BLUMEA 44 (1999) 351–379

NUMERICAL ANALYSIS OF THAI MEMBERS OF THE EUGENIA–SYZYGIUM GROUP (MYRTACEAE)

J. PARNELL

Herbarium, School of Botany, Trinity College, Dublin, Ireland

SUMMARY

Four different analyses of Thai *Syzygium* show, with very small discrepancies, that one smaller monophyletic and one larger polyphyletic group can be recognised. The smaller monophyletic group warrants sectional rank as *Syzygium* section *Jambosa* and consists of *S. anacardiifolium*, *S. aqueum*, *S. diospyrifolium*, *S. formosum*, *S. foxworthianum*, *S. jambos*, *S. lakshanakarae*, *S. malaccense*, *S. megacarpum*, *S. papillosum*, *S. pseudoformosum*, *S. pycnanthum*, *S. samarangense*, *S. scortechinii* and *S. siamense*.

Key words: *Eugenia*, *Jambosa*, *Syzygium*, cladistics, phenetic analysis, phylogenetic analysis.

INTRODUCTION

The Myrtaceae are a moderately sized, mostly southern hemisphere family containing between 132 and 150 genera and somewhere between 3,675 and 3,900 species (Johnson et al., 1988; Kochummen, 1995; Lemmens, 1995; Mabberley, 1997; Schmid, 1980; all provide slightly different estimates). Despite its modest size the family poses a disproportionately large number of complex taxonomic problems evident at many levels in the taxonomic hierarchy. For example, McVaugh (1968) described the species of American Myrtaceae as "distressingly alike in aspect and in most individual characteristics, making identification and classification of both genera and species a correspondingly difficult and tedious matter."

Within the Myrtaceae two unequally sized subfamilies were recognised by Niedenzu (1893): the Myrtoideae and Leptospermoideae. The Myrtoideae have only 60 genera (Nic Lughadha & Proença, 1996) but contain approximately two-thirds (2,375) of the species known in the entire family (Schmid, 1980). These species occur in the New and Old World tropics, are mainly shrubs or trees of wet-forests, usually have an inferior ovary, and almost always possess opposite broad leaves and a fleshy indehiscent fruit. The other subfamily, the Leptospermoideae, contains more genera (72) than the Myrtoideae but fewer species (1,300). Its species are largely Australasian, have alternate, spiral or opposite leaves, an inferior or superior ovary and, usually, a dry dehiscent fruit.

Niedenzu's (1893) subfamilial divisions are convenient and still used (e.g. Hora, 1978 and Nic Lughadha & Proença, 1996). However, several recent studies indicate that Niedenzu's classification is unsatisfactory (e.g. Johnson & Briggs, 1984 or Schmid, 1972, 1980). Johnson & Briggs (1984) developed an informal system of alliances and

suballiances through cladistic analysis of a fairly large and comprehensive dataset (Briggs & Johnson, 1979; Johnson & Briggs, 1984). Though the alliances appear well-supported their interrelationships and boundaries remain somewhat controversial. Johnson & Briggs (1984) conclude that Niedenzu's (1893) division of the family on the basis of fruit characteristics is not phylogenetically supportable: their alliances and suballiances cross traditional subfamilial boundaries.

Two alliances are of particular concern in this paper – the *Acmena* and Myrtoideae s.s. alliances (Johnson & Briggs, 1984) or, more specifically, the *Acmena* and *Eugenia* alliances (Johnson et al., 1988). Though traditionally placed close to each other in the fleshy fruited Myrtoideae, these alliances appear far apart phylogenetically (Johnson & Briggs, 1984). Indeed they could scarcely be much further apart in Johnson & Briggs' cladogram wherein the *Eugenia* alliance (including *Jossinia* Comm. ex DC.) is in a clade which incorporates the *Leptospermum* alliance whilst the *Acmena* alliance forms an altogether separate clade with, amongst others, the *Eucalyptus* alliance. However, in a later compromise classification both alliances are placed together in the subfamily Myrtoideae (Johnson et al., 1988). In this paper I adhere to this latter concept and use the term Myrtoideae in its traditional sense.

The *Acmena* and *Eugenia* alliances contain two of the most taxonomically confused genera in the Myrtaceae – *Eugenia* L. and *Syzygium* Gaertn. and a number of other genera which have, at one time or another, been cleaved off from or reunited with them. Schmid (1972) pointed out that there were about 35 generic names which have been or could be reduced to *Syzygium* s.l. and at least another 30 assignable to *Eugenia* s.l. Since then the number of segregates has increased with, for example, the description of *Waterhousea* Hyland and *Monimiastrum* J. Guého & A.J. Scott. Together these genera form a "vast array of more or less closely allied species" (Ashton, 1981). This 'array', dominated by *Eugenia* and *Syzygium*, is very large – Index Kewensis has over 3,000 species listed under *Eugenia* and over 1,000 under *Syzygium*. Undoubtedly this does not reflect the true balance in numbers of species between these genera when they are considered in the strict sense as even now many authors prefer, because of historical precedent, to ignore the differences between them.

Schmid (1972) provides the most comprehensive review of the status of *Syzygium* s.l. and makes clear why *Eugenia/Syzygium* were confused. Schmid's work summarises nearly all of the relevant references and arguments, and is therefore not repeated here. Table 1 summarises the relevant sections of Schmid (1972) together with Johnson & Briggs (1984).

As can be seen from Table 1:

1) there is some overlap in all of the most easily observable differential morphological characters between *Eugenia* and *Syzygium* (e.g. pubescence of vegetative and/or reproductive parts, fugacity of bracts and bracteoles, presence or absence of a pseudostipe and fusion of the cotyledons);

2) there are ± no meristic characters in this table. This reflects the fact that, as far as I know, there have been no published studies dealing with morphometric variation within the group;

3) there appear to be some truly diagnostic characteristics (e.g. pathway of vascular supply to the ovules, presence of siliceous rays, elongation of vessel-ray pits, carpel number), however none of these, with the possible exception of the latter, are easily

Table 1. Major differential and diagnostic characteristics of *Eugenia* and *Syzygium*. Data from Schmid (1972), Johnson & Briggs (1984) and personal observation. The percentage of species estimated to have a trait are shown and are largely taken from Schmid (1972).

Character	*Eugenia* s.s.	*Syzygium* s.l.
Pubescence of vegetative and/or reproductive parts	Usually present (90%)	Rarely present (5%)
Bracts and bracteoles	Present and conspicuous (80%)	Fugacous and inconspicuous (95%)
Pseudostipe[1]	Absent (99%)	Present (92%)
Cotyledons	Usually fused	Usually fusion absent
Degree of floral tube prolongation above the ovary	± Not prolonged	Prolonged
Perianth	Usually large and conspicuous	Usually small and inconspicuous
Surface of seed-coat	Smooth	Rough
Pathway of vascular supply to ovules	Transeptal	Axile
Vessels	Mostly grouped	Mostly solitary
Elongation of vessel-ray pits	Considerably elongated	Isodiametric or moderately elongated
Bordered pits from fibres	Absent	Present
Apotracheal parenchyma	Absent	Present
Paratracheal parenchyma	Confluent or banded	Scanty
Siliceous inclusions in rays	Present	Absent
Number of carpels	2	4–5
Number of seeds forming from ovules	Seeds reduced to 1 or 2	Most ovules forming seeds

1) Pseudostipe as defined by Wilson (1957) = pseudostalk of Henderson (1949).

observed. In general the work of Schmid (1972), Johnson & Briggs (1984) and others is based on a small fraction of the *Eugenia/Syzygium* complex; therefore their conclusions may be of limited application. On the other hand the sheer number of species involved make it unrealistic to expect the assemblage of a truly comprehensive dataset which will allow for an all-inclusive phylogenetic analysis and inarguable resolution of the controversy. Assemblage of data on a regional basis in the hope of eventual amalgamation offers the only potential solution. An additional difference between *Eugenia* and *Syzygium* not shown in Table 1 is that *Eugenia* is a largely tropical New World genus (though *Eugenia ahernia*na C.B. Rob. is found in the Philippines), whilst *Syzygium* is a strictly tropical Old World genus.

Although phenetic analysis of a partial dataset is not problematic – phylogenetic analysis of such data can only be justified if:

1) representatives cover the complete range of variability of the genus;

2) the apomorphies used are unique on a world-wide scale and not just to the study area in question (Esser et al., 1998);

3) a conservative approach to monophyletic group delimitation from any resulting cladogram is taken.

These criteria are met in this paper.

In the absence of contradictory evidence on monophyly, *Syzygium*, the more geographically confined and smaller of the two genera, appears the better genus to start analysis with and is the one dealt with here.

THE CONCEPT OF SYZYGIUM ADOPTED IN RECENT SOUTH-EAST ASIAN FLORAS

Historical precedent is a particular problem in South-East Asia where the genus *Syzygium* is centred. Ashton (1981) points out that "through the latter part of the 19th century a broad generic definition operated in Asiatic floras with all species being considered as part of *Eugenia*." Even nowadays some authors continue to utilise one or other generic name for the entire constellation of species (e.g. Kochummen, 1995) whilst others acknowledge that new combinations are needed but refuse to make them (Lemmens, 1995). Whilst such procedures promote nomenclatural stability they perpetuate the use of inaccurate names and make it psychologically more difficult to introduce the necessary new combinations.

Despite the strength of historical precedent there have been three concepts of the *Syzygium/Eugenia* complex adopted in recent South-East Asian floras.

Concept 1 — The simplest concept, and that which current evidence indicates is the most inaccurate, is the all-inclusive one used by Henderson (1949), Kochummen (1995) and Lemmens (1995) wherein a single enormous genus – *Eugenia* – is recognised. However, it is obvious from Schmid (1972) and Johnson & Briggs (1984) that this concept is an uninformative option, that *Eugenia* and *Syzygium* are distinguishable and therefore this concept has latterly been adopted merely for the sake of convenience. Henderson (1949) accepts convenience as a valid determinant of a classificatory structure and produces further evidence to suggest that *Eugenia* and *Syzygium* are not in reality distinct from each other. He also points out that such a single large genus is unwieldy and that "attempts to split it up have not met with conspicuous success"; he then proceeds to split *Eugenia* into five sections based on features of the calyx and androecium (Table 2).

Henderson's section *Fissicalyx* contains only two rare species, both endemic to Malaysia, and both of which have the same apomorphic characteristics – that their stamens are not attached to the rim of the calyx-tube but occur scattered over its sur-

Table 2. Sectional differences within *Eugenia* s.l. according to Henderson (1949). § = section.

Character	§ *Cleistocalyx*	§ *Syzygium*	§ *Acmena*	§ *Fissicalyx*	§ *Eu-eugenia*
Calyx-tube produced above ovary	Yes	Yes	Yes	Yes	No
Stamens marginal on disc lining calyx-tube	Yes	Yes	Yes	No	Yes
Anther-cells divaricate	No	No	Yes	No	No
Calyx calyptrate	Yes	No	No	No	No

Table 3. Henderson's (1949) division of his section *Syzygium* into groups. Henderson's calyx-tube characters include the pseudostipe: a character described in detail by Chantaranothai & Parnell (1994a).

Character	Group 1	Group 2	Group 3	Group 4	Group 5
Calyx-tube at least 1 cm long	Yes	Yes	No	No	Yes
Calyx-tube campanulate, funnel-shaped, urceolate or subglobose (1) or Calyx-tube fusiform, peg-shaped, clavate or gradually narrowed (2)	1	1	1	1	2
Inflorescence short	Yes	No	No	Usually no	Varied
Inflorescences few-flowered	Yes	No	No	No	Usually no

face and that the calyx-tube splits after anthesis into six or seven irregular segments. These differences are sufficiently great to suggest that Henderson was unduly conservative and that these species should, probably, be segregated into a separate genus. As Henderson's other four sections occur in Thailand phylogenetic analysis, as undertaken herein, is still justifiable. Sections *Cleistocalyx* and *Acmena* contain only their epithet bestowing genera and are separated by autapomorphies from the other sections (viz. calyx calyptrate and anthers divaricate, respectively). Section *Syzygium* contains the vast majority of species and was further subdivided by Henderson (1949) into five groups on the basis of calyx and inflorescence characters (Table 3).

Henderson's groups are entirely novel and do not correspond to previously recognised genera. Henderson was also the first author to use a biometric character (the length of the calyx-tube) to subdivide this complex. Other authors, for example Hung-ta & Ru-hwai (1982), divide the genus *Syzygium* into sections and series on the basis of differences in characters such as flower length, petals either free or fused, and inflorescence type. Unfortunately many species are too little known to allow them to be easily accommodated in their system.

Concept 2 — This second concept allows for separation of *Syzygium* from *Eugenia* but gives no recognition to any other generic segregates. It conforms to that of Alston (1931) (who, in his revision of Sri Lankan material, reduced *Jambosa* to synonymy with *Syzygium* whilst maintaining the split between *Syzygium* and *Eugenia*) and more or less to those of Ashton (1981), Hartley & Perry (1973), and Airy Shaw (1949).

Concept 3 — The third concept of the complex is the most divisive. In its most fissible form it was adopted by Merrill & Perry (1938, 1939) who accepted a large series of generic segregates including *Acmena* and *Cleistocalyx*. However, Merrill & Perry did not accept *Jambosa*: a genus or subgenus defined in a variety of ways [on the basis of its large leaves, flowers and fruits (Ridley, 1922)] or differing only in its centrifugal inflorescence (Bentham & Hooker, 1865). Similar concepts have been utilised by Kostermans (1981) and Hyland (1983). In addition, Merrill (1950a, b),

though not Perry (pers. comm. in Schmid, 1972), also accepted the further segregate *Jossinia*: a genus or subgenus distinguished from *Eugenia* largely by its broad staminiferous disc (Bentham & Hooker, 1865) as well as many seeded fruits and long radicle according to De Candolle (1828). Hô (1992) accepted aspects of Merrill & Perry's concept and recognised *Acmena, Cleistocalyx, Syzygium,* and *Eugenia.* None of these segregates were supported by Schmid (1972) who was unable to distinguish any of them.

Previous analyses of the *Eugenia/Syzygium* controversy have relied on alpha-taxonomic methods – little phenetic or phylogenetic analysis has been undertaken. This paper uses regional data fully representative of the *Syzygium* complex in a phenetic and phylogenetic analysis, with particular emphasis placed on the use of biometric characteristics and the question of homogeneity of *Syzygium.*

MATERIALS AND METHODS

Data were in part drawn from Chantaranothai (1989): further data were collected by re-examination of herbarium material cited in this work and of new material (Simpson et al., 1995, provides further details of these collections). The data used in this paper are based on examination of all the available material (c. 4,000 specimens) of these taxa from A, AAU, ABD, BK, BKF, BM, C, E, K, L, MEL, OXF, P, PDA, PSU, SING, TCD, U and on the basis of field-collections. There are 89 species in the *Acmena* and *Eugenia* alliances present in Thailand (Chantaranothai & Parnell, 1994a). Chantaranothai & Parnell (1994a), Parnell (1995), Parnell & Nic Lughadha (1992) and Parnell & Chantaranothai (1996) provide details of all of the taxa of *Syzygium* s.l. in Thailand and further details of the characters recorded for these species. A very few (5) species of *Syzygium* s.l., included in Chantaranothai & Parnell (1994a), are not dealt with because the available material was inadequate to allow them to be scored. Normally, for phylogenetic analysis such a procedure would be unacceptable, it being better to have representation of all taxa than all characters. However, as previously stated, this analysis uses a representative selection of species in the genus *Syzygium* which representation is not diminished by the above procedure. A list of the taxa included in this paper is given in Appendix Table 1.

Chantaranothai & Parnell (1994a) assigned these 84 species to genera as follows: 81 species to *Syzygium* s.s., 1 species to *Acmena* (autapomorphy, anthers divaricate), 4 to *Cleistocalyx* (autapomorphy, calyx calyptrate) and 3 to *Eugenia* (autapomorphy, inflorescence and/or vegetative parts hairy).

Of these genera only one, the most distinctive, *Acmena,* was defined as an outgroup for phylogenetic analysis, though representatives of *Eugenia* and *Cleistocalyx* were added as possible outgroups too.

ANALYSES UNDERTAKEN

A number of different types of analysis were undertaken. For morphometric phenetic analysis two fundamentally different techniques were used. Firstly the data were ordinated.

Principal Components Analysis (PCA) is designed for continuously varying or quantitative data or multistate data expressed on a scale of reasonable length. With most biological datasets PCA will extract as many summary axes as there are original variables in the data. However, only the first few axes represent effective summaries of the data. On the basis of Parnell & Needham (1998) only those components which satisfied Frontier's (1976) criterion as well as the first component which his scheme signals may be trivial but which still meets the heuristic Kaiser-Guttmann criterion were used. The programme used was Datadesk 5.0.1 (cf. Data Description Inc., Ithaca, New York).

Discriminant Analysis (DSC), was used to test whether pre-defined groups of species visualised on PCA were or were not statistically distinguishable. This multivariate extension of analysis of variance (Marriott, 1974) was performed in both stepwise and non-stepwise ways. Stepwise analysis was performed using all default options [minimisation of Wilks lambda, auto F-to-enter and F-to-remove in SPSS 6.1 (cf. Norusis & SPSS Inc., 1993)].

The following characters were measured for morphometric phenetic analysis (Table 4). In virtually all cases the characters were measured on from between one and four leaves and flowers from each of four separate specimens of the taxon (a list of most of the specimens measured is given in Chantaranothai, 1989). The values obtained per specimen were first averaged and then the average values per taxon were calculated: these latter values were then used in PCA and DSC. All leaf measurements were made on dried material and all flower measurements on flowers which had been soaked in water overnight or boiled for 10–15 minutes. It proved impossible to obtain a complete datamatrix which covered all Thai *Syzygium* species and because I do not like to interpolate data I decided to eliminate the taxa for which I had incomplete data from the analysis. The problems raised by this procedure for phylogenetic analysis have been discussed earlier. The remaining taxa are listed in Appendix Table 1.

Appendix Table 1.

Taxa included in this analysis. *A. acumenatissima, C. khaoaienensis, C. nigrans, C. operculatus, C. operculatus* subsp. *paniala, C. phengklaii, E. bracteata, E. xanthocarpa, S. abortivum, S. aksornii, S. albiflorum, S. anacardiifolium, S. angkai, S. angkae* subsp. *spissum, S. aqueum, S. aromaticum, S. attenuatum, S. attenuatum* subsp. *montanum, S. attenuatum* subsp. *circumcissum, S. balsameum, S. boiseanum* subsp. *longifolium, S. borneense* subsp. *myrtillus, S. borneense* subsp. *myrtillus, S. cacuminis, S. cacuminis* subsp. *inthanonense, S. cerasiforme, S. chavaran, S. cinereum, S. claviflorum, S. craibii, S. cuminii, S. diospyrifolium, S. duthieanum, S. dyerianum, S. fastigiatum, S. filiforme, S. formosum, S. foxworthianum, S. fruticosum, S. fuscescens, S. glaucum, S. globiflorum, S. grande, S. grande* subsp. *parviflorum, S. gratum, S. gratum* subsp. *confertum, S. helferi, S. hemsleyanum, S. hemsleyanum* subsp. *paucinervum, S. hulletianum, S. ixoroides, S. jambos, S. jasminifolium, S. kerrii, S. kurzii, S. laetum* subsp. *jugorum, S. laetum* subsp. *sublaetum, S. lakshanakarae, S. leptostemon, S. lineatum, S. maingayi, S. malaccense, S. megacarpum, S. mekongense, S. muelleri, S. nitrasarirakii, S. oblatum, S. oblatum* subsp. *laevicaule, S. pachyphyllum, S. pachysarcum, S. papillosum, S. pergmantaceum, S. polyanthum, S. praecox, S. praineanum* subsp. *minor, S. pseudoformosum, S. puttii, S. pycnanthum, S. pyrifolium, S. refertum, S. ridleyi, S. rigens, S. ripicola, S. samarangense, S. samarangense* subsp. *parviflorum, S. scortechinii, S. siamense, S. skiophilum, S. subrufum, S. subrufum* subsp. *smalianum, S. syzygiodes, S. tetragonum, S. thumra, S. thumra* subsp. *punctifolium, S. toddalioides, S. urophyllum, S. winitii, S. zeylanicum, S. zimmermannii.*

Table 4. Morphological characters measured for phenetic, morphometric analysis.

Character number	Character	Comments
1	Petiole length	The length of the petiole from the base of the petiole to the base of the lamina
2	Lamina width	Measured at right-angles to the axis at the point of maximum width
3	Lamina length	The length as measured along the midrib
4	Basal angle	The angle formed at the base of the leaves
5	Apical angle	The angle formed at the apex of the leaves
6	Midrib width	Measured at the point midway between the apex and base of the leaf
7	Number of secondary veins	Counted upwards from the second pair of veins at the base of the lamina
8	Distance between the secondary veins	Measured at the point midway between the midrib and the margin in a line parallel to the midrib
9	Number of intramarginal veins	Counted at the base of the lamina
10	Distance between the intramarginal veins and the margin	Measured at right angles to the margin at a point midway between the leaf base and leaf apex. If the leaf had 2 or 3 intramarginal veins then the innermost distance was measured. If the species had no intramarginal veins then the distance between the secondary veins and the margin was measured instead
11	Calyx-tube length	Measured from the base of the calyx-lobes to the base of the calyx-tube
12	Number of calyx-lobes	
13	Length of calyx-lobes	Measured along the mid-axis of the calyx-lobe
14	Number of petals	
15	Petal length	Measured along the long axis of the petal at its longest point
16	Outer stamen length	Measured from the junction of the calyx-tube to the tip of the anther
17	Number of glands on the back of the anther	
18	Style length	Measured from its junction with the calyx-tube to the tip of the stigma
19	Number of locules	
20	Number of ovules per locule	
21	Number of gland-dots per petal	

Initial analysis showed that characters 17 and 19 (Table 4) were insufficiently variable to allow for inclusion in a PCA or DCA; they were therefore not directly included in these analyses. Initial analysis also showed that characters 4, 8, 10, 13, 15, 16, 18, 20 and 21 were not normally distributed on an individual basis. (These variables were either markedly skewed, had abnormal kurtosis or did not fall along a straight line on a normal probability plot.) Transformation of the above characters was undertaken

following Manly (1986) and Marriott (1974). By trial and error the most appropriate transformation for all these characters, except for character 8, was found to be Log_{10}: character 8 required a reciprocal square root transformation. These transformations contracted the range of the transformed variables, bunching them closer together than in their untransformed state. (Initial analysis showed that PCA and DSC analyses run on an untransformed dataset produced an almost identical pattern of results as that run with the above transformations included, though the gap between Group 1 and 2 appeared much larger).

PCA (and DSC) require data which are approximately multivariate normal and are therefore inappropriate for categorical data (Manly, 1986). Suitable ordination methods are detailed in Causton (1988) and their advantages and disadvantages outlined in Kent & Coker (1992) who indicate that Detrended Correspondence Analysis (DCA) is probably the most widely used technique. Its derivation, application and possible flaws are further detailed in Ter Braak (1988), Hill (1973, 1979), Hill & Smith (1976), Minchin (1987), Oksanen & Minchin (1997) and Parnell & Waldren (1996). Parnell & Waldren (1996) show that DCA has wide application to taxonomic bistate data providing simple, clear plots of easy interpretability where both objects and characters are plotted to the same scale and are overlayable. Furthermore DCA can be used where the group under consideration does not form a complete phylogenetic unit. DCA was performed using the DCA option of CANOCO (Ter Braak, 1988). Although DCA can utilise mixed datasets, where some data are recorded on one scale and other data are recorded on another, it is clear that this practice results in ordinations which are very difficult to interpret. It is much preferable to utilise datamatrices where the data are all scored to the same scale. Also DCA is derived from CA: the latter technique preserves the χ^2 distances between objects, and hence ignores comparisons between values with double 0 ranks (Legendre & Legendre, 1983). So if 0 was of frequent occurrence plots could result with the intertaxon or intercharacter differences incorrectly scaled, i.e. characters should not bear a 0 code if it is avoidable. Characters were coded as possessing one of two alternative states (1) or (39), the latter corresponding to the range of the number of characters input (Parnell & Waldren, 1996). Table 5 lists the characters and their coding for this analysis. The same taxa were surveyed for this analysis as for PCA (Appendix Table 1). Very few morphometric characters were used in DCA. Those that were included were decomposed into naturally occurring groups: for example it was clear from examination of histograms of the raw data that a naturally occurring break in terms of stamen length is that between those flowers with outer stamens less than 15 mm and those with outer stamens 15 mm or more in length: this break then allowed these two states to be coded for DCA.

It is obviously impossible to express characters where more than a single natural break occurs in a single pair of dichotomous values. There is no perfect solution to this problem, though there are three potential approaches. Firstly, to introduce subdivisions within the coded range for each character (i.e., for a four state character, codes would be 1, 11, 22 and 34). Unfortunately, this procedure is inadvisable as implementations of DCA appear not to be absolutely scale independent. Secondly, to treat each potentially multistate character as a number of independent dichotomous characters – therefore a four state character is coded as four separate bistate characters.

Table 5. Characters recorded for weighted Detrended Correspondence Analysis (DCA). Characters were scored as one of two states, and coded as (1) or (39) and are given in that order below; the slash (/) is used to divide the different character states.

Character number	Character states
1	Trees / Shrubs
2	Plant without hairs / plant hairy
3	Midvein impressed or sunken / midvein not impressed or sunken
4	Midvein not raised with a channel in the middle of the vein / midvein raised with a channel in the middle of the vein
5	Midrib not keeled / midrib keeled
6	Intramarginal veins 0 / intramarginal veins not 0
7	Intramarginal veins 1 / intramarginal veins not 1
8	Intramarginal veins 2 / intramarginal veins not 2
9	Intramarginal veins 3 / intramarginal veins not 3
10	Midvein less than 1 mm wide / midvein more than 1 mm wide
11	Number of secondary veins 5–10 / number of secondary veins not 5–10
12	Number of secondary veins 11–15 / number of secondary veins not 11–15
13	Number of secondary veins 16–20 / number of secondary veins not 16–20
14	Number of secondary veins 21–25 / number of secondary veins not 21–25
15	Number of secondary veins 26–30 / number of secondary veins not 26–30
16	Calyx-tube narrowly funnel-shaped and contracted into a distinct pseudostipe / calyx-tube not narrowly funnel-shaped and contracted into a distinct pseudostipe
17	Calyx-tube broadly funnel-shaped and narrowed into a pseudostipe / calyx-tube not broadly funnel-shaped and narrowed into a pseudostipe
18	Calyx-tube clavate or narrowly funnel-shaped and narrowed into a long pseudostipe / calyx-tube not clavate or narrowly funnel-shaped and narrowed into a long pseudostipe
19	Calyx-tube cylindrical / calyx-tube not cylindrical
20	Calyx-tube broadly funnel-shaped and more than 10 mm long / calyx-tube not broadly funnel-shaped and more than 10 mm long
21	Calyx-tube funnel-shaped and less than 10 mm long / calyx-tube not funnel-shaped and less than 10 mm long
22	Calyx-tube funnel-shaped without a definite pseudostipe or with a pseudostipe less than 1 mm long / calyx-tube not funnel-shaped without a definite pseudostipe or with a pseudostipe less than 1 mm long
23	Pseudostipe absent / pseudostipe present
24	Number of gland-dots on a petal 0–50 / number of gland-dots on a petal not 0–50
25	Number of gland-dots on a petal 51–100 / number of gland-dots on a petal not 51–100
26	Number of gland-dots on a petal 101–150 / number of gland-dots on a petal not 101–150
27	Number of gland-dots on a petal 151–200 / number of gland-dots on a petal not 151–200
28	Number of gland-dots on a petal > 200 / number of gland-dots on a petal not > 200
29	Number of calyx-lobes 0 or 4 / number of calyx-lobes 5 or more
30	Number of petal-lobes 4 / number of petal-lobes 5 or more
31	Petal colour not polymorphic (always yellow) / petal colour polymorphism present (not always yellow)
32	Stamen length 0–14.9 mm / stamen length 15 mm or more
33	Style length approximately equal to outer stamen length / style length less than outer stamen length
34	Bracts absent or narrow, oblong or spathulate or linear or sheathing / bracts triangular or ovate
35	Two locules per ovary / sometimes more than two locules per ovary
36	Number of ovules per locule 5–10 / number of ovules per locule not 5–10
37	Number of ovules per locule 11–15 / number of ovules per locule not 11–15
38	Number of ovules per locule 16–20 / number of ovules per locule not 16–20
39	Number of ovules per locule > 20 / number of ovules per locule not > 20

This procedure has the disadvantage that it may overweight that particular character with respect to others in the datamatrix. Finally, the data can be broken down into the appropriate number of applicable states for each character and then re-aggregated within each coded character so that eventually there is one character less than the original number of states. So, for example, each state of a four state character would appear 3 times in the datamatrix in each of three separate characters. This procedure has the advantage that the weighting introduced by full dichotomisation is reduced but has the disadvantage that it eliminates the strict independency of the characters. Strict dichotomisation (sometimes called nominal variable coding) was used throughout (though a trial analysis of all three methods produced similar ordination patterns). The characters recorded for DCA are shown in Table 5.

As can be seen from Table 5 a number of characters are included (11–15, 24–28 and 36–39) which are effectively aggregated states of what might be considered a semi-continuous or multistate variable. The divisions chosen for these variables were based on study of frequency distribution histograms which indicated relevant inflexion points or gaps over the range of values for the characters chosen: these points were then used as interstate boundaries.

It is not possible to use the data from the characters listed in the above table directly in a cladistic or phylogenetic analysis. It is obvious from Table 5 that some of the characters used in the DCA analysis could be considered as decomposed (= dichotomised) versions of what could be coded as multistate characters (e.g. characters 6, 7 and 8). As previously indicated such decomposition of multistate characters is essential for DCA analysis. However, it is clear that this procedure leads to an increase in the step length per character in the parsimony results (Van Welzen & Esser, 1997). Therefore it was decided to adopt the most rigorous view possible of the homoplasy of the characters involved. This meant, for example, that it was assumed that the number of intramarginal veins (0, 1, 2, 3) were thought to be different expressions of the same character when in fact there is no evidence that they are or that they are not: similar arguments apply to other characters – for example to the different shapes of the calyx-tube. The character set utilised for phylogenetic analysis is similar to that used by DCA except for the inclusion of characters 2, 10 and 23 (Table 6) – twigs terete or angled, number of flowers per inflorescence and bract shape, respectively. These latter characters were not included in DCA as they were too variable to be successfully coded for this type of analysis. As can be seen from Table 6 the number of characters used for phylogenetic analysis is small. However, consideration of the species involved has not revealed any more widely available characters and the use of artificial or derived characters (through decomposition of multistate characters or the construction of ratios) appears unjustifiable. Three outgroups were included in the analysis – *Acmena, Eugenia,* and *Cleistocalyx*. All characters were coded as unordered. Phylogenetic analysis was performed using PAUP (Swofford, 1993). Uncertainties were coded as ? and polymorphisms as { }. Relative Apparent Synapomorphy Analysis (RASA) (Lyons-Weiler et al., 1996; Lyons-Weiler, 1997) showed that there was a strong phylogenetic signal present in the decomposed, binary DCA datamatrix (t_{RASA} = 35.8; df = 2481; $p < 0.001$).

Table 6. Characters coded for phylogenetic analysis. Bistate characters were coded as either 0 or 1, respectively, the slash (/) is used to divide these different character states. The values used for multistate characters are given as bracketed values below.

Character number	Character states
1	Trees / Shrubs
2	Twigs terete (0) / twigs angled (1) / twigs compressed (2)
3	Plant without hairs / plant hairy
4	Leaf venation with one or more pairs of intramarginal veins extending from the base of the lamina and close and parallel to the leaf margin (0), secondary veins arching upwards to form a false intramarginal vein, true intramarginal veins absent (1), true and false intramarginal veins absent, though ultimate branches of secondary veins sometimes rejoining (2)
5	Midvein impressed or sunken (0) / midvein flat or raised (1) / midvein raised with a channel in the middle of the vein (2)
6	Midrib keeled / midrib not keeled
7	Intramarginal veins 0 (0) / intramarginal veins 1 (1) / intramarginal veins 2 (2) / intramarginal veins 3 (3)
8	Midvein less than 1 mm wide / midvein more than 1 mm wide
9	Number of secondary veins 5–10 (0) / number of secondary veins 11–15 (1) / number of secondary veins 16–20 (2) / number of secondary veins 21–25 (3) / number of secondary veins 26–30 (4)
10	Number of flowers per inflorescence 1 (0) / number of flowers per inflorescence 3 (1) / number of flowers per inflorescence > 3 (2) / number of flowers per inflorescence 2 (3)
11	Calyx-tube narrowly funnel-shaped and contracted into a distinct pseudostipe (0) / calyx-tube broadly funnel-shaped and narrowed into a pseudostipe (1) / calyx-tube clavate or narrowly funnel-shaped and narrowed into a long pseudostipe (2) / calyx-tube cylindrical (3) / calyx-tube broadly funnel-shaped and more than 10 mm long (4) / calyx-tube funnel-shaped and less than 10 mm long (5) / calyx-tube funnel-shaped without a definite pseudostipe or with a pseudostipe less than 1 mm long (6)
12	Pseudostipe present / pseudostipe absent
13	Number of gland-dots on a petal 0–50 (0) / number of gland-dots on a petal 51–100 (1) / number of gland-dots on a petal 101–150 (2) / number of gland-dots on a petal 151–200 (3) / number of gland-dots on a petal > 200 (4)
14	Anthers divaricate / anthers not divaricate
15	Anthers opening apically / anthers not opening apically
16	Seed coat free / seed coat not free (coded on the basis of the literature)
17	Embryo undivided / embryo divided (coded on the basis of the literature)
18	Number of calyx-lobes 4 (0) / number of calyx-lobes 5 (1) / number of calyx-lobes > 5 (2)
19	Number of petal lobes 4 (0) / number of petal lobes 5 (1) / number of petal lobes > 5 (2)
20	Petal colour not polymorphic (always yellow) / petal colour polymorphism present (not always yellow)
21	Stamen length 0–5 mm (0) / stamen length 5.1–9.9 mm (1) / stamen length 10–20 mm (2) / stamen length 20.1–30 mm (3) / stamen length > 30.1 mm (4)
22	Style length approximately equal to outer stamen length (0) / style length less than outer stamen length (1) / style length greater than outer stamen length (2)
23	Bracts absent (0) / bracts triangular or ovate (1) / bracts narrow, oblong, spathulate, or linear (2) / bracts membranous, enclosing the inflorescence (3)
24	Two locules per ovary / sometimes more than two locules per ovary
25	Number of ovules per locule 5–10 (0) / number of ovules per locule 11–15 (1) / number of ovules per locule 16–20 (2) / number of ovules per locule > 20 (3)

RESULTS

(i) *Phenetic analyses*

Two types of phenetic analysis were undertaken.

1) Continuous data: PCA & DSC — The PCA of the raw dataset covering the whole of Thai *Syzygium* s.l. resulted in two significant axes which met the 'a priori' criteria for significance (see above) being extracted: these had eigenvalues of 7.4 and 2.5 and the associated percentage of the total variance they accounted for was 39% and 13%, respectively. A plot of the species scores on these two axes is presented in Fig. 1a and the associated eigenvectors in Table 7.

As can be seen from Fig. 1a it is clear that there is no biometric support for separation of any segregate genera from *Syzygium* s.s. However, this is unsurprising as these genera were segregated by Chantaranothai & Parnell (1994a) on the basis of bistate characteristics which cannot be and were not included in this PCA analysis. Therefore the PCA analysis was re-run excluding the above segregate genera. The results of this analysis are shown in Fig. 1b, 1c & 1d; the appropriate eigenvectors are in Table 8. Again, only the first two axes extracted were found to be significant accounting for 53% of the variance in combination (eigenvalues of 7.8 & 2.3, respectively). A com-

Table 7. Eigenvectors for the first two principal component axes extracted for the analysis Thai *Syzygium* s.l. Character numbers and the reasons for transformation of some are given in Table 4 and associated discussion.

Character number	Characters	Eigenvectors for Axis 1	Eigenvectors for Axis 2
1	Petiole length	−0.039	0.436
2	Lamina width	−0.272	0.237
3	Lamina length	−0.295	0.154
4	Log_{10} angle at base of leaf	−0.168	−0.210
5	Apical angle	−0.044	0.113
6	Midrib width	−0.226	0.269
7	Number of secondary veins	0.063	−0.085
8	$1/\sqrt{}$Distance between veins	0.286	−0.220
9	Number of intramarginal veins	−0.037	−0.251
10	Log_{10} distance between the intramarginal veins and the margin	−0.289	0.167
11	Calyx-tube length	−0.263	−0.289
12	Number of calyx-lobes	0.080	−0.365
13	Log_{10} length of calyx-lobes	−0.308	0.004
14	Number of petals	0.078	−0.255
15	Log_{10} petal length	−0.325	−0.154
16	Log_{10} stamen length	−0.298	−0.231
18	Log_{10} style length	−0.270	−0.283
20	Log_{10} number of ovules per locule	−0.284	−0.057
21	Log_{10} number of gland-dots per petal	−0.248	−0.111

Table 8. Eigenvectors for the first two principal component axes extracted for the analysis Thai *Syzygium* s.s. Character numbers and the reasons for transformation of some are given in Table 4 and associated discussion.

Character number	Characters	Eigenvectors for Axis 1	Eigenvectors for Axis 2
1	Petiole length	−0.015	0.409
2	Lamina width	−0.266	0.287
3	Lamina length	−0.292	0.169
4	Log_{10} angle at base of leaf	−0.170	−0.169
5	Apical angle	−0.048	0.142
6	Midrib width	−0.261	0.251
7	Number of secondary veins	0.053	0.019
8	$1/\sqrt{}$Distance between veins	0.277	−0.202
9	Number of intramarginal veins	−0.069	−0.061
10	Log_{10} distance between the intramarginal veins and the margin	−0.285	0.225
11	Calyx-tube length	−0.265	−0.278
12	Number of calyx-lobes	0.105	−0.354
13	Log_{10} length of calyx-lobes	−0.305	−0.092
14	Number of petals	0.082	−0.338
15	Log_{10} petal length	−0.321	−0.134
16	Log_{10} stamen length	−0.295	−0.231
18	Log_{10} style length	−0.268	−0.286
20	Log_{10} number of ovules per locule	−0.278	−0.086
21	Log_{10} number of gland-dots per petal	−0.241	−0.170

parison of Tables 7 and 8 shows that the elimination of the segregate genera *Acmena*, *Cleistocalyx* and *Eugenia* affected the eigenvector scores but did not significantly alter the relative importance of the characters for each axis. So, for example, characters number 1, 5, 7, 9 and 14 (petiole length, apical angle, number of secondary veins, number of intramarginal veins and number of petals) are low scoring (and hence unimportant) contributors to Axis 1 in both analyses and characters number 7, 9 and 20 (the number of secondary veins, the number of intramarginal veins and Log_{10} number of ovules per locule) are insignificant in both analyses for Axis 2. Therefore discussion of the results of the PCA for *Syzygium* s.s. is applicable, with only slight modification to *Syzygium* s.l.

It is also clear from Table 8 that many characters have similar absolute eigenvector scores on Axes 1 and 2 and that no single character stands out. The most important characters on Axis 1 for the PCA of *Syzygium* s.s. (Table 8) are characters number 3, 10, 13, 15 and 16 (lamina length, Log_{10} distance between the intramarginal veins and the margin, Log_{10} length of calyx-lobes, Log_{10} petal length, Log_{10} stamen length); for Axis 2 characters number 1, 12 and 14 (petiole length, number of calyx-lobes and number of petals) score strongly. It is clear that Axes 1 and 2 are determined, to some extent, by characters related to size; species whose organs are larger or more numerous

are located on the left-hand (negatively scoring) side of Fig. 1b, 1c & 1d. I have chosen to separate out plants on the left-hand side of this plot (Fig. 1b) as constituting a distinct group from those on the right-hand side. It could be argued that these groups are artificial and defined on the basis of size. However, this would not be entirely accurate as, for example, plants on the left-hand side of these plots also tend to have leaves whose apex and base are relatively obtuse and, as they do not score particularly strongly on Axis 2, these plants do not have petioles as long proportionately as might be expected from their larger leaves. In summary, plants on the left-hand side of the plots in Fig. 1 have larger, broader leaves with a larger midrib, secondary veins which are relatively distant from each other, intramarginal veins well-removed from the margin, larger calyces, petals (and these have a larger number of gland-dots), ovules, and longer stamens and styles. In addition the left-hand group (Fig. 1) contains all the species which have more than 1 gland-dot on the anther (a character excluded from PCA as it was too invariable) and none of the species which have more than 5 petals. Other data can be added to the basic plot of Fig. 1a. Fig. 1b shows the same PCA analysis as Fig. 1a but with the superimposition of an extra characteristic – all species in Thailand which occur only above 1,000 m are highlighted. To these highlighted data are added in Fig. 1c those species which are endemic to Thailand. As can be seen from both these figures, species endemic to Thailand or only found in highland areas are all located on the right-hand side of the plot; none occur in the left-hand group indicated on these plots. For the sake of convenience the group on the left-hand side of Fig. 1 is henceforth referred to as Group 2 and that on the right as Group 1. The species which occur in the smaller group, Group 2, are given in Table 9.

Leaving aside the desirability of using distinctions based on size to erect categorical distinctions, it is not inarguably clear on the basis of the plots shown in Fig. 1b, 1c & 1d that there is a clear morphological discontinuity between species of *Syzygium* in Groups 1 and 2. However, DSC analysis allows testing of the differences between the groups. Stepwise DSC resulted in the two groups being distinguished from each other with a F-value of 99.9 with 3 and 87 degrees of freedom (alternatively $\chi^2 = 131$ with three degrees of freedom for a test of the null hypothesis that there is no difference between the group means). Therefore the groups are highly significantly different from each other at > 99.9% confidence. Stepwise DSC utilised only three morphometric characters to maximise group distinctions viz. characters number 1, 3 and 4 (petiole length, lamina length and the Log_{10} angle at base of leaf). It is therefore possible to

Fig. 1. Plot of species scores for the first two Principal Component Axes extracted in two separate analyses (*Syzygium* s.l. is shown in Fig. 1a and *Syzygium* s.s. in Fig. 1b, 1c & 1d). U1 and U2 are Axes 1 and 2, respectively. For Fig. 1a the first two Principal Component Axes together account for 47.9% of the total variance: the symbols used in Fig. 1a are as follows: O = *Syzygium* s.s. sp.; x = *Cleistocalyx* sp.; \ = *Eugenia* sp.; – = *Acmena* sp. For Fig. 1b, 1c & 1d the first two Principal Component Axes account for 52.7% of the total variance. Also in these figures O is used to indicate Group 1 and ● to indicate Group 2. In Fig. 1b the position of three species discussed further in the text, *S. kurzii, S. lakshanakarae* and *S. zimmermanii,* are indicated by the first letters of their specific name. In Fig. 1c \ is used to indicate those species which occur only above 1,000 m and in Fig. 1d to indicate species which occur only above 1,000 m or are endemic to Thailand. Species in Group 2 are listed in Table 9 and those in Groups 1 and 2 in Appendix Table 1.

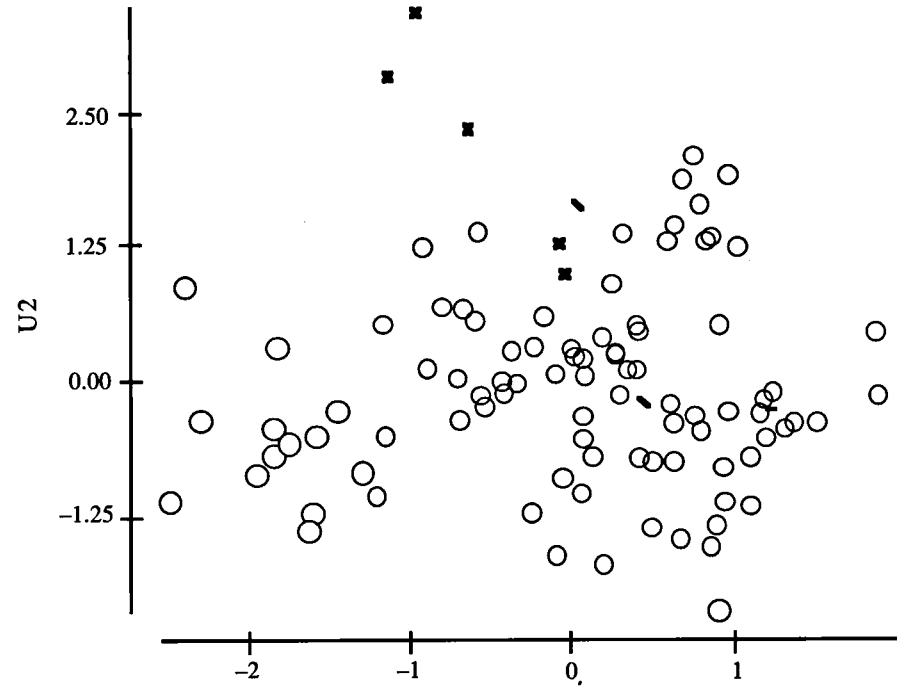


Fig. 1a

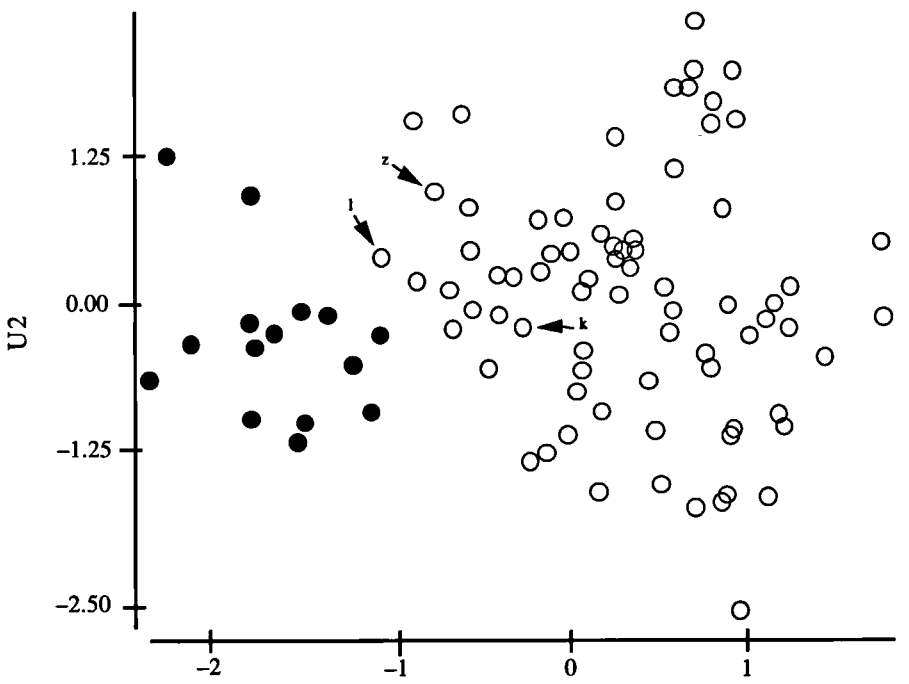


Fig. 1b

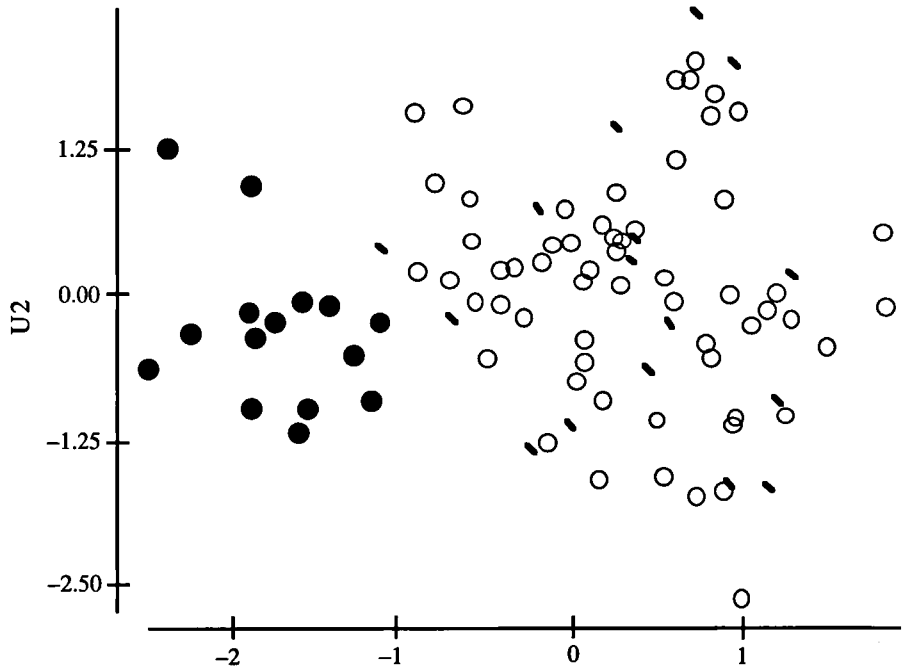


Fig. 1c

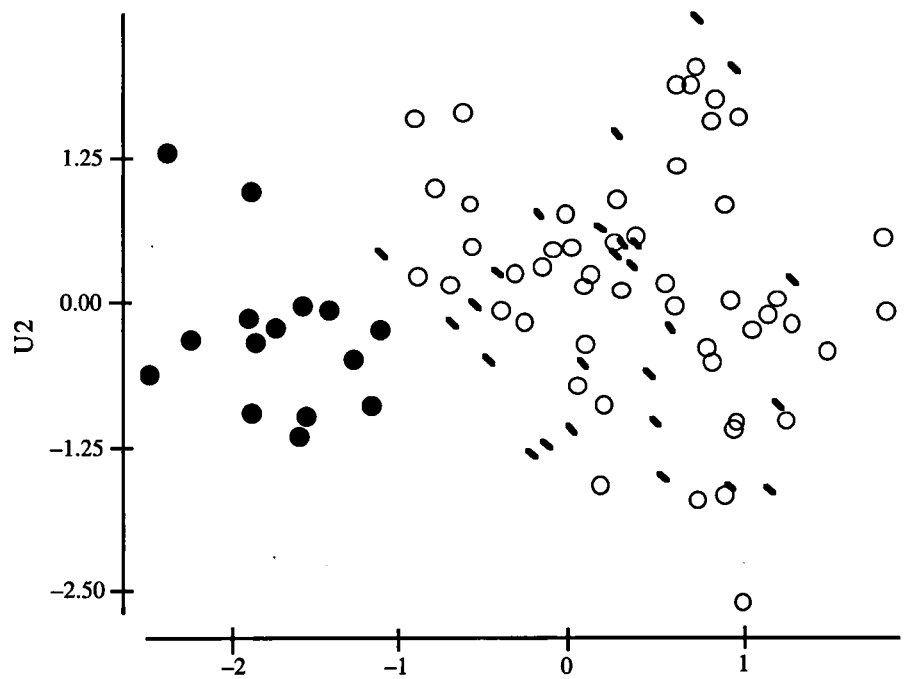


Fig. 1d

Table 9. Species placed in Group 2 as seen on PCA and DCA plots (Fig. 1b, 1c, 1d & Fig. 2). Species in Group 2 are all those listed under *Syzygium* s.s. in Appendix Table 1 and which are not listed below. + indicates a definite indication of membership of Group 2 on the basis of a particular analysis; ? questionable membership of that group; and – no evidence for membership of that group.

Species in Group 2	PCA	DCA
S. anacardiifolium	+	+
S. aqueum	+	?
S. diospyrifolium	+	+
S. formosum	+	+
S. foxworthianum	+	?
S. jambos	+	+
S. lakshanakarae	–	+
S. malaccense	+	+
S. megacarpum	+	+
S. papillosum	+	–
S. pseudoformosum	+	+
S. pycnanthum	+	+
S. samarangense	+	+
S. scortechinii	+	+
S. siamense	+	+

distinguish the two groups designated on PCA by measurement of these three characteristics alone. Stepwise DSC gives the following standardised discriminant function coefficients which allow determination of group membership for any taxa not in the analysis through multiplication of the appropriate coefficients by the appropriate variable (Table 10). Taxa with a total score above 2 fall into Group 2 and taxa scoring below 2 into Group 1.

Unfortunately, though stepwise analysis allows for efficient discrimination between the groups, it does not allow a picture to be built-up of the overall morphological differences between the groups. A non-stepwise or simultaneous analysis of the same datamatrix allowed construction of Fig. 2 which presents the means for each variable together with their 99.9% confidence limits: note that limits for transformed data are naturally asymmetric (Sokal & Rohlf, 1995). Fig. 2 is constructed such that it is safe

Table 10. Standardised and non-standardised discriminant function co-efficients for the DSC analysis of *Syzygium* s.s.

Character	Standardised	Non-standardised
Petiole length	0.796	0.304
Lamina length	0.388	0.009
Log_{10} angle at base of leaf	0.546	3.002
Constant		–7.97

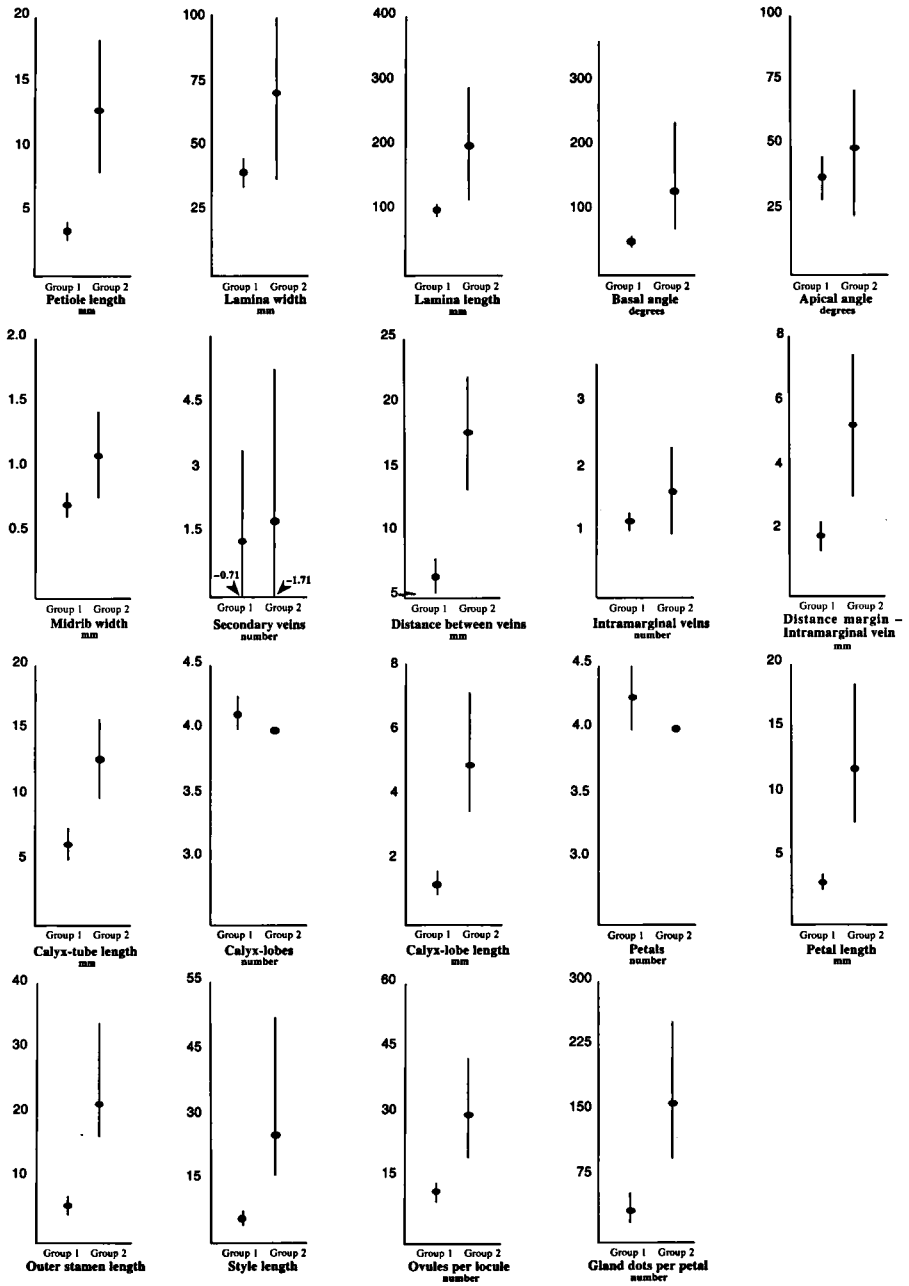


Fig. 2. Mean values of morphological characteristics for Groups 1 and 2 of Thai *Syzygium* s.s., together with 99.9% confidence limits.

to say that all variables with visibly non-overlapping confidence limits are significantly different from each other over the two groups at least 99.9% confidence. In fact the confidence limits overlap for 5 characters only (number 5, 7, 9, 12 and 14 of Tables 7 and 8): viz apical angle, number of secondary veins, number of intramarginal veins, number of calyx-lobes and number of petals. The visible differences between the other variables are variable but all pairs are significantly different from each other with 99.9% confidence: indeed the lowest univariate F-statistic for between group difference for these other variables is 24.8 (with 1 and 89 degrees of freedom) for character number 22 (gland-dot number per petal). From Fig. 2 it is clear that species in Group 1 differ from those in Group 2 in that they have smaller, narrower, more shortly petioled leaves with a narrower midrib, secondary veins which are relatively distant from each other, intramarginal veins close to the margin, smaller calyces, petals (and these have fewer gland-dots), ovules and shorter stamens and styles. Simultaneous DSC analysis yielded a slightly more significant test of inter-group mean difference than had been achieved through the stepwise analysis ($\chi^2 = 142$ with 19 degrees of freedom for a test of the null hypothesis that there is no difference between the group means).

2) Binary data: DCA — DCA analysis provided further confirmation that the two groups seen on PCA and confirmed by DSC are distinguishable. Table 9 presents a comparison of the species separated into Group 2 on DCA and PCA. As can be seen from this table the species separated out on DCA are almost identical to those separated out by PCA with the addition of one taxon (*S. lakshanakarae*) and the exception of *S. aqueum* and *S. papillosum.* The characters which DCA highlights as separating these species from the rest are characters 28, 31 and 32 of Table 5 (Fig. 3b). Therefore plants in Group 2 have usually more than 200 gland-dots per petal, colour polymorphism in the corolla, and stamens more than 15 mm long. In a similar manner to PCA DCA has separated plants placed into Group 2, on the basis of some size-related characters and some which are not size-related. The fact that DCA has isolated a similar group of species to PCA, but on the basis of a radically different dataset, strengthens the argument that this group is in fact distinct. A comparison of the plots of species from the two analyses (Fig. 1b, 3a) shows that there are three differences in the composition of species placed in Group 2 by these analyses (Table 9). *Syzygium aqueum* and *S. papillosum* are placed in Group 2 by PCA but is omitted from this group on DCA, though it is in fact spatially very close to Group 2 (arrowed Fig. 3a); the principle reason for the spatial separation of these species from the rest of Group 2 in the DCA plot is that

Fig. 3. DCA of Thai *Syzygium* s.l. The first plot (Fig. 3a) is the species plot for Axis 1 against 2 whilst the second plot (Fig. 3b) is the plot of characters. Axes 1 and 2 together account for at least 23% of the variance. Following from Fig. 1b two groups of species are distinguished from each other in these plots. The position of individual members of Group 1 is indicated by the first four letters of each species' name and this group's points are indicated in black. The groups are identical to those seen in Fig. 1b with the exception of *S. aqueum* and *S. papillosum* which appear not to belong to this group on the basis of this analysis and the peripheral position of *S. kurzii* and *S. zimmermanii* which are shaded in the figure. In Fig. 3b the characters which contribute to the separation of Group 1 from the rest are highlighted in black and are further discussed in the text, numbering as per Table 5.

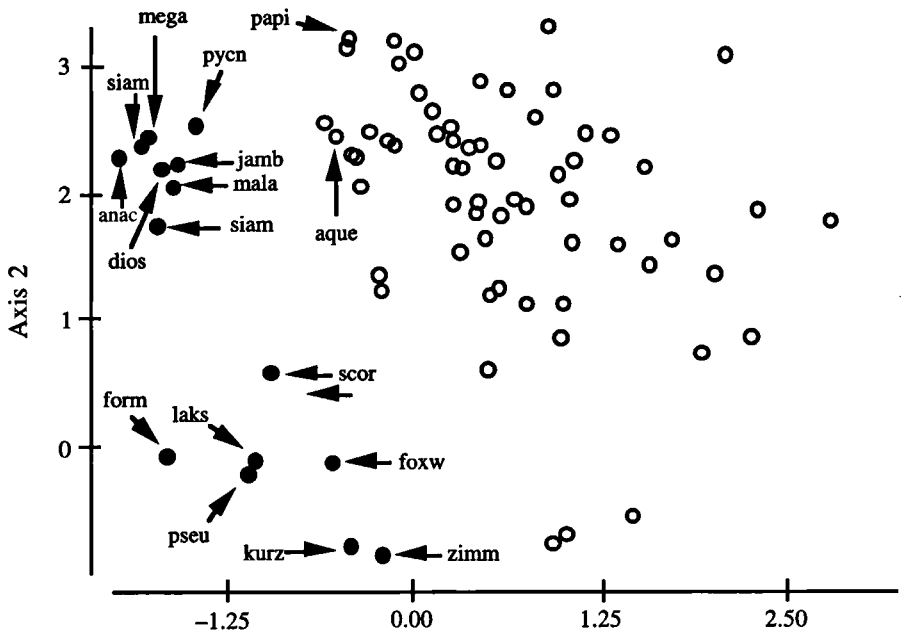

mega
papi
pycn
siam
jamb
mala
anac
siam
aque
dios
scor
form
laks
foxw
pseu
kurz
zimm
Axis 2
Axis 1
3
2
1
0
-1.25
0.00
1.25
2.50

Fig. 3a

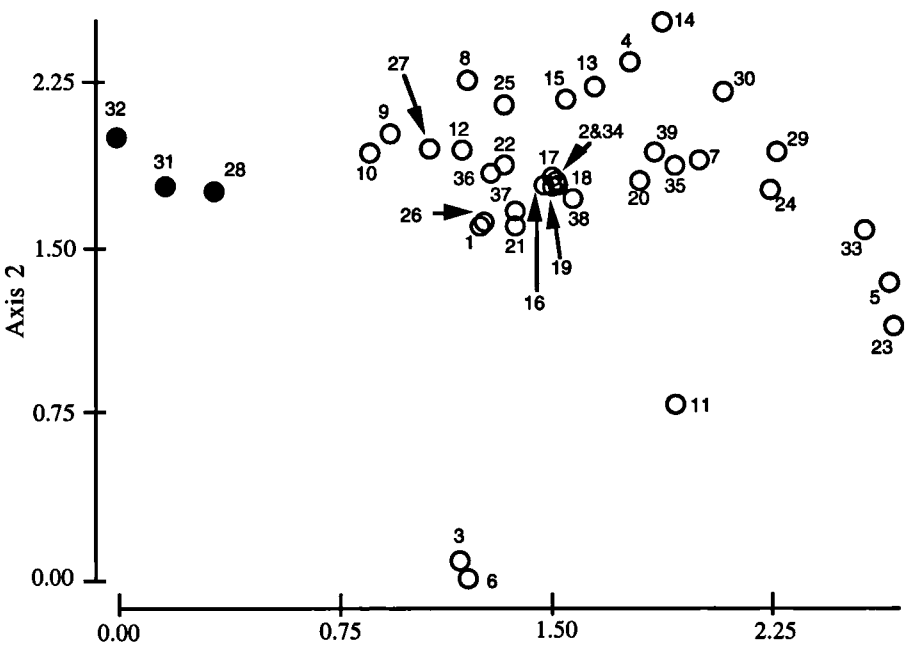

Axis 2
Axis 1
2.25
1.50
0.75
0.00
0.00
0.75
1.50
2.25
32
31
28
27
8
25
15
13
4
14
30
9
12
22
2&34
39
29
10
36
17
18
7
37
35
24
26
38
20
1
21
19
16
33
5
23
11
3
6

Fig. 3b

they have rather few gland-dots per petal. *Syzygium lakshanakarae,* omitted from Group 2 on the basis of PCA but therein included on the basis of DCA, is not a peripheral inclusion to Group 2 but placed near the core of the group, possessing both long stamens and highly glandular petals. In addition it is worth noting that *S. kurzii* and *S. zimmermanii* (shaded in Fig. 3a) are peripheral to Group 2, on DCA having been so placed largely on the basis of having a combination of the absence of an intramarginal vein and having a midrib which is impressed or sunken; they share almost no other common features with other species in Group 2. Overall there appears to be little case for their inclusion in Group 2. As the DCA plot accounts for rather less of the variance than the PCA plot and as Fig. 3b shows rather few characteristics as being involved in the creation of Group 2 on DCA I feel that much more importance should be attached to the PCA plot than to the DCA plot. In other words, *S. lakshanakarae* should not be included in Group 2 solely on the basis of the DCA nor should *S. aqueum* and *S. papillosum* be excluded.

Summary of phenetic analysis

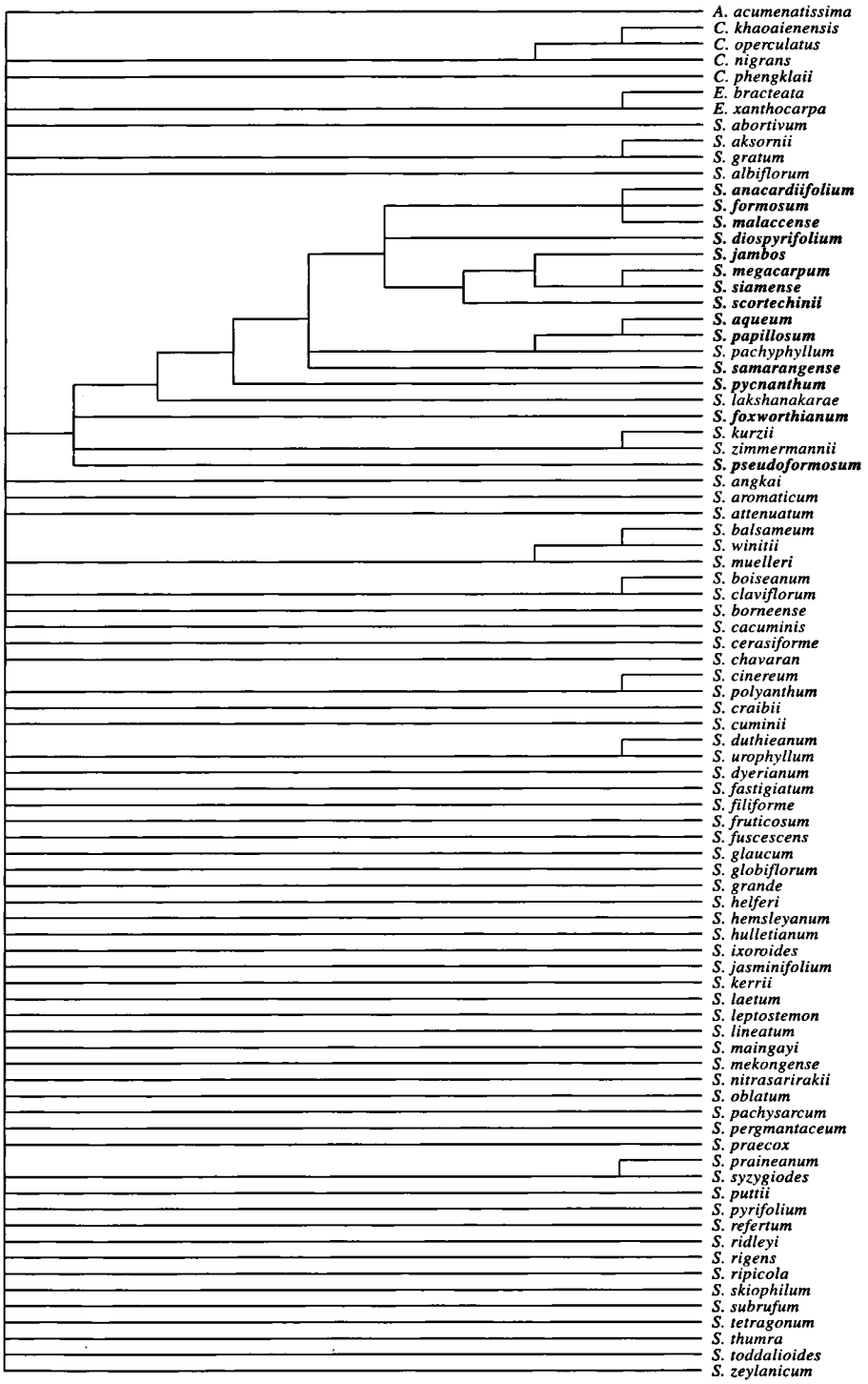
DCA confirmed the results of PCA. Both suggest that there is a well-recognised group of species separable from all the others in Thai *Syzygium.* The composition of the group is not unequivocal, as there are some species which DCA indicates should be placed in Group 1 which PCA (and DSC) does not so indicate. This is not an unexpected result as the analyses use different character matrices. Evidently, phenetic analyses such as PCA, DSC and DCA may be unable to produce absolute, universal, boundaries between species of *Syzygium* in Thailand; rather they strongly suggest that divisions do exist and DSC allows unknown species to be placed in one or another group with ease. Further datasets must be added to these data to confirm or disprove the reality of the divisions hinted at in the above analyses.

(ii) *Phylogenetic analysis*

RASA analysis of the PAUP datamatrix showed a significant phylogenetic pattern as being present ($t_{RASA} = 29$; df = 3,317; $p < 0.001$). The data were firstly heuristically analysed using 100 replicates of random addition, with TBR branch-swopping, steepest decent not in effect and with MAXTREES set to 2,000. The 2,000 trees were then subjected to SPR branch swopping under the same conditions but with MAXTREES set to 5,000; all trees obtained were of length 189 steps. This process was repeated seven times and no trees shorter than 189 step were found.

This approach, rather than a fully exhaustive heuristic search, seems the best one to adopt. This is because Chase et al. (1997) has suggested that it may be best, for large datasets, especially those that are difficult to resolve, to stop branch swopping at an early stage of the analysis. Also Chase has argued that continual branch swopping within such datasets results in trees which are in fact too short when compared with trees produced from analysis of combined datamatrices and he questions the validity of trying to optimise nodes for which there is rather little support.

Fig. 4. Strict consensus trees for phylogenetic analysis of Thai *Syzygium.* Species belonging to phenetically defined Group 2 are indicated on the figure in bold-face type.


A. acumenatissima
C. khaoaienensis
C. operculatus
C. nigrans
C. phengklaii
E. bracteata
E. xanthocarpa
S. abortivum
S. aksornii
S. gratum
S. albiflorum
S. anacardiifolium
S. formosum
S. malaccense
S. diospyrifolium
S. jambos
S. megacarpum
S. siamense
S. scortechinii
S. aqueum
S. papillosum
S. pachyphyllum
S. samarangense
S. pycnanthum
S. lakshanakarae
S. foxworthianum
S. kurzii
S. zimmermannii
S. pseudoformosum
S. angkai
S. aromaticum
S. attenuatum
S. balsameum
S. winitii
S. muelleri
S. boiseanum
S. claviflorum
S. borneense
S. cacuminis
S. cerasiforme
S. chavaran
S. cinereum
S. polyanthum
S. craibii
S. cuminii
S. duthieanum
S. urophyllum
S. dyerianum
S. fastigiatum
S. filiforme
S. fruticosum
S. fuscescens
S. glaucum
S. globiflorum
S. grande
S. helferi
S. hemsleyanum
S. hulletianum
S. ixoroides
S. jasminifolium
S. kerrii
S. laetum
S. leptostemon
S. lineatum
S. maingayi
S. mekongense
S. nitrasaririrakii
S. oblatum
S. pachysarcum
S. pergmantaceum
S. praecox
S. praineanum
S. syzygiodes
S. puttii
S. pyrifolium
S. refertum
S. ridleyi
S. rigens
S. ripicola
S. skiophilum
S. subrufum
S. tetragonum
S. thumra
S. toddalioides
S. zeylanicum

The most parsimonious trees found were all of step length 189 with a consistency index of 0.27, a retention index of 0.63 and a homoplasy index of 0.73. Despite this relatively poor resolution, the strict consensus tree and the 50% majority rule consensus tree (latter not shown) (Fig. 4) showed some non-pectinate structure.

As can be seen from the strict consensus tree *Eugenia* is supported as a discrete clade. Of greatest interest is the support given by the strict consensus tree to the separation of species constituting the vast majority of species in Group 2. Character reconstruction showed that the species in this clade are distinguished largely by their possession of character states 2, 3 and 4 of character 21 (stamens more than 10 mm long), by usually possessing a midvein more than 1 mm wide, usually by character state 4 of character 13 (possessing a large number of gland-dots on the petals), by character state 4 of character 11 (calyx-tube broadly funnel-shaped and more than 10 mm long) and by character state 3 of character 25 (number of ovules per locule > 20). Interestingly, these are the same characters which were drawn out by DCA as being important in designation of Group 2. The inclusion of *S. pachyphyllum* in this clade does not reflect the very different shape of the calyx-tube in this species (state 2 of character 11 as opposed to states 4 or 5 for the other species), its polymorphic anther states (states 1 or 2 of character 21, state 1 being unique in this group) as opposed to the usual states 2 or 3, nor a number of other morphological differences. The other species in this analysis are not shown to form a single coherent group – the species forming, by and large, individual separate clades. Undoubtedly the huge number of trees found and the number of clades suggested by the consensus analysis must reflect the fact that the phylogenetic signal which is present is very dispersed; given the RASA values it cannot be the case that the signal is just obscured by 'noise'. The strong suggestion that Group 2 forms a unit cannot, I think, be ignored when it is taken together with the evidence from the other analysis; however, it is clear that the support for substructure in the rest of the tree is, at best, weak and no other groupings should be recognised. Logicians might argue that recognition of Group 2 has the effect that you must recognise the others at the same level – I disagree: nature is seldom logical and there is no reason to pre-suppose that the patterns it produces will infallibly be so either. To apply such arguments to cases such as *Syzygium* would appear foolish.

The composition of Group 2 suggested by phylogenetic analysis is somewhat dissimilar to that suggested by the other three analyses; *S. pachyphyllum*, *S. kurzii* and *S. zimmermannii* are all suggested for inclusion. The rationale for exclusion of the latter two species has already been discussed and I feel that the balance of evidence is not strong enough yet to warrant their inclusion or that of *S. pachyphyllum*, but does confirm the inclusion of *S. laksanakarae*, whose potential position in Group 2 was indicated by DCA. As most macro-morphological characters for which comprehensive data are available have been included in this analysis it appears that greater resolution of the interspecies relationships in *Syzygium* by phylogenetic analysis will need to involve new data sources – molecular sequence data will probably prove particularly useful.

The results of this analysis generally support those of the PCA, DSC and DCA analyses. On the basis of the strict consensus tree Group 1 contains those species delimited as belonging to that group on PCA, with the inclusion of *S. laksanakarae*.

CONCLUSIONS

All analyses presented support division of *Syzygium* s.s. in Thailand into groups. Phenetic analysis suggests that there are two groups of species present; phylogenetic analysis is supportive of the smaller of the two groups recognised phenetically as also being a phylogenetic unit. Overall the analyses show some very small discrepancy in the constitution of this smaller group (Group 2) which on balance consists of *S. anacardiifolium*, *S. aqueum*, *S. diospyrifolium*, *S. formosum*, *S. foxworthianum*, *S. jambos*, *S. lakshanakarae*, *S. malaccense*, *S. megacarpum*, *S. papillosum*, *S. pseudoformosum*, *S. pycnanthum*, *S. samarangense*, *S. scortechinii* and *S. siamense*.

The three phenetic analyses suggest that the groups are retrieved on the basis of a number of key characteristics, though there are wide ranging differences. As the groups are retrievable through a number of different analyses they therefore appear to have the attributes of polythetic taxa (Sneath & Sokal, 1973; Pauw & Linder, 1997). Unfortunately, phenetic methods do not provide ranking criteria and it is very difficult to say at what level in the taxonomic hierarchy these polythetic groups should be recognised. As Pauw & Linder (1997) point out for *Widdringtonia,* most characters show overlap between groups and are therefore 'traits' (traits are features which may assume the value observed in either of the groups being investigated). However, there are some non-overlapping features (at least in terms of mean values) which do not overlap and are not polymorphic and which can be termed 'characters'. According to Cracraft (1989) phylogenetic species are "minimal diagnosable units", i.e. units diagnosed on the basis of at least one character. Though the phylogenetic species concept is flawed (Cornet & Metz, 1993) it is still widely used. If applied here this would imply that the groups recognisable here, consisting of a number of such units which are diagnosable on the basis of a number of characters, could therefore be ranked at generic level.

Group 2 is similar, both in morphology and in species composition, to the *Jambosa* genus. This assemblage has medium to large flowers, spreading petals and large calyx-lobes (Merrill, 1950b). However, Merrill expressed ambivalence as to the distinctiveness of *Jambosa* DC. whilst admitting that it might be a real genus. The synonymy of the species included in Group 2 indicates that some, though not all, of the species in that group have been placed in *Jambosa*. The work discussed herein suggests that *Jambosa* as defined here is distinct and warrants recognition. However, for a variety of reasons, including the promotion of nomenclatural stability, I think that it is inadvisable to reinstate *Jambosa* as a genus at present, preferring on the basis of the present evidence sectional level. Therefore I recognise *Syzygium* sect. *Syzygium* and *Syzygium* sect. *Jambosa* and admit that the former section may be artificial and warrant further segregation. However, if further analyses from other geographic regions lend support to my conclusions then *Jambosa* may well have to be re-instated, albeit in modified form, as a genus.

Phenetic analysis has also shown that the two groups are different ecologically and in terms of distribution. Group 2 contains only one species endemic to Thailand (*S. lakshanakarae*) and no species occurring above 1,000 m in altitude whereas Group 1 contains many endemics and all species occurring only above 1000 m: most of the species in Group 2 are therefore lowland and fairly widely distributed. This can be

related to the tectonic history of this region of SE Asia which has recently been clarified (Metcalf, 1996). From the latter analysis it appears that Whitmore's (1981, 1998) expression of the widely held view of a dry Pleistocene period (Williams et al., 1993), allowing migration of lowland species across the Sunda Continental Shelf, is very likely correct. Certainly the data in this paper lend further support to this hypothesis as almost 50% of the species in Group 2 (the lowland group) are found on either Sumatra, Borneo and/or Java. It must be remembered that Whitmore's (1981) hypothesis will have an impact on the distribution of all species in the region – not just those at low altitude. For example, species of high altitudes are likely to have extended their altitudinal ranges during the Pleistocene by moving into lower altitudes: however, I would suggest that it is unlikely that this downward shift was ever sufficiently great to allow such species to migrate across the continental shelf. This hypothesis is supported by the data for Thai *Syzygium* where the species which occur only at high altitude have a high frequency of endemism to Indochina and do not cross the Sunda shelf. Further biometric work may also reveal support for the division of *Syzygium* into phenetic units which relate to rises in Eustatic sea-level and the presence of well-known demarcation knots (Parnell, in press).

A final conundrum is posed but remains unanswered by these data. It is clear that there are a rather large number of high altitude *Syzygium* species endemic to Thailand but very few lowland endemics. This difference could come about through a variety of mechanisms. For example the difference between the groups could reflect:

1) different evolutionary rates at different altitudes due to different niche breadths and/or different intensities of competition or (as seems likely)

2) different evolutionary patterns due to differences in breeding biology (Chantaranothai & Parnell, 1994b; Nic Lughadha & Proença, 1996) or

3) an inherent morphological uniformity in lowland species which makes differences between them more difficult to detect than in upland species or

4) a relative lack of knowledge of lowland species.

It is essential to try to get a measure of how important these different factors are, as only then a robust taxonomic framework can be generated for *Syzygium* and *Eugenia*.

ACKNOWLEDGEMENTS

I thank Prof. P. Chantaranothai who collected much of the raw data and Dr. H.-J. Esser for helpful comments on an early draft of this paper.

REFERENCES

Airy Shaw, H.K. 1949. Additions to the flora of Borneo and other Malay islands, XX. The Myrtaceae of the Oxford University Expedition to Sarawak, 1932. Kew Bull. 1949: 117–125.

Alston, A.H.G. 1931. Myrtaceae. In: Supplement to the Handbook to the Flora of Ceylon 6: 109–121. Dulau, London.

Ashton, P.S. 1981. Myrtaceae. In: M.D. Dassanayake (ed.), A Revised Handbook to the Flora of Ceylon 2: 403–472. A.A. Balkema, Rotterdam.

Bentham, G. & J.D. Hooker. 1865. Genera Plantarum. Vol. 1, 2. Reeve & Co. London.

Briggs, B.G. & L.A.S. Johnson. 1979. Evolution of the Myrtaceae – evidence from inflorescence structure. Proc. Linn. Soc. NSW. 102: 157–256.

Causton, D.G. 1988. Introduction to Vegetation Analysis. Unwin Hyman, London.
Chantaranothai, P. 1989. The Taxonomy of Eugenia L. sensu lato (Myrtaceae) in Thailand. Unpublished PhD Thesis, University of Dublin, Trinity College.
Chantaranothai, P. & J. Parnell. 1994a. A revision of Acmena, Cleistocalyx, Eugenia s.s. and Syzygium (Myrtaceae) in Thailand. Thai Forest Bull. 21: 1–123.
Chantaranothai, P. & J. Parnell. 1994b. The breeding biology of some Thai Syzygium species. Trop. Ecol. 35: 199–208.
Chase, M.W., A.V. Cox, V. Savolainen & C.M. Morton. 1997. Large DNA sequence matrices, phylogenetic signal, and feasibility: an empirical approach. In: A. Diccon & R.T. Pennington (eds.), Systematics: 57. First Biennial International Conference of the Systematics Association. Published for The Systematics Association by the Royal Botanic Gardens, Edinburgh.
Cornet, D.J. & J.A.J. Metz. 1993. Internodons as equivalence classes in genealogical networks: building-blocks for a rigorous species concept. In: D.J. Cornet (ed.), Reconstructing Species. Demarcations in genealogical Networks: 93–111. Instituut voor Theoretische Biologie, Rijksherbarium, Leiden.
Cracraft, J. 1989. Speciation and its ontology: the empirical consequences of alternative concepts for understanding patterns and processes of differentiation. In: D. Otte & J.A. Endler (eds.), Speciation and its consequences: 28–59. Sunderland, Sinaur Associates.
De Candolle, A.P.P.L. 1828. Prodromus systematis naturalis regni vegetabilis ... Vol. 3. Treuttel & Würtz, Paris.
Esser, H.-J., P. van Welzen & T. Djarwaningsih. 1998. A phylogenetic classification of the Malesian Hippomanae (Euphorbiaceae). Syst. Bot. 22: 617–628.
Frontier, S. 1976. Étude de la décroissance des valeurs propres dans une analyse en composantes principales: comparison avec le modèle du bâton brisé. J. Exp. Mar. Biol. Ecol. 25: 67–75.
Hartley, T.G. & L.M. Perry. 1973. A provisional enumeration of species of Syzygium (Myrtaceae) from Papuasia. J. Arnold Arbor. 54: 160–227.
Henderson, M.R. 1949. The genus Eugenia (Myrtaceae) in Malaya. Gdns. Bull. (Singapore) 12: 1–293.
Hill, M.O. 1973. Reciprocal averaging: an eigenvector method of ordination. J. Ecol. 61: 237–250.
Hill, M.O. 1979. DECORANA – A Fortran Program for Detrended Correspondence Analysis. New York.
Hill, M.O. & A.J.E. Smith. 1976. Principal components analysis of taxonomic data with multistate discrete characters. Taxon 25: 249–255.
Hô, P. 1992. Myrtaceae. In: P. Hô (ed.), Câyco Viêtnam. An Illustrated Flora of Vietnam 2, 1: 41–74. Montreal.
Hora, F.B. 1978. Myrtaceae. In: V.H. Heywood (ed.), Flowering Plants of the World: 161–162. Oxford University Press.
Hung-ta, C. & M. Ru-hwai. 1982. Additions to the Myrtaceous flora of China. Acta Bot. Yunnanica 4: 17–25.
Hyland, B.P.M. 1983. A revision of Syzygium and allied genera (Myrtaceae) in Australia. Aust. J. Bot., Suppl. Ser. 9: 1–164.
Johnson, L.A.S. & B.G. Briggs. 1984. Myrtales and Myrtaceae – a phylogenetic analysis. Ann. Missouri Bot. Gard. 71: 700–756.
Johnson, L.A.S., B.G. Briggs, P.G. Wilson, B.P.M. Hyland & J. Thompson. 1988. Myrtaceae. In: B.D. Morley & H.R. Toelken (eds.), Flowering Plants in Australia: 175–185. Rigby, Willoughby, Australia.
Kent, M. & P. Coker. 1992. Vegetation Description and Analysis. A Practical Approach. Boca Raton. Ann Arbor.
Kochummen, K.M. 1995. Eugenia. In: F.S.P. Ng (ed.), Tree Flora of Malaya 3: 172–247. Longman, London. Reprint of the 1978 edition.
Kostermans, A.J.G.H. 1981. Eugenia, Syzygium and Cleistocalyx (Myrtaceae) in Ceylon. Quart. J. Taiwan Mus. 34: 117–188.

Legendre, L. & P. Legendre. 1983. Numerical Ecology. Elsevier, Amsterdam.
Lemmens, R.H.M.J. 1995. Syzygium Gaertner. In: R.H.M.J. Lemmens, I. Soerianegara & W.C. Wong (eds.), Plant Resources of South-east Asia 5. Timber trees: Minor commercial timbers: 441–474. Backhuys Publishers, Leiden.
Lyons-Weiler, J.F. 1997. RASA 2.1 software and documentation for the Mac. http:loco.biology.unr.edu/archives/rasa/rasa.html updated 9/18/96.
Lyons-Weiler, J.F., G.A. Hoelzer & R.J. Tausch. 1996. Relative Apparent Synapomorphy Analysis RASA I: the statistical measurement of phylogenetic signal. Mol. Biol. Evol. 13: 749–757.
Mabberley, D.J. 1997. The Plant-Book. 2nd ed. Cambridge University Press.
Manly, B.M. 1986. Multivariate Statistical Methods. Chapman & Hall, London.
Marriott, F.H.C. 1974. The Interpretation of Multiple Observations. Academic Press, London.
McVaugh, R. 1968. The genera of American Myrtaceae – An interim report. Taxon 17: 354–418.
Merrill, E.D. 1950a. On the synonymy of Jossinia reinwardtiana (Blume) Blume. J. Arnold Arbor. 31: 329–333.
Merrill, E.D. 1950b. Readjustments in the nomenclature of Philippine Eugenia species. Philipp. J. Sci. 79: 351–430.
Merrill, E.D. & L.M. Perry. 1938. The Myrtaceae of China. J. Arnold Arbor. 19: 191–247.
Merrill, E.D. & L.M. Perry. 1939. The myrtaceous genus Syzygium Gaertner in Borneo. Mem. Amer. Acad. Arts Sci. 18: 135–202.
Metcalf, I. 1996. Pre-Cretaceous evolution of SE Asian terranes. In: R. Hall & D. Blundell (eds.), Tectonic Evolution of SE Asia: 97–122. Geological Society Special Publication 106, London.
Minchin, P.R. 1987. An evaluation of the relative robustness of techniques for ecological ordination. Vegetatio 69: 89–107.
Nic Lughadha, E. & C. Proença. 1996. A survey of the reproductive biology of the Myrtoideae (Myrtaceae). Ann. Missouri Bot. Gard. 83: 480–503.
Niedenzu, F. 1893. Myrtaceae. In: A. Engler & K. Prantl (eds.), Die natürlichen Pflanzenfamilien, Teil 3, Abteilung 7: 57–105. Engelmann, Leipzig.
Norusis, M.J. & SPSS Inc. 1993. SPSS for Windows. Professional Statistics. Release 6.0. SPSS, Illinois.
Oksanen, J. & P.R. Minchin. 1997. Instability of ordination results under changes in input data order: explanations and remedies. J. Veg. Sci. 8: 447–454.
Parnell, J. 1995. Two new taxa for the Flora of Thailand. Novon 5: 59–60.
Parnell, J. (In press). The conservation of biodiversity: aspects of Ireland's role in the study of tropical plant diversity with particular reference to the study of the flora of Thailand and Syzygium. In: B. Rushton (ed.), Biodiversity: the Irish Dimension. Royal Irish Academy, Dublin.
Parnell J. & P. Chantaranothai. 1996. A new combination in Cleistocalyx. Novon 6: 201.
Parnell, J. & M. Needham. 1998. Morphometric variation in Irish Sorbus. Watsonia 22: 153–161.
Parnell, J. & E. Nic Lughadha. 1992. Notes on Thai Myrtaceae. Kew Bull. 47: 703–706.
Parnell, J. & S. Waldren. 1996. Detrended correspondence analysis in the ordination of data for phenetics and cladistics. Taxon 45: 71–84.
Pauw, C.A. & H.P. Linder. 1997. Tropical African cedars Widdringtonia, Cupressaceae: systematics, ecology and conservation status. Bot. J. Linn. Soc. 123: 297–319.
Ridley, H.N. 1922. The Flora of the Malay Peninsula. L. Reeve & Co., London.
Schmid, R. 1972. A resolution of the Eugenia-Syzygium controversy (Myrtaceae). Amer. J. Bot. 59: 423–436.
Schmid, R. 1980. Comparative anatomy and morphology of Psiloxylon and Heteropyxis, and the subfamilial and tribal classification of Myrtaceae. Taxon 29: 559–595.
Simpson, D.A., J. Parnell, P. Chantaranothai & D. Middleton. 1995. The Royal Botanic Gardens Kew, Khon Kaen University and Trinity College Dublin expeditions to Thailand 1990 and 1993. Thai Forest Bull. 23: 50–61.
Sneath, H.A. & R.R. Sokal. 1973. Numerical Taxonomy. Freeman & Company, San Francisco.
Sokal, R.R. & F.J. Rohlf. 1995. Biometry. The Principles and Practice of Statistics in Biological Research. W.H. Freeman, New York.

Swofford, D.L. 1993. PAUP: Phylogenetic Analysis using Parsimony, Version 3.1. Washington.
Ter Braak, C.J.F. 1988. CANOCO – a Fortran Program for Canonical Community Ordination by Partial Detrended Canonical Correspondence Analysis, Principal Components Analysis and Redundancy Analysis, Version 2.1. Wageningen.
Van Welzen, P.C. & H.-J. Esser. 1997. Review: R.Ch. Kruijt [1996. A taxonomic monograph of Sapium Jacq., Anomostachys (Baill.) Hurus., Duvigneaudia J. Léonard and Sclerocroton Hochst. (Euphorbiaceae tribe Hippomaneae)]. Blumea 42: 256–257.
Whitmore, T.C. 1981. Paleoclimate and Vegetation History. In: T.C. Whitmore (ed.), Wallace's Line and Plate Tectonics: 36–42. Clarendon Press, Oxford.
Whitmore, T.C. 1998. The phytogeography of Malesian Euphorbiaceae. In: J. Dransfield, M.J.E. Coode & D.A. Simpson (eds.), Plant Diversity in Malesia III: 389–404. Royal Botanic Gardens, Kew.
Williams, M.A.J., D.L. Dunkerley, P. de Decker, A.P. Kershaw & T. Stokes. 1993. Quaternary Environments. Arnold, London.
Wilson, K.A. 1957. A taxonomic study of the genus Eugenia (Myrtaceae) in Hawaii. Pacif. Sci. 11: 161–180.