching, seed form, the nature of fruit epidermis, axillary scale characters and stigma form.

In their review for Borneo, Wong & Sugau (1996) enumerated 42 species, including 20 that they newly described. This was a big contrast to the previous work by Leenhouts (1962, 1984) for Malesia, where he only enumerated 15 species for Borneo.

The species concepts of Leenhouts were considered too broad by Conn & Brown (1993) and Wong & Sugau (1996). An example stated by Wong & Sugau (1996) is that Leenhouts (1962) had only one species accepted in the whole section *Racemosae* for Malesia, having reduced many previously described species to synonymy. Wong & Sugau (1996) found many of such taxa morphologically distinct and possible to key out using both vegetative and flower and fruit characters, and resurrected many corresponding names from synonymy.

Scope of the present study

The bizarre contrast between the results obtained by Leenhouts (1962) and Wong & Sugau (1996), the demonstration of what appears to be likely impractical species concepts in Leenhouts (1962) by both Conn & Brown (1993) and Wong & Sugau (1996), and the lack of any recent detailed reviews for Malesia other than Borneo, precipitated the present inquiry. Just as species concepts were apparently too broad, so the inclusion of morphologically very different sections in the same genus was perhaps also questionable.

At the fundamental level, the distinctness of the subgeneric groups recognised as *Fagraea* sections *Cyrtophyllum*, *Fagraea* and *Racemosae* require further consideration. This is in view of somewhat conspicuous characters distinguishing them, as summarised by Wong & Sugau (1996), and very preliminary possibilities for generic distinction suggested by limited molecular evidence discussed by Struwe & Albert (1997). This problem is investigated using a molecular analysis of as many taxa as possible from the Malay Peninsula and Borneo, and augmented by molecular results for other taxa in the same subtribe and tribe from other geographical regions. The extents of morphological distinction among the so-called sections are also re-examined.

Finally, the merits and intricacies of classifying the several morphologically distinct groups of species within *Fagraea* s.l. as either sections or distinct genera will be discussed. From this, recommendations would be made for an appropriate classification of the group.

Sectional classification of *Fagraea* s.l.

The recognition of infrageneric groups within *Fagraea* has a convoluted history. Blume (1838) erected *Fagraea* section *Cyrtophyllum*, which included the two species *Cyrtophyllum peregrinum* Reinw. (this is in fact a synonym of the earlier published and better known *F. fragrans*) and *C. speciosum* Blume (synonymous with *Picrophloeus javanensis* Blume, to which also some authors have applied the name *F. elliptica* Roxb.). Blume had evidently done this because the many-branched cymes with much smaller flowers in these taxa were rather different from the sparsely branched cymes with larger flowers of *F. ceilanica* Thunb., the type species of the genus.

Later, Blume (1850) included two of his species, *F. kimangu* Blume and *F. picrophloea* Blume (both synonyms of his *P. javanensis*) in his *Fagraea* section *Eufagraea*, with a number of other species that had a very different inflorescence form, where cyme-like clusters of flowers were borne along an elongate main inflorescence axis that superficially resembled a complex raceme (described by Blume as "*cymis in racemum terminalem longissimum*"). In the same paper, Blume (1850) placed *F. ceilanica* (spelled as *F. zeylanica*) and associated species in "section *Fagraea verae*".

Bentham (1856) appears to be the first person who clearly differentiated three groups within *Fagraea*, as section *Parviflorae* Benth. (= Blume's section *Cyrtophyllum*), section *Racemosae* (which is that group with raceme-like inflorescences in Blume's section *Eufagraea*), and section *Corymbosae* Benth. (= Blume's section *Fagraea verae*). Subsequent authors also provided infrageneric names. Miquel (1857) employed subgenus *Cyrtophyllum* and subgenus *Eufagraea*, in place of Blume's sections *Cyrtophyllum* and *Fagraea verae*, respectively. Solereder (1892), on the other hand, named section *Pseudoracemosae* and section *Pseudocorymbosae* for Bentham's *Racemosae* and *Corymbosae*, respectively.

In his revision for the Flora Malesiana, Leenhouts (1962) recognised three infrageneric groupings that he called sections *Fagraea* (containing the type and by far the largest number of species), *Cyrtophyllum* (as designated by Blume) and *Racemosae* (as proposed by Bentham). Leenhouts (1962) gave very few characters for sectional distinction, some of which were somewhat inconsistent (see Table 1).

Wong & Sugau (1996) also used these sections as circumscribed by Leenhouts (1962) for their account of *Fagraea* in Borneo, but added more distinguishing characters and provided a key to sections. They used such characters as fruit size, ease of epidermis detachment on drying, seed form, inflorescence branching, characters of the petiole base (including whether scale-like structures develop at the leaf axils), stigma form and the extent of stamen and style exsertness in open flowers.

Alternative taxonomic interpretations

Several species placed in *Fagraea* s.l. by Leenhouts (1962) and other authors have been the basis of other generic names. Although Cammerloher (1923) considered *Cyrtophyllum* a synonym of *Fagraea*, Ridley (1923) had distinguished the two genera, using the former name for *C. lanceolatum* (Wall.) DC., *C. peregrinum* and *C. giganteum* (Ridl.) Ridl. The equivalents for these names in *Fagraea* are *F. lanceolata* Wall. (a synonym of *F. wallichiana* Benth.), *F. peregrina* Blume (a synonym of *F. fragrans*), and *F. gigantea* Ridl., respectively.

Leenhouts (1962) suggested that *Cyrtophyllum* and *Picrophloeus* Blume are synonyms of *Fagraea* section *Cyrtophyllum*. He also placed the names *Utania* G.Don, *Kuhlia* Reinw. and *Kentia* Steud. (the latter two illegitimate as a later homonym and a name lacking a description, respectively) as synonyms of *Fagraea* section *Racemosae*.

Table 1. Diagnostic characteristics of three sections in *Fagraea* s.l. according to Leenhouts (1962).

Character / Section	sect. *Fagraea*	sect. *Cyrtophyllum*	sect. *Racemosae*
leaves	auriculate or not	not auriculate	not auriculate
stipules	*(not explicitly stated)*	*(not explicitly stated)*	connate in an ochrea
inflorescence form	corymbose, dichasial, glomerulous or with solitary flowers; "usually with a pair of strong branches in the upper leaf axils, therefore the inflorescence as a whole mostly sessile"	corymbose	racemiform, with a number of decussate pairs of small cymes
inflorescence position	always terminal	(*not explicitly stated, but* including terminal and axillary types in different taxa)	always terminal
flower size	*(not explicitly stated)*	small	*(not explicitly stated)*
corolla form	*(not explicitly stated)*	tubular	*(not explicitly stated)*
stamens, style	hardly or not exserted	far exserted (except in *F. umbelliflora*)	not or only slightly exserted
fruit shape, size	*(not explicitly stated)*	globular, small	*(not explicitly stated)*

Materials and methods

Herbarium materials and studies. Study specimens were loaned from, or examined at, the following herbaria: Royal Botanic Gardens, Kew (K), Forest Research Institute Malaysia, Kepong (KEP), Nationaal Herbarium Nederland (L), Forest Research Centre, Sabah (SAN), Singapore Botanic Gardens (SING) and the University of Malaya (KLU). A handlens (×20 magnification) and binocular microscope (×40 magnification) were used to examine material.

Field collections and processing. Additional material collected yielded specimens that were oven-dried at 55ºC for 2–4 weeks; and flowers, fruits and leafy branches preserved in AWG solution (70% alcohol + 28% distilled water + 2% glycerol) for

study. Leaf material of 30 taxa of *Fagraea* collected from various locations were preserved in silica gel for subsequent DNA extraction (Table 2).

DNA extraction, gene regions and primers. Total DNA extraction from silica-dried or fresh leaf tissue samples followed the CTAB (cetyltrimethylammonium bromide) method of Doyle & Doyle (1987), or made use of *DNeasy* plant kits (Qiagen) following the manufacturer's protocol (Qiagen 2003–2009). Gene regions chosen for the phylogenetic analyses were: ITS (Internal Transcribed Spacer) (nuclear) and the *trn*L intron, *trn*L–F spacer and *ndh*F region (chloroplast). These regions have been widely used to study phylogenetic relationships at the family level and lower (Chassot et al. 2001; Davis et al. 2001; Gielly & Taberlet 1996; Ranker et al. 2003; Struwe et al. 2002; Thiv et al. 1999; Hagen & Kadereit 2001; Yuan & Küpfer 1995, 1997; Yuan et al. 1996, 2003). The whole ITS region (ITS1, 5.8S rDNA & ITS2) was amplified with universal primers: ITS 1 (forward) and ITS 4 (reverse), following White et al. (1990). For the *trn*L intron and *trn*L–F regions, universal primer sets (C & D and E & F, respectively) given by Taberlet et al. (1991) were used. Two new primer sets for the *ndh*F gene sequences were constructed at conserved regions following alignment of *ndh*F sequences for three taxa obtained from the GenBank: *Fagraea* sp. (AJ 235830), *Anthocleista grandiflora* (AJ 235829) and *Potalia resinifera* (AJ 235831). Each of these two sets of primers (GB1 Fwd & GB1 Rev and GB2 Fwd & GB2 Rev) was estimated to amplify about 900–1000 bp of the *ndh*F gene. Sequences of all primers used in this study are listed in Appendix A.

Polymerase Chain Reaction (PCR). PCR for all the regions were performed in a Whatman Biometra T Gradient or Perkin Elmer GeneAmp PCR System 9600 thermocycler. A total reaction mix of 50 µl was used, containing 25 µl of PCR master mix (2x Go Taq® Green Master Mix), 2 µl (50 mM) each of forward and reverse primer, 1–3 µl of DNA template and RNAse free water topped up to a final volume of 50 µl. Parameters for PCR amplification were: 1 cycle of 3 min at 94 ºC, linked to 30 cycles of 10 s at 94ºC, 20 s at 55ºC, 90 s at 72ºC, followed by 4 min at 72ºC to complete primer extension. PCR-amplified samples were electrophoresed on agarose gel and samples that had a clear single desired band were selected for sequencing. Purification of PCR products used the QIAquick PCR purification kit (Qiagen) following the manufacturer's protocol. When multiple bands of different sizes were observed in some cases, a gradient PCR (with different annealing temperatures) was carried out to find the optimal temperature that amplified a single desired band. The sample with such band was then selected for purification and sequencing. Occasionally, samples still produced multiple bands after PCR despite optimising the annealing temperature. For these, the PCR product was electrophoresed on agarose gel (1.2–1.5 %) for about 60 to 90 min. The target band was then excised and purified using a Qiaquick Gel extraction Kit (Qiagen) following the manufacture's protocol. This product was then sequenced.

Table 2. Voucher specimens and GenBank accession numbers of sequences for *Fagraea* s.l. taxa representing the different "sections" collected for the present study.
Fagraea sp. A, *F.* sp. B and *F.* sp. C are new Malay Peninsula taxa diagnosed in the present study, whereas two other unidentified Bornean taxa were labeled as *F.* sp 1 and *F.* sp 2. *Fagraea elliptica* is represented by two accessions from Peninsular Malaysia and Borneo. Sections: C (*Cyrtophyllum*), F (*Fagraea*), R (*Racemosae*). Vouchers with 'SAN' numbers were deposited in the Forest Research Centre Sandakan, Sabah (SAN); all other vouchers were deposited in the University of Malaya herbarium (KLU). The two partial non-overlapping *ndh*F sequences are represented with two GenBank accession numbers.

Taxon	Section	Voucher specimen	GenBank accessions for (ITS) – (*trn*L–F) – (*ndh*F: 2 partial sequences)
F. belukar	C	Postar & Ahmad SAN 147987	JX283355–JX217749–JX283385 & JX283414
F. caudata	C	Low LYW 213	JX283356–JX217750–JX283386 & JX283415
F. collina	C	Low LYW 260	JX283357–JX217751–JX283387 & JX283416
F. elliptica 1	C	Low LYW 358	JX283358–JX217752–JX283388 & JX283417
F. elliptica 2	C	Postar et al. SAN 147993	JX283359–JX217753–JX283389 & JX283418
F. fragrans	C	Sugumaran SM 212	JX283360–JX217754–JX283390 & JX283419
F. gigantea	C	Sugumaran SM 193	JX283361–JX217755–JX283391 & JX283420
F. wallichiana	C	Low LYW 206	JX283362–JX217756–JX283392 & JX283421
F. auriculata	F	Sugumaran SM 240	JX283363–JX217757–JX283393 & JX283422
F. carnosa	F	Lee DLKP 30	JX283364–JX217758–JX283394 & JX283423
F. crassifolia	F	Low LYW 244	JX283365–JX217759–JX283395 & JX283424
F. crenulata	F	Sugumaran SM 246	JX283366–JX217760–JX283396 & JX283425
F. curtisii	F	Low LYW243	JX283367–JX217761–JX283397 & JX283426
F. gardenioides	F	Sugumaran SM 170	JX283368–JX217762–JX283398 & JX283427
F. imperialis	F	Sugumaran SM 238	JX283369–JX217763–JX283399 & JX283428
F. oblonga	F	Sugumaran SM 165	JX283370–JX217764–JX283400 & JX283429
F. renae	F	Sugumaran SM 177	JX283371–JX217765–JX283401 & JX283430
F. resinosa	F	Postar et al. SAN 147998	JX283372–JX217766–JX283402 & JX283431
F. ridleyi	F	Low LYW 227	JX283373–JX217767–JX283403 & JX283432
F. splendens	F	Zahid ZMS 42	JX283374–JX217768–JX283404 & JX283433
F. stonei	F	Wong et al. SAN 147989	JX283375–JX217769–JX283405 & JX283434
F. sp. A	F	Low LYW 138	JX283376–JX217770–JX283406 & JX283435
F. sp. B	F	Sugumaran SM164	JX283377–JX217771–JX283407 & JX283436
F. sp. 1	F	Postar et al. SAN 149702	JX283378–JX217772–JX283408 & JX283437
F. cuspidata	R	Seligi & Lingkong SAN 145303	JX283379–JX217773–JX283409 & JX283438
F. racemosa	R	Sugumaran SM 248	JX283380–JX217774–JX283410 & JX283439
F. spicata	R	Postar & Ahmad SAN 147985	JX283381–JX217775–JX283411 & JX283440
F. volubilis	R	Sugumaran SM 206	JX283382
F. sp. C	R	Sugumaran SM 201	JX283383–JX217777–JX283412 & JX283441
F. sp. 2	R	Postar SAN 149705	JX283384–JX217778–JX283413 & JX283442

Sequencing PCR products and data authentication. PCR products were sequenced on an Applied Biosystems 3730xl DNA Analyser with BigDye®Terminator ver. 3.1 Sequencing Kit with: 25 cycles of 96ºC for 10 s, 50ºC for 5 s and 60ºC for 4 min, at rapid thermal ramp for 1ºC/s. Sequencing in both directions used the same primers as for PCR. All sequence data were manually checked by eye with the corresponding electrophoregrams. Ambiguous base pairs (those with multiple peaks) were cross-checked with the complementary sequence data (i.e., forward sequence was cross-checked with reverse sequence and vice versa). This could be done because forward and reverse strands that were sequenced had an approximate overlap of about 80%. Samples that had very noisy sequence data were not used in the analyses.

Sequence alignment. Sequence data were initially aligned using ClustalX 2.0.10 (Larkin et al. 2007). Aligned sequences were then manually adjusted using the software Bioedit (Hall 1999). The boundaries of the ITS, *trn*L intron, *trn*L–F spacer and *ndh*F were determined by comparison with the published outgroup sequence, *Anthocleista grandiflora,* that was also used in the analyses (Table 3). The *trn*L intron and *trn*L–F spacer sequences were aligned and combined into a single data matrix (hereafter referred to as the "*trn*L–F data set"). Numerous single and multibase insertions or deletions (indels) were introduced for the ITS and *trn*L–F data sets but no indels were needed for aligning the *ndh*F data set. The aligned data was then saved in a Nexus file format and phylogenetic analyses were performed using PAUP* version 4.0b4a for Macintosh (Swofford 2001) and MrBayes v.3.1 (Huelsenbeck & Ronquist 2001).

Phylogenetic assessments: scope and experimental design. Two phylogenetic assessments were conducted. The first assessment included analyses with individual data sets (ITS, *trn*L–F or *ndh*F) and a combined data set (ITS + *trn*L–F + *ndh*F). This assessment aims to evaluate if (a) reasonable support can be found for any monophyletic groups forming within the ingroup; (b) if such clades correspond to the sections of *Fagraea* s.l. circumscribed by existing classification; and (c) if these clades are consistent (or congruent) among the topologies representing the different gene regions. The combined data analysis was performed to evaluate whether clades were better resolved in comparison with single-gene analyses. This assessment was performed with 29 taxa of *Fagraea* s.l. (8 from section *Cyrtophyllum*, 16 from section *Fagraea* and 5 from section *Racemosae*) and one taxon as outgroup (*A. grandiflora*). The 29 DNA sequences of *Fagraea* s.l. from all four gene regions included in these analyses are new (details in Table 2). The sequences for *A. grandiflora* were obtained from GenBank (Table 3).

The second assessment was done with only the ITS data set where sampling was expanded to 45 taxa. This assessment was done to evaluate whether clades formed among so-called *Fagraea* species within the ingroup can be reasonably circumscribed as distinct genera. The selected ingroup taxa are all from Potalieae and include the well recognised genera *Anthocleista* and *Potalia* (in the same subtribe, Potaliinae, as *Fagraea*) as well as *Lisianthius* (subtribe Lisianthiinae) (Weaver 1972; Sytsma 1988). The two outgroup taxa belong to two other distantly related tribes within the

Table 3. List of species for which GenBank accessions were included for this study, including tribal and subtribal classification in Gentianaceae (Struwe et al. 2002) and voucher information. Herbarium acronyms: *E* – Royal Botanic Garden Edinburgh, Edinburgh, Scotland, UK; *HNWP* – Northwest Plateau Institute of Biology, Xining, Qinghai, China; *F* – Field Museum of Natural History, Chicago, Illinois, USA; *NEU* – Université de Neuchâtel, Neuchâtel, Switzerland; *NY* – New York Botanic Garden, Bronx, New York, USA; *UPS* – Uppsala University, Uppsala, Sweden; *US* – Smithsonian Institutution, Washington, District of Columbia, USA.

Taxa	Current tribe–subtribe	Voucher (*Herbarium*)	GenBank accession (and gene region)
Anthocleista amplexicaulis	Potalia–*Potaliinae*	Woulhauser PBZT (*NEU*)	AJ489863 (ITS)
Anthocleista grandiflora	Potalia–*Potaliinae*	Callmander s.n. (*NEU*) Callmander s.n. (*NEU*) Callmander s.n. (*NEU*) Bremer 3098 (*UPS*)	AJ489864 (ITS) AJ490190 (*trn*L (UAA) intron) AY251777 (*trn*L-F spacer) AJ235829 (*ndh*F)
Exacum affine	Exaceae	Miller et al. 6201 (*E*)	AJ489879 (ITS)
Fagraea berteroana	Potalia–*Potaliinae*	L. Struwe 1219 (*NY*)	DQ449918 (ITS)
F. ceilanica	Potalia–*Potaliinae*	L. Struwe 1300 (*NY*)	FJ23257 (ITS)
F. elliptica	Potalia–*Potaliinae*	Takeuchi 7122 (*NY*)	FJ232579 (ITS)
F. macroscypha	Potalia–*Potaliinae*	Beaman et al. 8867 (*US*)	FJ232573 (ITS)
F. salticola	Potalia–*Potaliinae*	Pullen 326 (*US*)	FJ232571 (ITS)
Gentiana algida	Gentianeae–*Gentianinae*	Liu 1257 (*HNWP*)	DQ398659 (ITS)
Lisianthius brevidentatus	Potalia–*Lisianthiinae*	Ortiz 1664 (*F*)	FJ32569 (ITS)
Lisianthius cuspidatus	Potalia–*Lisianthiinae*	Lewis 895 (*F*)	FJ32567 (ITS)
Lisianthius laxiflorus	Potalia–*Lisianthiinae*	Struwe & Specht 1153 (*NY*)	FJ232552 (ITS)
Potalia amara	Potalia–*Potaliinae*	S. Mori 24123 (*NY*)	DQ449919 (ITS)
Potalia elegans	Potalia–*Potaliinae*	P. Berry 7434 (*NY*)	DQ449920 (ITS)
Potalia resinifera	Potalia–*Potaliinae*	B. Stahl 1872 (*NY*)	DQ449921 (ITS)

Gentianaceae, namely, Exaceae (*Exacum affine*) and Gentianeae (*Gentiana algida*) (Struwe et al. 2002). The ITS region was chosen because it was relatively more informative than the other two data sets; moreover, many more ITS sequences were available in GenBank compared to *trn*L–F (intron + spacer) and *ndh*F sequences. The ingroup included 35 *Fagraea* taxa, two *Anthocleista* taxa, three *Lisianthius* taxa and three *Potalia* taxa. Of the *Fagraea* sequences used in this assessment, 30 are new

sequences, including *F. volubilis* (which was excluded from the first assessment as sequencing results for the *ndh*F region was very poor) (Table 2). All other sequences, including five sequences of *Fagraea* and the outgroup sequences, were obtained from GenBank (Table 3).

Parsimony analyses in PAUP* were performed with heuristic searches where all characters were unordered and unweighted ('Fitch parsimony'; Fitch 1971). All gaps in the sequence were treated as missing data. Starting trees were obtained using simple stepwise addition sequences, with one tree held at each step with tree bisection–reconnection (TBR) branch swapping algorithm, MULTREES option in effect, accelerated transformation (ACCTRAN), branches with zero length collapsed and topological constraints not enforced. Separate and combined analyses of the three data sets (ITS, *trn*L–F and *ndh*F) were performed and strict consensus trees were generated. The bootstrap method (Felsenstein 1985) was used to estimate robustness of the various clades revealed in the consensus tree. Bootstrap values were estimated from 1000 replicates of full-heuristic searches using simple addition sequence and TBR branch swapping with a set 'MAXTREES' limit of 10000 trees per bootstrap replicate. Branches less than 50% value were collapsed. The consistency index (CI) (Kluge & Farris 1969), the retention index (RI) and rescaled consistency index (RC) (Farris 1989) were also calculated using PAUP* as measures of character fit to the phylogenetic trees.

For Bayesian inference analyses, an appropriate evolutionary model was selected using MrModeltest 2.2 (Nylander 2004) together with PAUP* (Swofford 2001). The programme (MrModeltest 2.2) specifically tests the 24 models available common to PAUP* and MrBayes (Huelsenbeck & Ronquist 2001) with a given data set. The Akaike Information Criterion tests selected the General Time Reversible model with gamma distribution of rates for all individual data sets (ITS, *trn*L–F, *ndh*F and ITS-expanded). Two independent runs of 1.5 million generations were performed each with four MCMC (Markov Chain Monte Carlo) chains. One tree was sampled every 500 generations and the first 750 trees (burn-in=750) were excluded from the analyses which amounts to 25% of trees sampled. A 95% credible set of trees was generated by including all trees with the highest posterior probabilities until the cumulative posterior probabilities (PP) was 95% (Huelsenback et al. 2001). A 50% majority rule consensus of these generated trees was used to estimate the PP of each clade. PP values above p=0.95 were considered to be statistically significant (Huelsenback & Ronquist 2001; Larget & Simon 1999; Lewis 2001; Rannala & Yang 1996; Kauff & Lutzoni 2002).

The incongruence length difference test (ILD; Farris et al. 1994, 1995) as the partition homogeneity test (PHT) implemented in PAUP* was employed to test the null hypothesis that the three data sets (ITS, *trn*L–F and *ndh*F) were homogeneous with respect to phylogenetic information. PHT was performed with 100 replicates of heuristic searches, Maxtree=1000 and TBR branch swapping. It has been suggested that PHT p values more than 0.01 indicate that the data sets are congruent and if combined will either improve or will not affect the accuracy of the phylogenetic information (Cunningham et al. 1998).

Results

Sequence variation

The aligned data matrix of the ITS data set with 30 taxa had 639 characters. Out of these, 89 characters were variable but parsimoniously uninformative whereas 64 characters (10% of total characters) were phylogenetically informative. The unaligned length of ITS sequences among the 30 taxa of *Fagraea* (Table 2) used in this study varied from 622 to 631 bp. These data resulted in uncorrected pairwise sequence divergences ranging from 0 (*F*. sp 1 vs. *F*. *gardenioides*, *F. curtisii* vs. *F. crassifolia*) to 13.9% (*F. oblonga* vs. *F*. sp. 2).

The individual *trn*L–F sequences (*trn*L intron + *trn*L–F spacer) in *Fagraea* used in this study ranged from 652 to 836 bp. The length varied mainly at several A–T rich regions where alignment could not be readily done due to repeated motifs. These highly variable regions which accounted for 156 bp were excluded from the analyses. The final aligned *trn*L–F data matrix had a total of 811 characters where 60 characters were variable and parsimoniously uninformative, and 49 characters (6% of total characters) were phylogenetically informative. These data resulted in uncorrected pairwise sequence divergences ranging from 0 (*F. auriculata* vs. *F. imperialis*, *F. gigantea* vs. *F. caudata*, *F. wallichiana* vs. *F. caudata*, *F. wallichiana* vs. *F. gigantea*, *F. spicata* vs. *F. cuspidata*, *F.* sp. C vs. *F. cuspidata* and *F. spicata* vs. *F.* sp. C) to 3.5% (*F. stonei* vs. *F. fragrans*).

The two new primer pairs that were used to amplify two partial sequence regions of *ndh*F produced 1832 bp in total. The primer pair "GB1" produced 899 bp whereas the "GB2" pair produced 933 bp. Based on the *ndh*F sequence of *A. grandiflora* (GenBank Acc. AJ235829), the 899 bp sequence data corresponds with positions 161 to 1059 whereas the 933 bp sequence data corresponds with positions 1182 to 2114. The *ndh*F sequences were easily aligned and the total 1832 characters had 114 characters that were variable and parsimoniously uninformative whereas 81 characters (4.4% of total characters) were informative. These data resulted in uncorrected pairwise sequence divergences ranging from 0 (*F. racemosa* vs. *F.* sp. C, *F. cuspidata* vs. *F.* sp. C, *F*. sp. 2 vs. *F.* sp. C, *F. splendens* and *F.* sp. B and *F. imperialis* vs. *F. auriculata*) to 2% (*F. resinosa* vs. *F.* sp. C, *F. resinosa* vs. *F. racemosa*, *F. resinosa* vs. *F. spicata*, *F. resinosa* vs. *F*. sp. 2).

The combined data set of ITS, *trn*L–F and *ndh*F resulted in 3282 characters. The data matrix had 263 variable and parsimoniously uninformative characters whereas 194 characters (5.9% of total characters) were informative.

The aligned matrix of the expanded ITS data set (45 taxa) had 697 characters where 131 characters were variable and parsimoniously uninformative, and 185 characters (26.5% of total characters) were phylogenetically informative. These data resulted in uncorrected pairwise sequence divergences ranging from 0 (*F. ridleyi* vs. *F. gardenioides*, *F.* sp. 1 vs. *F. gardenioides*, *F.* sp. 1 vs. *F. ridleyi*, *F. curtisii* vs. *F. crassifolia* and *F. macroscypha* vs. *F. auriculata*) to 14% (*F*. sp. 2 vs. *F. imperialis*).

Phylogenetic analyses

Data set characteristics are presented in Table 4. In all the analyses, parsimony and Bayesian tree topologies were generally congruent and bootstrap (BS)-supported nodes mostly also had high support of posterior probabilities (PP). The trees presented were obtained from maximum parsimony analyses and congruent branch support values obtained from the Bayesian analyses (PP) are stated.

Analyses of the ITS data set

Parsimony analyses of the 30-taxon ITS data set resulted in 3520 trees with a tree length of 211 (strict consensus shown in Fig. 1), CI (Consistency Index) of 0.8152 and RI (Retention Index) of 0.8691. Four monophyletic groups can be recognised, referred to as the Fagraea, Racemosa, Gigantea and Elliptica clades. In these analyses, all included taxa representing *Fagraea* section *Fagraea* were represented on the Fagraea clade except *F. crenulata*, which is placed in a tetrachotomy with three other clades. Within the Fagraea clade, two monophyletic sister groups are formed where one group is smaller, consisting of three taxa, and the rest form a bigger group with 12 taxa. The 3-taxon group is well-supported (BS = 100%; PP = 1.00) with *F. resinosa* sister to the branch with *F. auriculata* and *F. imperialis*. Within the 12-taxon group, there are two smaller well-supported subgroups (*F. crassifolia* and *F. curtisii*; *F. splendens* and *F.* sp. B) as well as other species whose relationships are unresolved.

Species included in the study representing section *Racemosae* (Racemosa clade) form a strongly supported monophyletic group (BS = 100%; PP = 1.00). Within the Racemosa clade, *F.* sp 2 is sister to an equivocal clade of only moderate support (BS = 69%) with four unresolved taxa. Section *Cyrtophyllum* is paraphyletic with the eight representative taxa split into two distinct monophyletic groups. The groups comprising four taxa each are referred to as the Elliptica and Gigantea clades. The Elliptica clade is sister to the rest of *Fagraea* s.l. and well-supported with BS (91%) and moderately supported with PP (0.82) values. Within the Elliptica clade, *F. belukar* is sister to a clade with moderate BS (65%) but strong PP (0.99) support containing the three other taxa. The Gigantea clade is well supported (BS = 81%; PP = 1.00) and also well resolved where *F. fragrans* is the most basal taxon. The branches within this clade received moderate BS and good PP support.

Analyses of the trnL–F data set

The resulting 7224 most parsimonious trees with the *trn*L–F data set were 135 steps long (strict consensus shown in Fig. 2) with a CI of 0.8519 and RI of 0.9206. The clades formed are similar to those observed for the ITS tree; however, the interspecific relationships within each clade were less resolved with the *trn*L–F data set. Again, as in the ITS tree, *F. crenulata* was outside of the Fagraea clade. *Fagraea crenulata* is basal to both the Elliptica and Fagraea clades together and the relationship is well supported (BS = 82%; PP = 1.00). The rest of the species representing *F.* section *Fagraea* (the Fagraea clade) form a strongly supported clade (BS = 99%; PP = 1.00). Within the Fagraea clade, *F.* sp. A and *F. gardenioides* were unresolved, *F. crassifolia* and *F. curtisii* form a strongly supported group (BS = 100%; PP = 1.00) and the rest also form

Table 4. Characteristics of the parsimony-based analyses with individual and combined data sets.

Characteristics	ITS	*trn*L–F (*trn*L intron + *trn*L–F spacer)	*ndh*F (2 combined partial regions)	ITS + *trn*L–F + *ndh*F	ITS (expanded data set)
Number of taxa	30	30	30	30	45
Total characters	639	811	1832	3282	697
Constant characters	486	702	1637	2825	381
Parsimony informative characters	64	49	81	194	185
Variable characters	89	60	114	263	131
Most parsimonious trees	3520	7224	1	4	2145
Tree length	211	135	242	591	582
Consistency Index, CI *(values excluding uninformative sites in parentheses)*	0.82 *(0.66)*	0.85 *(0.73)*	0.83 *(0.68)*	0.83 *(0.68)*	0.73 *(0.65)*
Retention Index, RI	0.87	0.92	0.91	0.90	0.81
Rescaled Consistency Index, RC	0.71	0.78	0.76	0.74	0.59

a well-supported group (BS = 84%; PP = 1.00). In this latter group, *F. ridleyi* is placed in a trichotomy with two well-supported subclades. These subclades have three and seven taxa, respectively. The 3-taxon subclade comprising of *F. resinosa*, *F. imperialis* and *F. auriculata* were well resolved where *F. resinosa* was sister to the other two. In the 7-taxon subclade, *F.* sp. B and *F. splendens* formed a strongly supported group while the rest were unresolved.

Section *Cyrtophyllum* is paraphyletic, with four out of eight species represented forming the Elliptica clade and three other species forming the Gigantea clade. The Elliptica clade was basal to the Fagraea clade and the relationship was only moderately supported with BS (60%) but well supported with PP (0.99) values. The Elliptica clade itself was only weakly supported with 51% BS and a moderate PP support (0.93). The relationship within the Elliptica clade was unresolved.

The Gigantea and Racemosa clades plus *F. fragrans* formed a well-supported group (BS = 97%; PP = 1.00). The Gigantea clade received strong branch support (BS = 97%; PP = 1.00) but was internally unresolved. The remaining taxon, *F. fragrans*, was placed in a trichotomy with the Racemosa and Gigantea clades. The Racemosa clade included all the five representative taxa of section *Racemosae* and was monophyletic with moderate BS (67%) and good PP (0.96) support. The *F. cuspidata* and *F. spicata* subgroup within the Racemosa clade received moderate BS (66%) and good PP (0.99) support; the other three taxa were not resolved.

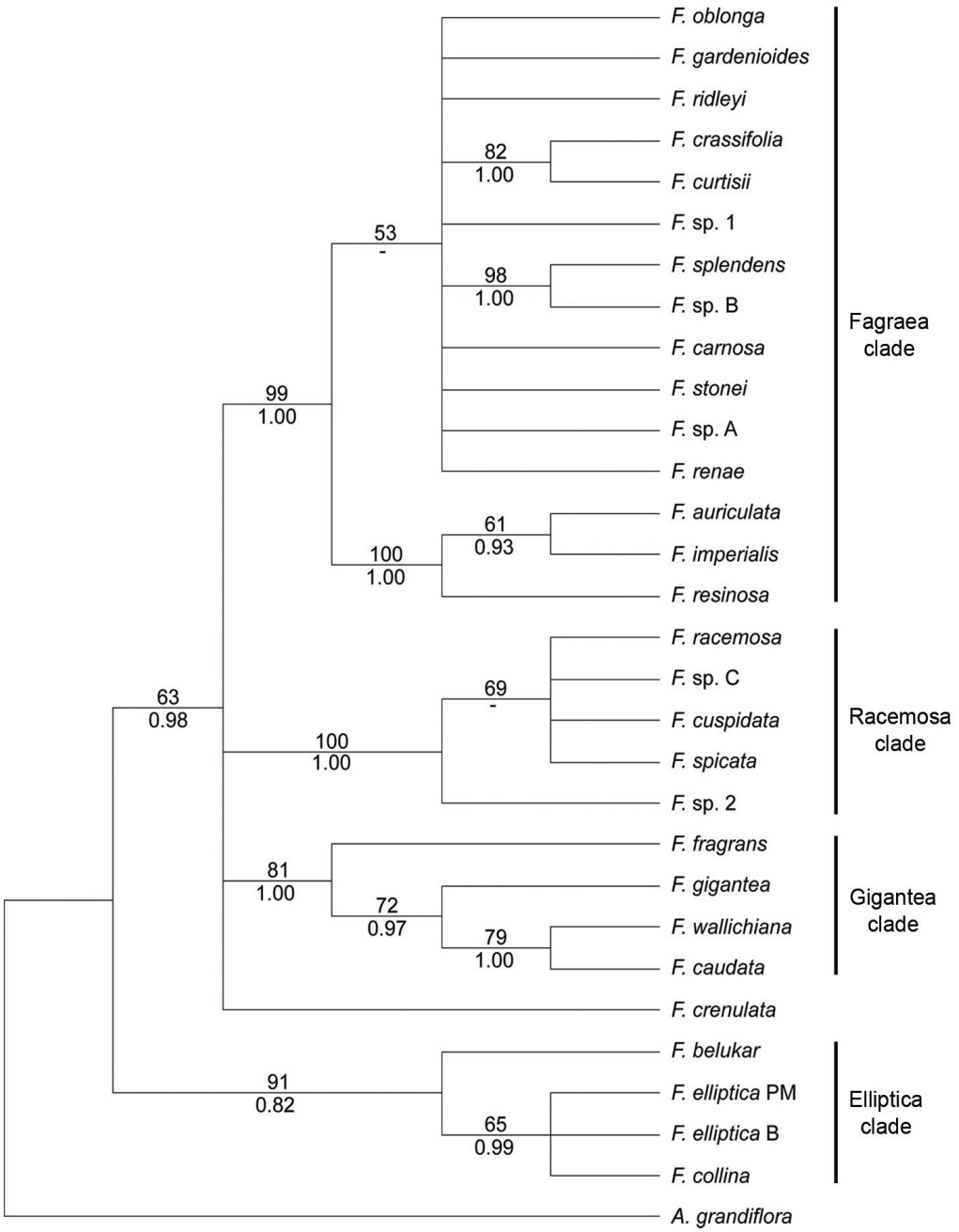

Fig. 1. Strict consensus of 3520 equally parsimonious trees based on the ITS sequence data. The numbers above and below the branches denote Bootstrap and Bayesian Posterior Probability values, respectively. Length (L) = 211; consistency index (CI) = 0.8152; retention index (RI) = 0.8691. *A*. = *Anthocleista*; *F.* = *Fagraea*. Different accessions of *F. elliptica* are indicated (PM = Peninsular Malaysia, B = Borneo).

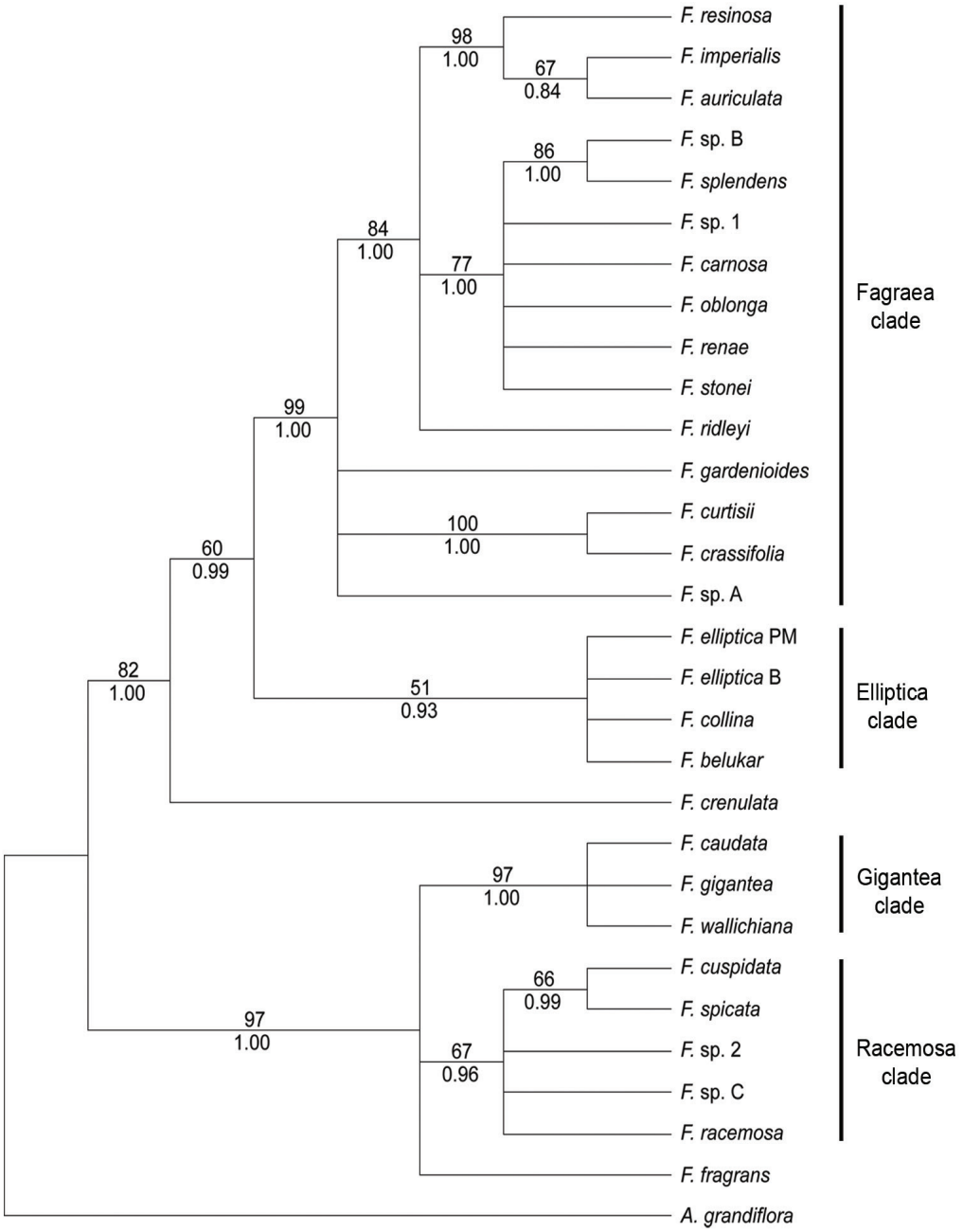

Fig. 2. Strict consensus of 7224 equally parsimonious trees based on the *trn*L–F sequence data. The numbers above and below the branches denote Bootstrap and Bayesian Posterior Probability values, respectively. Length (L) = 135; consistency index (CI) = 0.8519; retention index (RI) = 0.9206. *A*. = *Anthocleista*; *F.* = *Fagraea*. Different accessions of *F. elliptica* are indicated (PM = Peninsular Malaysia, B = Borneo).

Analyses of the ndhF data set

Parsimony analysis of the *ndh*F data set produced a single most parsimonious tree with tree length of 242 (Fig. 3), a CI of 0.8347 and RI of 0.9109. The four major clades common to the previous two analyses were also observed here and all had good BS and PP support. The *F.* section *Fagraea* (Fagraea clade + *F. crenulata*), was a well-supported monophyletic group, with *F. crenulata* basal within the clade. There were two well-supported sister groups, with four and eleven taxa, respectively. The 4-taxon group was well resolved with good branch support where *F. gardenioides* was basal. The 11-taxon group has two well supported subgroups: *F.* sp. B plus *F. splendens*; and a subgroup with *F. resinosa*, *F. imperialis* and *F. auriculata.* These groups were also observed in the analyses with the ITS and *trn*L–F data sets. The relationships among the other taxa in this clade were unresolved.

Section *Cyrtophyllum* was paraphyletic where four out of eight taxa formed the Elliptica clade, while a further three taxa grouped to form the Gigantea clade. The Elliptica clade was sister to the Fagraea clade, both forming a group with good branch support (BS = 75%; PP 0.99). The Elliptica clade itself was well supported (BS = 97%; PP = 1.00) and the interspecies relationships within the clade were well resolved with moderate branch support.

The Gigantea clade was well supported and was sister to the Racemosa clade. However, the interspecies relationships within the Gigantea clade were not resolved. The Racemosa clade was monophyletic and received good branch support (BS = 100%; PP = 1.00). Within the Racemosa clade, *F.* sp. C, *F. racemosa* and *F*. sp 2 formed a group with moderate BS (66%) but good PP (0.98) support. The remaining two species were unresolved. *Fagraea fragrans* resolved as a basal taxon for the Gigantea + Racemosa clades.

Analyses of the combined ITS, trnL–F and ndhF data sets

The result of the ILD test was not significant (p=0.03), indicating that the null hypothesis of data set homogeneity could not be rejected. The three data sets were thus combined into a single matrix and parsimony and Bayesian analyses were performed. Parsimony searches on the combined data set produced four trees having a tree length of 591 (strict consensus shown in Fig. 4) with a CI of 0.8274 and RI of 0.8979. The combined data set tree was most congruent to the tree resulting from the *ndh*F data set. *Fagraea* section *Fagraea* (Fagraea clade + *F. crenulata*) was monophyletic and strongly supported (BS = 96%; PP = 1.00) with *F. crenulata* as the basal taxon for the clade. *Fagraea crenulata* is then sister to two well-supported groups, one smaller comprising of four taxa and another bigger with eleven taxa. These groups were also observed in the *ndh*F tree but the internal resolutions were slightly better in this tree. The 4-taxon group was well resolved and received good branch support. The 11-taxon group was divided into two subclades with one unresolved taxon (*F. ridleyi*). The *F. resinosa*, *F. imperialis* and *F. auriculata* clade was well resolved with good branch support. *Fagraea resinosa* was sister to the latter two and this clade was also observed in all the individual data set analyses. The interspecies relationships among the taxa in the remaining group were not resolved except for two taxa, *F. splendens* and *F.* sp. B,

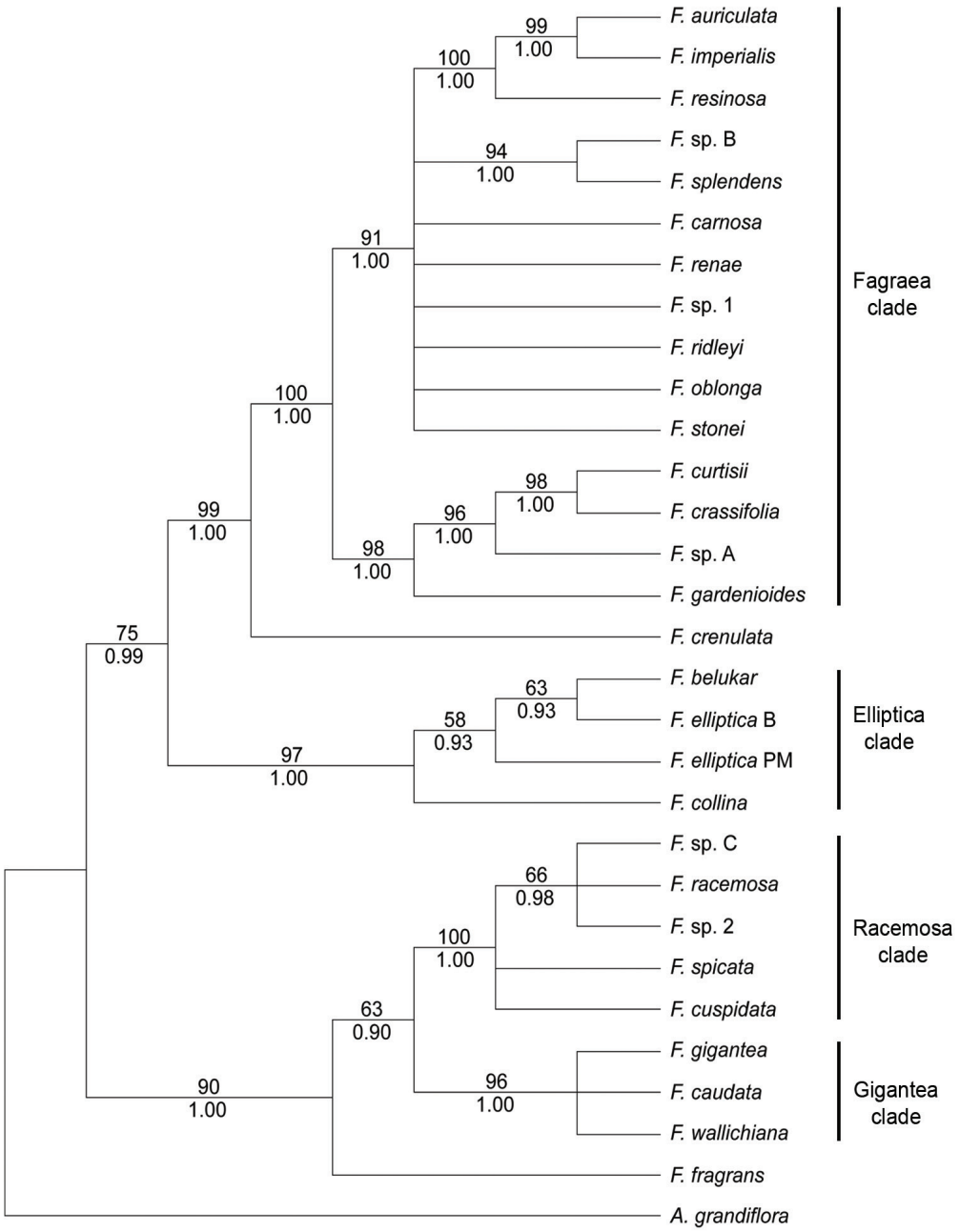

Fig. 3. Single most parsimonious tree based on the *ndh*F sequence data. The numbers above and below the branches denote Bootstrap and Bayesian Posterior Probability values, respectively. Length (L) = 242; consistency index (CI) = 0.8347; retention index (RI) = 0.9109. *A.* = *Anthocleista*; *F.* = *Fagraea*. Different accessions of *F. elliptica* are indicated (PM = Peninsular Malaysia, B = Borneo).

that formed a well-supported subclade. This subclade was also found in the all the individual data set analyses.

The section *Cyrtophyllum* was paraphyletic and resolved into two distinct clades (Elliptica and Gigantea) as also observed in all individual data set analyses. The Elliptica clade was sister to the Fagraea clade with good support (BS = 82%; PP = 0.95). The Elliptica clade was itself strongly supported (BS = 100%; PP = 1.00), wherein *F. collina* was sister to a clade with weak support (BS = 59%; PP = 0.69) that contained the remaining three taxa.

The Gigantea and Racemosa clades were sisters to each other with good branch support (BS = 99%; PP = 1.00). The Gigantea clade was internally well resolved, with *F. fragrans* basal to the remaining three taxa. This relationship received moderate BS support (65%) but good PP support (0.96). The Racemosa clade was well supported (BS = 100%; PP = 1.00), with *F.* sp 2 sister to a weakly formed clade (BS = 58%). Within this clade, *F. cuspidata* and *F. spicata* formed a moderately supported subclade (BS = 64%; PP = 0.94).

Analyses of the expanded ITS data set

Parsimony analyses of the expanded ITS data set produced a total of 2145 most parsimonious trees having tree length of 582 with a CI of 0.7337 and RI of 0.8089 (strict consensus shown in Fig. 5). As seen in the individual and combined data set results, *Fagraea* s.l. was divided into four major clades. *Fagraea crenulata* was placed in a tetrachotomy with the Fagraea, Racemosa and Gigantea clades. A similar pattern was also seen in the 30-taxon ITS tree (Fig. 1).

Within the Fagraea clade, *F. berteroana* and *F. salticola* form a well-supported subclade (BS = 92%; PP = 0.99). These two taxa are sister to a subclade with two other monophyletic groups, where one is larger with 13 taxa and the other is a smaller 4-taxon clade. The 13-taxon clade is moderately supported with BS = 67% but well supported in the Bayesian analyses (PP = 0.98). Within this 13-taxon clade there are two smaller subclades (*F. splendens* and *F.* sp. B; *F. crassifolia* and *F. curtisii*) with good BS and PP support, as well as several other species whose relationships are unresolved. The 4-taxon clade is well supported with both BS and PP values. Within this clade, *F. resinosa* is sister to a moderately supported (BS = 62%; PP = 0.91) subclade.

The Racemosa clade which represents all the taxa selected from section *Racemosae*, is monophyletic and received good support (BS = 100%; PP = 1.00). Within this clade, *F.* sp 2 was sister to the rest of the taxa in a weakly formed clade (BS = 71%).

As with all other individual data set analyses, section *Cyrtophyllum* is paraphyletic and was split into two distinct monophyletic groups, viz., the Elliptica and Gigantea clades. The Gigantea clade was moderately supported with BS = 78% but well-supported in the Bayesian analyses (PP = 1.00). The clade was internally well resolved with moderate BS and strong PP support. The widespread *F. fragrans* was placed as the basal taxon in the Gigantea clade.

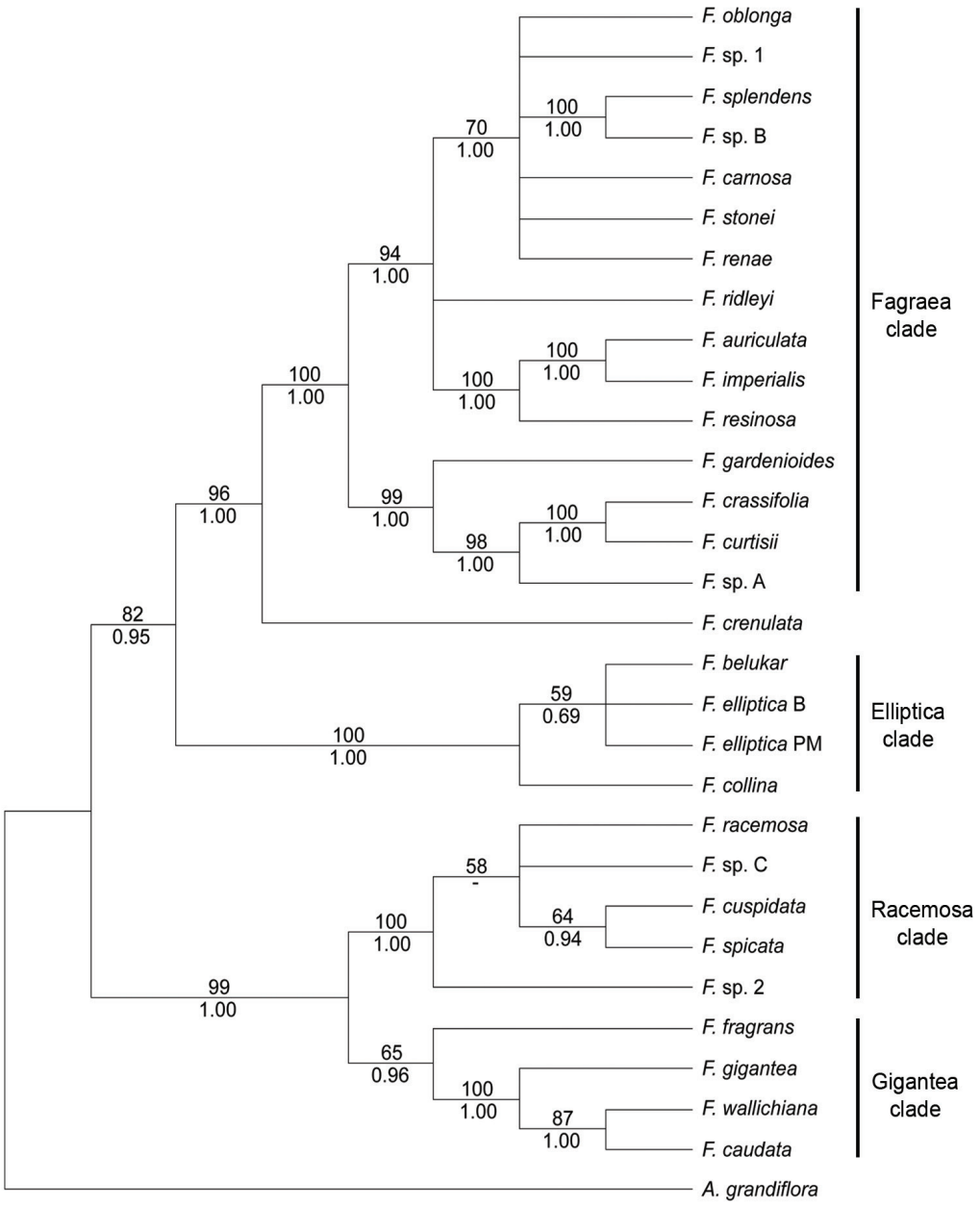

Fig. 4. Strict consensus of four equally parsimonious trees based on the combined ITS, *trn*L–F and *ndh*F sequence data. The numbers above and below the branches denote Bootstrap and Bayesian Posterior Probability values, respectively. Length (L) = 591; consistency index (CI) = 0.8274; retention index (RI) = 0.8979. *A*. = *Anthocleista*; *F.* = *Fagraea*. Different accessions of *F. elliptica* are indicated (PM = Peninsular Malaysia, B = Borneo).

The position of the Elliptica clade, sister to the rest of *Fagraea* s.l., was well supported (BS = 98%; PP = 0.99). The Elliptica clade itself received moderate branch support (BS = 77%; PP = 0.81) and within this clade, *F. belukar* was sister to the rest. The three accessions of *F. elliptica* from Peninsular Malaysia, Borneo and New Guinea along with *F. collina* were clustered together, receiving moderate supported with BS (72%) but good support with PP (0.99).

The other three genera represented in the ingroup, *Anthocleista*, *Potalia* and *Lisianthius*, were each monophyletic. *Anthocleista* and *Potalia* both formed a group sister to *Fagraea* s.l. with good support (BS = 97%; PP = 1.00), whereas *Lisianthius* was sister to *Anthocleista*, *Potalia* and *Fagraea* s.l. in turn, also with good support (BS = 90%; PP = 0.96).

Implications of the molecular phylogenetic analyses

Clade correspondence to named taxonomic sections

Regardless of which gene regions were employed in the study, including the individual and combined data set analyses, as well as the different methods used (MP and Bayesian), the results show *Fagraea* s.l. segregating into four reasonably well-supported monophyletic groups. Among these, the Fagraea clade and the Racemosa clade closely correspond to the sectional classification, i.e., *Fagraea* section *Fagraea* and *F.* section *Racemosae*, respectively.

The position of the bizarre *F. crenulata* is interesting. It was unresolved in a polytomy in both the ITS and the ITS–expanded data sets. It also failed to resolve with the Fagraea clade in the analyses of the *trn*L–F data set. Only the analyses with the *ndh*F and the combined data set supported the placement of *F. crenulata* as a basal taxon to the Fagraea clade.

Taxa sampled as *Fagraea* section *Racemosae* were shown to be monophyletic in all the analyses. The clade received strong BS and PP support in analyses with all the data sets except in the *trn*L-F data set where it received moderate support for BS (67%).

Fagraea section *Cyrtophyllum* as defined by Leenhouts (1962) was paraphyletic and resolved as two distinct clades, i.e., the Elliptica and Gigantea clades, in all the analysed data sets. The inclusion of *F. fragrans* within the Gigantea clade was shown in the analyses with the ITS, ITS-expanded and combined data sets. In the *trn*L-F data set analyses, the position of *F. fragrans* was unclear and in the *ndh*F data set analysis, *F. fragrans* was sister to the Gigantea + Racemosa clades.

Thus, there was large but incomplete correspondence between well-formed monophyletic groups in the series of analyses conducted with the various molecular data sets and the existing taxonomic "sections" of *Fagraea* s.l. Whereas sections *Fagraea* and *Racemosae* appear well-defined, the section *Cyrtophyllum* appears to be an artificial grouping of two natural groups, and *F. crenulata* appears to be somewhat isolated within *Fagraea* s.l.

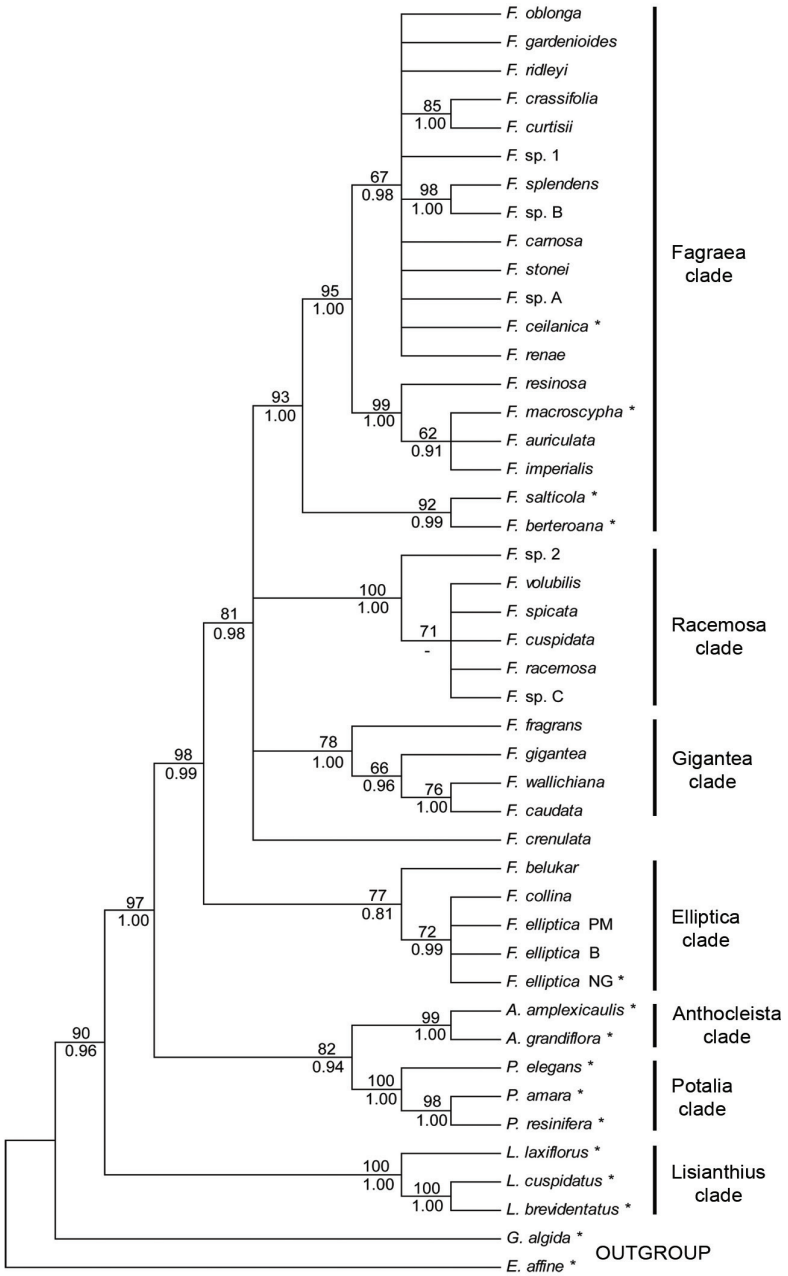

Fig. 5. Strict consensus of 2145 equally parsimonious trees based on the expanded ITS sequence data. The numbers above and below the branches denote Bootstrap and Bayesian Posterior Probability values, respectively. Length (L) = 582; consistency index (CI) = 0.7337; retention index (RI) = 0.8089. *A.* = *Anthocleista*; *E.* = *Exacum*; *F.* = *Fagraea*; *G.* = *Gentiana*; *L.* = *Lisianthius*; *P.* = *Potalia*. Different accessions of *F. elliptica* are indicated (PM = Peninsular Malaysia, B = Borneo, NG = New Guinea). An asterisk after a name indicates the taxon's sequence data obtained from GenBank (see Methods and Materials).

Potentially synapomorphic morphological characters for clades

An attempt to identify unique, non-molecular characters defining monophyletic groups (likely synapomorphies in cladistic terms) was made. This is shown in summary form in Table 5. This attempt might also be viewed against the pre-existing notion that *Fagraea* s.l. was not particularly well distinguished as a natural group with a large defining character suite. Prevost (1978), surveying modularity in growth architecture among tropical woody plants, remarked that *Fagraea* s.l. was quite polymorphic in the sense that it included many examples of architectural models. The examples she provided included *F. crenulata*, *F. fragrans* and *F. racemosa* s.l. (she cited an Australian-New Guinea provenance for this), with the models of Fagerlind, Aubréville and Roux, respectively. We have surveyed multiple species for the tree architectural character and find distinct forms representing each of the identified lineages or clades in our phylogenetic study.

From the present survey, it can be appreciated that the section *Racemosae* appears to have the most number of unique character-states among the sections compared within *Fagraea* s.l. It is thus highly distinct by morphological characteristics from the other two sections and from *F. crenulata*. *Fagraea crenulata* appears to diverge from all other *Fagraea* s.l. in a number of characteristics, including Fagerlind's architectural model with modular branch construction (as mentioned by Hallé et al. 1978 and Prevost 1978), and the presence of stem/branch prickles and a serrulate leaf margin. Likewise, two groups appear morphologically well distinguished within the section *Cyrtophyllum*— one with Scarrone's architectural model, with orthotropic branch complexes and terminal flowering; the other with Aubréville's architectural model, with branches that extend plagiotropically by apposition and axillary flowering. Likely synapomorphies were identifiable for the Fagraea clade, the Racemosa clade and the Gigantea clade, but not for the Elliptica clade, which resolved most basally in the ITS analyses (Fig. 1 & 5).

THE FAGRAEA CLADE (*Fagraea* s.s., excluding *F. crenulata*) — The growth habits of members in this group are erect, scrambling, climbing or scandent shrubs or smallish trees, which are also facultatively hemi-epiphytes. All the other distinguished groups as well as *F. crenulata* are free-standing trees and do not have scrambling, climbing or hemi-epiphytic habit. *Anthocleista* and *Potalia* are free-standing trees, with a few climbers found in the former genus (Struwe et al. 2001; Struwe & Albert 2004). Members of *Fagraea* s.s. produce copious amounts of creamy-pale to yellowish latex in the fruit pericarp (visible especially when fresh fruits are cut or bruised). The other groups as well as *F. crenulata* either have small amounts of translucent gummy latex or (like the Racemosae clade, *Anthocleista* and *Potalia*) do not produce latex at all. The seed shape in *Fagraea* s.s. is ellipsoid-rounded (similar to the condition in *Anthocleistus* and *Potalia*: Struwe et al. 2001; Struwe & Albert 2004), compared to polygonal in all the other three groups of *Fagraea* s.l. as well as *F. crenulata*.

FAGRAEA CRENULATA — The Fagerlind's tree architectural model is found only in *F. crenulata* within *Fagraea* s.l. The model applies to trees that have a monopodial trunk

with episodic growth as well as branching tiers. The branch modules are terminated by an inflorescence. The other groups have Scarrone's, Aubréville's or Roux's models. Perhaps the character that makes this species bizarre among *Fagraea* s.l. is the presence of prickles on the stems and branches. Prickles are completely absent in all other species within *Fagraea* s.l. and the Potaliinae. Another aberrant character which unmistakably distinguishes this species is the serrulate leaf margin. All other species in *Fagraea* s.l., *Anthocleista* and *Potalia* (Struwe et al. 2001; Struwe & Albert 2004) have entire leaf margins. In terms of morphology, *F. crenulata* is arguably the most enigmatic species in *Fagraea* s.l.

Notwithstanding, *Fagraea* s.s. and *F. crenulata* do share a number of characters (Table 5), although most of these are not exclusive to them and can be found in other groups or genera. A possible link is that the petiolar sheaths of a leaf pair in both *Fagraea* s.s. and *F. crenulata* do not fully fuse. The slight fusion at the extreme edges of the pairing sheaths do not form a consistent cup-like structure (ochrea) around the stem as in *Anthocleista* and *Potalia* (Struwe et al. 2001; Struwe & Albert 2004; Struwe, pers. comm.) and other taxa of *Fagraea* s.l.

THE RACEMOSA CLADE — Roux's architectural model applies to all members of the Racemosa clade. Species with this growth model have a monopodial orthotropic stem / trunk with continuous growth. In comparison, all other taxa in *Fagraea* s.l. have episodic stem / trunk growth. The branches on the stem/trunk in the Racemosa clade are opposite and decussate but leaf arrangement on the branches is secondarily distichous; in all the other groups of *Fagraea* s.l., the leaves are opposite-decussate in arrangement. The branches in the Racemosa clade are plagiotropic, ending with a terminal inflorescence, whereas the branches in other groups are orthotropic complexes or plagiotropic by apposition. The terminal buds of vegetative shoots in the Racemosa clade are not conspicuously covered with any resinous substances whereas all the others (including *Anthocleista* and *Potalia*: Struwe, pers. comm.) have creamy to yellowish resin covering the shoot apices.

Among Racemosa clade members, the inflorescence is generally a pendulous elongate panicle with cymose branching where the branching pairs are condensed and distinctly shorter than the rachis. In comparison, the inflorescences in the other groups bear only a solitary flower or are branched cymes in which the longest basal branches are nearly as long as the rachis. The mature fruit colour in the Racemosa clade are generally pale to dark brown, whereas in the other groups they are yellow-orange to red-scarlet or creamy pale grey-green to white. The fruits in the Racemosa clade (and also *Anthocleista* and *Potalia*) do not exude any conspicuous latex, whereas fruits in all the other *Fagraea* s.l. groups produce a gummy latex. Also, fruits in the Racemosa clade (and also in *Anthocleista* and *Potalia*: Struwe, pers. comm.) in both fresh and dried specimens, have a rather intact fruitwall epidermis that does not easily come off; on the other hand, in all the other *Fagraea* s.l. groups, the epidermis separates easily as a thin, tough and translucent peel. The mature fruit wall in dried herbarium samples is firm and retains its rounded structure (as in *Anthocleista* and *Potalia*: Struwe, pers. comm.), whereas in other groups the fruit wall breaks down and crumples as it dries.

Table 5. A comparison of various habit, plant-architectural and morphological characters found in distinct groups of *Fagraea* s.l. resolving as monophyletic groups and an isolated lineage in molecular phylogenetic analyses in the present study. Specially diagnostic character-states which are potentially synapomorphic for the identified clades (in the context of *Fagraea* s.l.) are given in bold italics. Other states found which are unusual for (absent in) other (non-Malesian) members of the subtribe, are marked with an asterisk. Thus, traits marked by both an asterisk and bold italics are likely clade / generic synapomorphies in the context of the subtribe in general. A few characters placed in brackets (column 1) appear not to be of special taxonomic utility at this level of classification.

Monophyletic groups in molecular analyses (present work)	**Elliptica clade**	**Gigantea clade**	**Fagraea clade**	***F. crenulata***	**Racemosa clade**
Sectional name fide Leenhouts (1962)	*Cyrtophyllum*	*Cyrtophyllum*	*Fagraea*	*Fagraea*	*Racemosae*
Growth habit	free-standing trees, never scrambling or climbing or hemi-epiphytic	free-standing trees, never scrambling or climbing or hemi-epiphytic	erect, *scrambling, *climbing or *scandent shrubs or small trees but these also *facultative hemi-epiphytes	free-standing trees, never scrambling or climbing or hemi-epiphytic	free-standing trees, never scrambling or climbing or hemi-epiphytic
General architecture	Scarrone's model	***Aubréville's model***	Scarrone's model	***Fagerlind's model***	***Roux's model***
Trunk / stem growth	episodic	episodic	episodic	episodic	***continuous***
Trunk / stem bark	becoming fissured in older trees or smooth to scaly-dippled; lacking prickles	becoming fissured in older trees; lacking prickles	smooth to lightly scaly-dippled; lacking prickles	becoming fissured and ***densely set with prickles***	becoming fissured in older trees; lacking prickles
Branches on stem / trunk	orthotropic complexes	***plagiotropic by apposition***	orthotropic complexes	***plagiotropic by substitution and modular***	***plagiotropic***
Vegetative terminal buds	yellowish resinous	yellowish resinous	creamy yellowish resinous	creamy yellowish resinous	***non-resinous***
Leaf arrangement on branches	decussate	decussate	decussate	decussate	***secondarily distichous***
Leaf margin	entire	entire	entire	***serrulate-crenulate***	entire
Petiolar sheaths	fused at node into a cuplike ochrea	fused at node into a cuplike ochrea	*not fused to slightly fused at extreme edges, not forming a cuplike ochrea	*not fused to slightly fused at extreme edges, not forming a cuplike ochrea	fused at node into a cuplike ochrea

Inflorescence, general form	branched cymes (basal branches longest, nearly as long as rachis, mostly rebranched)	branched cymes (basal branches longest, nearly as long as rachis, mostly rebranched)	solitary flowers / 1–few-flowered cymes / branched cymes (basal branches longest, nearly as long as rachis, mostly rebranched)	branched cymes (basal branches longest, nearly as long as rachis, mostly rebranched)	***elongate panicle with cymose branching (branches several pairs, condensed, distinctly shorter than rachis)***
Inflorescence, position	terminal	***axillary**	terminal	terminal	terminal
(Number of flowers per inflorescence)	several to many	several to many	several to many; in some taxa reduced to *single flowers	many	several to many
(Corolla size)	very small (up to 10 mm wide at mouth)	very small (up to 10 mm wide at mouth)	very small to *large (over 40–50 mm wide at mouth)	medium	very small to medium (up to 25 mm wide at mouth)
Stamen exsertness	*long-exsert (typically > 70% exsert)	*long-exsert (typically > 70% exsert)	slightly to medium-exsert	medium exsert	not to medium-exsert
Style exsertness	*medium- to long-exsert (typically > 40% exsert)	*medium- to long-exsert (typically > 40% exsert)	not to slightly exsert	not to slightly exsert	not to slightly exsert
Stigma structure & form	knoblike: stigma base not expanding conspicuously; stigmatic surface with 2 very slightly distinct lobes resembling twin mounds	knoblike: stigma base not expanding conspicuously; stigmatic surface with 2 very slightly distinct lobes resembling twin mounds	*stigma base expanding into a circular platelike rim that is often undulating; stigmatic surface weakly to distinctly 2-lobed	*stigma base expanding into a circular platelike rim that is often undulating; stigmatic surface weakly to distinctly 2-lobed	*stigma base expanding into a circular platelike rim; stigmatic surface moundlike or weakly to distinctly 2-lobed
(Fruit size)	very small (< 10 mm diameter)	very small (< 10 mm diameter)	very small to *big (> 40 mm diameter)	medium (10–15 mm across or more)	very small to medium (< 15 mm diameter)
Fruit colour at maturity	yellow-orange to red-scarlet	yellow-orange to red-scarlet	creamy pale grey-green to white	pale grey-green	***pale to dark brown***
Latex in fruit epidermis / fruitwall	small amounts of *translucent gummy latex	small amounts of *translucent gummy latex	***copious creamy pale yellowish latex***	small amounts of *translucent gummy latex	***no latex***
Fruit epidermis	*separating easily as a thin, tough, translucent 'peel'	*separating easily as a thin, tough, translucent 'peel'	*separating easily as a thin, tough, translucent 'peel'	*separating easily as a thin, tough, translucent 'peel'	***not separating from the fruit wall easily***
Fruitwall at maturity	*soft	*soft	*soft	*soft	***firm***
Seed shape	*polygonal	*polygonal	***ellipsoid-rounded***	*polygonal	*polygonal

THE GIGANTEA CLADE — The general architecture of members of the Gigantea clade follows Aubréville's model in which the monopodial main trunk shows episodic growth with opposite-decussate phyllotaxis, and the branches extend plagiotropically by apposition (this is sometimes referred to as Terminalian branching) (Hallé et al. 1978; Prevost 1978). The other Malesian groups in *Fagraea* s.l. follow Scarrone's, Fagerlind's or Roux's models. The position of the inflorescence in the Gigantea clade is axillary, whereas in all other groups of *Fagraea* s.l. and its subtribe (including *Anthocleista* and *Potalia*: Struwe et al. 2001; Struwe & Albert 2004), the inflorescence is terminal.

THE ELLIPTICA CLADE — The Gigantea and Elliptica clades have several morphological similarities that distinguish them from other taxa in *Fagraea* s.l. The exsertness of the filament from the corolla tube is very prominent within these two clades, typically more than 70% of the total length of the filament, whereas in the others, the filament is not exserted at all or only slightly to moderately so. Also, style exsertness in these two clades is typically more than 40% of the total length of the pistil, whereas in the other groups it is either not or only slightly exserted (refer to Appendix B for a detailed comparison of filament and style exsertness among the taxa). The structure of the stigma in these two clades is knoblike or capitate and the base of the stigma does not expand conspicuously, as in *Anthocleista* and *Potalia* (Struwe, pers. comm.). In *Fagraea* s.s., *F. crenulata* and the Racemosa clade, the stigma structure is peltate due to the base of the stigma expanding conspicuously into a circular plate-like rim. Further, the fruits in the Gigantea and Elliptica clades turn yellow-orange to red-scarlet upon maturity whereas in the other two clades as well as *F. crenulata*, the fruits ripen creamy grey-green to white or dark brown. The fruits are also generally smaller in these two clades compared to the others. Given the distinctness of the Gigantea and Elliptica clades as expressed in the topologies resulting from the molecular analyses (Fig. 1–5), any apparently shared character-states they have must be considered homoplasious in nature (i.e., similarity not due to common ancestry). Furthermore, inflorescence position is consistently terminal in the Elliptica clade, and axillary in the Gigantea clade.

Fagraea: complex genus or several genera?

The analysis with the expanded ITS data set (Fig. 5) shows that the monophyletic groups are comparable in distinctness to several recently reviewed or revised ingroup genera, i.e., *Anthocleista*, *Potalia* and *Lisianthius* (Struwe et al. 2002; Struwe & Albert 2004; Weaver 1972; Sytsma 1988), which are well-established genera of the same tribe. From the systematic and taxonomic points of view, therefore, the respective clades are best recognised as separate genera because of their phylogenetic resolution as monophyletic groups or isolated lineages, and their equivalence to other well-established genera of the same tribe.

A survey of possible morphological markers also suggests that *Fagraea* s.l.

contains morphologically well-distinguished groups of taxa with clear-cut boundaries. It was possible to ascribe potential morphological synapomorphies for the four monophyletic groups in *Fagraea* s.l. and for *F. crenulata*.

In summary, *Fagraea* s.l. is morphologically too divergent to be considered as a single genus and even the sectional classification available (Leenhouts 1962) appears to be incompletely circumscribed. The proposed concepts here for recognising the main lineages as genera are clearly applicable to Sundaland species, which represent the major geographical core of the complex, as these were well-represented in the present analyses. Notwithstanding an increased clarity now available for sorting out *Fagraea* s.l., various remarks on remaining problems and suggestions for future work may be made. The two species, *F. berteroana* and *F. salticola* (New Guinea and the south-west Pacific islands) included in the expanded ITS analysis, were resolved within the Fagraea clade (= *Fagraea* s.s.). These species are among several from east Malesia that form a group morphologically distinguished from others in *Fagraea* s.s. by a fleshy ring on the inside of the corolla tube, at the insertion level of the stamens (Leenhouts 1962; Struwe et al. 2002). Better taxon sampling that includes more species from other parts of Malesia and the south-west Pacific islands in future phylogenetic analyses may give better resolution and confidence about clade relationships for this group. The potential inclusion of taxa from beyond the so-called boundaries of Malesia, such as *F. schlechteri* Gilg & Gilg-Ben. from New Caledonia, which Prevost (1978) has observed to have the modular growth model of Koriba (so far not encountered within the Malesian *Fagraea* complex), also promises greater insight. Also, the exploration of further gene regions in similar studies should be interesting.

ACKNOWLEDGEMENTS. This paper is based in part on the doctoral research of the senior author (MS) at the University of Malaya, supervised by and augmented with research findings of the second author (KMW). The University of Malaya provided financial support for the fieldwork and phylogenetic study through Fundamental Research Grant FS 264/2007C. We thank Dr. Chandran Somasundram for kindly allowing use of equipment and space in his laboratory; Y.W. Low and Zulkapli Ibrahim for specially assisting with field work and material for DNA extraction; and N. Doraisamy, R. Sujatha, D.K.P. Lee and W.L. Goh for assistance in the lab. The Keepers and Curators of the BO, K, KEP, KLU, L, SAN and SING herbaria are thanked for loans of, and permission to study, specimens. Much thanks are also due to the following for their kind assistance rendered: Assoc. Prof. Normawati Haron and Dr. K.T. Yong (KLU); Dr. J.T. Pereira, Dr. R.C. Ong, Dr. R. Nilus, M. Postar, S. Ahmad and J. Jamirus (SAN); Joffre Hj Ali Ahmad, Muhd Ariffin (BRUN); Dr. A.S. Kamariah (University of Brunei Darussalam); Dr. R.C.K. Chung (KEP); Dr. Benjamin van Ee and Dr. Wenheng Zhang (Harvard University Herbaria); Dr. B.J. Conn (NSW). Research for, and preparation of, this work have specially benefitted through advice and comments from Prof. Rofina Yasmin and Prof. A.L. Lim (University of Malaya); Prof. Lena Struwe (Rutgers University), who also provided plant materials and literature; Prof. Nianhe Xia (South China Institute of Botany); Dr. Koichi Kamiya (Ehime University); Prof. Charles Davis (Harvard University Herbaria); and Prof. Christian Puff (University of Vienna). MS wishes to thank Dr. R. Shanti for her patience and support; Dr. M. Vijayndra, who patiently proofread most of his dissertation; and many friends, colleagues and associates at the University of Malaya's Rimba Ilmu Botanic Garden. We are grateful for

the support of Prof. Mohd Sofian Azirun, Dean of Science, and Prof. Rosli Hashim, Head of the Institute of Biological Sciences, University of Malaya, as well as the previous Deans, Prof. Hj Mohamed Abdul Majid and Prof. Amru Nasrulhaq Boyce.

References

Bentham, G. (1856) Notes on Loganiaceae. *J. Proc. Linn. Soc., Bot.* 1: 52–115.

Blume, C.L. (1838) *Rumphia*. Vol. 2. Leiden.

Blume, C.L. (1850) *Stirpium exoticarium novarum vel minus cognitarium ex vivis aut siccis brevis exposito et descriptio*. Pp. 163–173. Lugduni-Batavorum.

Burkill I.H. (1936) *A Dictionary of the Economic Products of the Malay Peninsula.* Fourth print (2002). Kuala Lumpur: Ministry of Agriculture and Co-Operatives. Vol. 1: 1009–1013.

Cammerloher, H. (1923) Die Loganiaceen und Buddleiaceen Niederländisch-Indiens. *Bull. Jard. Bot. Buitenzorg* Ser. III, 5 (4): 295–338.

Chassot, P., Nemomissa, S., Yuan, Y.-M., Küpfer, P. (2001) High paraphyly of *Swertia* L. (Gentianaceae) in the Gentianella-lineage as revealed by nuclear and chloroplast DNA sequence variation. *Pl. Syst. Evol.* 229: 1–21.

Conn, B.J. & Brown, E.A. (1993) Review of *Fagraea gracilipes* complex (Loganiaceae). *Telopea* 5 (2): 363–374.

Cunningham, C.W., Omland, K.E. & Oakley, T.H. (1998) Reconstructing ancestral character states: a critical reappraisal. *Trends Ecol. Evol.* 13: 361–366.

Davis, C.C., Anderson, W.R. & Donogue, M.J. (2001) Phylogeny of Malphigiaceae: Evidence from chloroplast *ndh*F and *trn*L–F nucleotide sequence. *Amer. J. Bot.* 88: 1830–1846.

Downie, S.R. & Palmer, J.D. (1992) Restriction site mapping of the choloplast inverted repeat: a molecular phylogeny of the Asteridae. *Ann. Missouri Bot. Gard.* 79: 266–283.

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

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

Farris, J.S., Källersjö, M., Kluge, A.G. & Bult, C. (1994) Testing significance of incongrunce. *Cladistics* 10: 315–319.

Farris, J.S., Källersjö, M., Kluge, A.G. & Bult, C. (1995) Constructing a significant test of incongruence. *Syst. Biol.* 44: 570–572.

Felsenstein, J. (1985) Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* 39: 783–791.

Fitch, W.M. (1971) Toward defining the course of evolution: minimal change for a specific tree topology. *Syst. Zool.* 20: 406–416.

Gielly, L. & Taberlet, P. (1996) A phylogeny of the European gentians inferred from chloroplast *trn*L (UAA) intron sequences. *Bot. J. Linn. Soc.* 120: 57–75.

Hagen, K.B. von & Kadereit, J.W. (2001) The phylogeny of *Gentianella* (Gentianaceae) and its colonization of the Southern hemisphere as revealed by nuclear and chloroplast DNA sequence variation. *Org. Divers. Evol.* 1: 61–79.

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

Hallé, F., Oldeman, R.A.A. & Tomlinson, P.B. (1978) *Tropical trees and forests. An architectural analysis.* Berlin: Springer-Verlag.

Huelsenback, J.P. & Ronquist, F. (2001) MrBayes: bayesian inference of phylogenetic trees. *Biometrics* 17: 754–755.

Huelsenback, J.P., Ronquist, F. & Hall, B. (2001) MrBayes: A program for the Bayesian of phylogeny (Instruction Manual). http://golab.unl.edu/teaching/SBseminar/manual.pdf (accessed September 25, 2009).

Jensen, S.R. (1992) Systematic implications of the distribution of iridoids and other chemical compounds in the Loganiaceae and other families of the Asteridae. *Ann. Missouri Bot. Gard.* 79: 284–302.

Jensen, S.R. & Schripsema, J. (2002) Chemotaxonomy and pharmacology of Gentianaceae. In: Struwe, L. & Albert, V.A. (eds.), *Gentianaceae–systematics and natural history*. Pp. 573–631. U.K., Cambridge: Cambridge University Press.

Kauff, F. & Lutzoni, F. (2002) Phylogeny of the Gyalectales and Ostropales (Ascomycota, Fungi): among and within other order relationships based on nuclear ribosomal RNA small and large subunits. *Molec. Phylogenet. Evol.* 25: 138–156.

Kluge, A.G. & Farris, J.S. (1969) Quantitative phyletics and the evolution of anurans. *Syst. Zool.* 18: 1–32.

Larget, B. & Simon, L. (1999) Markov chain Monte Carlo algorithms in the Bayesian analyses of phylogenetic trees. *Molec. Biol. Evol.* 16: 750–759.

Larkin, M.A., Blackshields, G., Brown, N.P., Chenna, R., McGettigan, P.A., McWilliam, H., Valentin, F., Wallace, I.M., Wilm, A., Lopez, R., Thompson, J.D., Gibson, T.J. & Higgins, D.G. (2007) Clustal W and Clustal X version 2.0. *Bioinformatics* 23: 2947–2948.

Leenhouts, P.W. (1962) Loganiaceae. In: Steenis, C.G.G.J. van (ed.), *Flora Malesiana.* Series 1 Vol. 6. Pp. 293–336. Groningen: Wolters-Noordhoff Publishing.

Leenhouts, P.W. (1984) A new *Fagraea* from Borneo (Loganiaceae). *Blumea* 29: 423–424.

Leeuwenberg, A.J.M. & Leenhouts, P.W. (1980) Taxonomy. In: Leeuwenberg, A.J.M. (ed.), *Engler and Prantl's Die natürlichen Pflanzenfamilien, Angiospermae. Ordnung Gentianales, Fam. Loganiaceae.* Vol. 28b (1). Pp. 8–96. Berlin: Duncker & Humblot.

Lewis, P.O. (2001) Phylogenetic systematics turns over a new leaf. *Trends Ecol. Evol.* 16: 30–37.

Macmillan H. F. (1991) *Tropical Planting and Gardening*. Sixth edition. Pp. 115 & 265. Malayan Nature Society, Kuala Lumpur.

Miquel, F.A.W. (1857) *Flora van Nederlandsch-Indië*. Vol. 2. Amsterdam.

Molina, J. & Struwe, L. (2009) Utility of secondary structure in phylogenetic reconstruction using *nr*DNA ITS sequences – an example from *Potalieae* (Gentianaceae: Asteridae). *Syst. Bot.* 34 (2): 414–428.

Motley T.J. (2004) The ethnobotany of *Fagraea* Thunb. (Gentianaceae): The Timber of Malesia and the scent of Polynesia. *Econ. Bot.* 58 (3): 396–409.

Nylander, J.A.A. (2004) MrModeltest 2.2. Evolutionary Biology Centre, Uppsala University. Software programme distributed by the author.

Olmstead, R.G., Bremer, B., Scott, K.M. & Palmer, J.D. (1993) A parsimony analysis of the *Asteridae* sensu lato based on *rbc*L sequences. *Ann. Missouri Bot. Gard.* 80: 700–722.

Perry, L.M. (1980) *Medicinal plants of East and Southeast Asia.* Cambridge: The MIT Press.

Prevost, M.-F. (1978) Modular construction and its distribution in tropical woody plants. In: Tomlinson, P.B. & Zimmermann, M.H. (eds) *Tropical Trees as Living Systems.* Pp. 223–231. New York: Cambridge Univ. Press.

Qiagen Sample & Assay Technologies (2003–2009) Handbook for DNA extraction. *http://www1.qiagen.com/.*

Quisumbing, E. (1978) *Medicinal plants of the Phillipines.* Philippines: Katha Publishing Co., Inc.

Ranker, T.A., Geiger, J.M.O., Kennedy, S.C., Smith, A.R., Haufler, C.H. & Parris, B.S. (2003) Molecular phylogenetics and evolution of the endemic Hawaiian genus *Adenophorus* (Grammitidaceae). *Molec. Phylogenet. Evol.* 26: 337–347.

Rannala, B. & Yang, Z. (1996) Probability distribution of molecular evolutionary trees: a new method of phylogenetic inference. *J. Molec. Evol.* 43: 304–311.

Ridley, H.N. (1923) *The Flora of the Malay Peninsula*. Vol. 2. Pp. 415–422. London: L. Reeve & Co.

Solereder, H. (1892) Loganiaceae. In: Engler, A. & Prantl, K. (eds.), *Die natürlichen Pflanzenfamilien* 4(2): 1–50. Germany, Leipzig: Engelmann.

Struwe, L., Albert, V.A. & Bremer, B. (1994) Cladistics and family level classification of Gentianales. *Cladistics* 10: 175–206.

Struwe, L. & Albert, V.A. (1997) Floristics, cladistics and classification: three case studies in Gentianales. In: Dransfield, J., Coode, M.J.E. & Simpson, D.A. (eds.), *Plant Diversity in Malesia, III.* Pp. 321–352. Royal Botanic Gardens, Kew.

Struwe, L., Kadereit, J., Klackenberg, J., Nilsson, S., Thiv, M., von Hagen, K.B. & Albert, V.A. (2002) Systematics, character evolution and biogeography of Gentianaceae, including a new tribal and subtribal classification. In: Struwe, L. & Albert, V.A. (eds.), *Gentianaceae–Systematics and Natural History*. Pp. 21–309. U.K., Cambridge: Cambridge University Press.

Struwe, L. & Albert, V.A. (2004) A monograph of neotropical *Potalia* (Gentinaceae: Potalieae). *Syst. Bot.* 29: 670–701.

Swofford, D.L. (2001) *PAUP*: Phylogenetic analyses using parsimony and other methods, version 4.0*. Sunderland, Massachussetts: Sinauer Associates.

Sytsma, K.J. (1988) Taxonomic revision of the Central American *Lisianthius skinneri* species complex (Gentianaceae). *Ann. Missouri Bot. Gard.* 75: 1587–1602.

Taberlet, P., Gielly, L., Pautou, G. & Bouvet, J. (1991) Universal primers for amplification of the three non-coding regions of chloroplast DNA. *Pl. Molec. Biol.* 17: 1105–1109.

Thiv, M., Struwe, L. & Kaderiet, J.W. (1999) The phylogenetic relationships and evolution on the Canarian laurel forest endemic *Ixanthus viscosus* (Aiton) Griseb. (Gentianaceae): evidence from *mat*K and ITS sequences, and floral morphology and anatomy. *Pl. Syst. Evol.* 218: 299–317.

Watson, J.G. (1935) Tembusu. *Malayan Forester* 4: 136–137.

Weaver, R.E. (1972) A revision of a neotropical genus *Lisianthus* (Gentianaceae). *J. Arnold Arb.* 53: 76–100.

White, T.J., Burns, T., Lee, S. & Taylor, J. (1990) Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. In: Innis, M.A., Gelfand, D.H., Sninsky, J.J. & White, T.J. (eds.), *PCR Protocols: A Guide to Methods and Applications.* Pp. 315–322. San Diego: Academic Press.

Wong, K.M. & Sugau, J.B. (1996) A revision of *Fagraea* (Loganiaceae) in Borneo, with notes on related Malesian species and 21 new species. *Sandakania* 8: 1–93.

Yuan, Y.-M. & Küpfer, P. (1995) Molecular phylogenetics of the subtribe Gentianinae (Gentianaceae) inferred from the sequences of the internal transcribed spacer (ITS) of nuclear ribosomal DNA. *Pl. Syst. Evol.* 196: 206–226.

Yuan, Y.-M. & Küpfer, P. (1997) The monophyly and rapid evolution of *Gentiana* sect. *Chondrophyllae* Bunge s.l. (Gentianaceae): evidence from the nucleotide sequences of the internal transcribed spacer of nuclear ribosomal DNA. *Bot. J. Linn. Soc.* 123: 25–43.

Yuan, Y.-M., Küpfer, P. & Doyle, J.J. (1996) Infrageneric phylogeny of the genus *Gentiana* (Gentianaceae) inferred from nucleotide sequences of the internal transcribed spacer (ITS) of nuclear ribosomal DNA. *Amer. J. Bot.* 83: 641–652.

Yuan, Y.-M., Wolhauser, S., Möller, M., Chassot, P., Mansion, G., Grant, J., Küpfer, P. & Klackenberg, J. (2003) Monophyly and relationships of the tribe Exaceae (Gentianaceae) inferred from nuclear ribosomal and chloroplast DNA sequences. *Molec. Phylogenet. Evol.* 28: 500–517.

Appendix A. Sequences of primers used in PCR for amplifying the ITS, *trn*L intron, *trn*L–F spacer and *ndh*F regions.

PCR amplified regions	Primer names	Sequences (5' — 3')	Approximate size of amplification in PCR	Source
ITS	ITS 1	TCC GTA GGT GAA CCT GCG G	700–750 bp	White et al. 1990
	ITS 4	TCC TCC GCT TAT TGA TAT GC		
*trn*L intron	'C'	CGA AAT CGG TAG ACG CTA CG	300–400 bp	Taberlet et al. 1991
	'D'	GGG GAT AGA GGG ACT TGA AC		
*trn*L–F spacer	'E'	GGT TCA AGT CCC TCT ATC CC	300–400 bp	Taberlet et al. 1991
	'F'	ATT TGA ACT GGT GAC ACG AG		
*ndh*F	GB 1 Fwd	CTT TCA TTC CAC TTC CAG TTC CT	900–1000 bp	This study
	GB 1 Rev	TAT AGG GTG AAT AGC CAA GAA GCC		
	GB 2 Fwd	AAA GCC AAA ATA TGG TTC TTA TGG G	900–1000 bp	This study
	GB 2 Rev	AAA TAA ATA GAA GAA AAT ATA AGA AGA AAT GCG		

Appendix B. Style and filament exsertness from the corolla tube in selected species of *Fagraea* s.l., representing sections *Cyrtophyllum*, *Fagraea*, and *Racemosae.*
The three blocks of species from top to bottom in the Table, correspond to *Fagraea* sections *Cyrtophyllum*, *Fagraea* and *Racemosae*, respectively. *CA* = *Cyrtophyllum* (Axillary Inflorescence), *CT* = *Cyrtophyllum* (Terminal Inflorescence), *F*. = *Fagraea*, *R*. = *Racemosae*.

Species – and section in *Fagraea* s.l.	Corolla tube length (mm)	Style length (mm)	Style protrusion (mm)	Style protrusion (%)	Filament length (mm)	Filament protrusion (mm)	Filament protrusion (%)
F. fragrans - CA	(4–)6–8	(14–)18–22	(8–)10–12(–14)	57–64	(10–)12–16(–17)	(8–)10–12(–13)	76–80
F. gigantea - CA	7–8	(12–)18–22	(5–)12–14	42–64	13–15	12–13	87–92
F. wallichiana - CA	(12–) 20–25	(34–)42–45(–55)	22–25(–30)	55–65	(27–)30–38	20–23(–28)	73–74
F. elliptica - CT	3–5	(6) –7–9	3–4	44–50	(4–)7–8	(4–)7–8	100
F. auriculata - F	60–82	70–90	8–10	11–12	45–60	15–23	33–38
F. carnosa - F	106–140	123–126	0–17	0–13	12–15	5	33–41
F. crassifolia - F	26–30	28–30	0–2	0–7	–	–	–
F. crenulata - F	15–18	15–20(–23)	0–5	0–22	12–14	7–8	57–58
F. curtisii - F	35–55	c. 60	c. 5	c. 8.3	32–40	c. 10	c. 25
F. gardenioides - F	40–53	50–55	2–10	4–18	20–27	12–17	60–63
F. imperialis - F	90–160	90–115	0–5	0–4	(55–)80–90	(5–)27–40	9–44
F. littoralis - F	25–32	32–35	3–7	9–20	23–25	12–13	c. 52
F. oblonga - F	21–36	22–30	1–6	5–20	18–22	c. 8	36–44
F. renae - F	23–34	28–32	0–5	0–16	20–26	10–14	50–54
F. ridleyi - F	32–37	40–45	c. 8	18–20	25–28	c. 10	36–40
F. splendens - F	25–37	40–45	8–15	20–33	20–25	7–8	32–35
F. tubulosa - F	75–93	85–88	5–10	6–12	15–18	c. 10	56–67
F. sp. B - *F*	26–36	35–42	6–9	14–26	20–24	6–10	30–42
F. sp. D - *F*	43–60	65–70	10–22	15–31	40–50	17–20	40–43
F. sp. E - *F*	22–30	–	–	–	16–20	5–6	30–31
F. maingayi - R	16–25	12–23	0	0	19–23	2–6	11–26
F. peninsularis - R	c. 14	17–19	3–5	18–26	6–7	1–2	17–28
F. racemosa - R	10–22	15–20(–25)	3–5	20	11–17(–20)	6–8	40–55
F. volubilis - R	15–25	17–20	0–2	0–10	11–13	0–3	0–23
F. sp. C - *R*	21–26	22–25	0–1	0–4	10–15	0	0
F. sp. G - *R*	20–25	20–28	0–3	0–11	15–16	4–5	27–31