

Multivariate analysis of leaf shape patterns in Asian species of the *Uvaria* group (Annonaceae)

CONOR MEADE^{1,2*} and JOHN PARNELL¹

¹Department of Botany, Trinity College Dublin, Ireland

²Institute of Bioengineering and Agroecology, National University of Ireland, Maynooth, Co. Kildare, Ireland

Received November 2002; accepted for publication July 2003

Multivariate analysis of leaf radian measurements was used to investigate variation in leaf shape among 34 Asian species of the *Uvaria* group, a large palaeotropical group of climbing Annonaceae characterized by imbricate petals and stellate hairs. Raw data were normalized by conversion into 15 ratio characters and using the \log_{10} transformation. All species surveyed showed a unique leaf-shape 'bauplan'. The ratio character with the greatest discriminating power in both the Principal Components Analysis and Discriminant Analysis (DA) results was a measure of the shape of the leaf base. Ratio characters with the highest factor loadings for principal components 1 and 2 clearly separated the sampled taxa when plotted against one another and provided support for the retention of several taxa as distinct species or varieties. Classification of cases into taxa using DA yielded a correct classification rate of only 52% for the ratio-transformed data; however, division of taxa in the dataset into smaller subgroups defined by discrete morphological characters significantly increased the accuracy of case identification to between 67 and 100% of cases correctly classified, depending on the group. Case identification using DA on \log_{10} -transformed data was higher than for the ratio values in the entire dataset (61.7%) and the larger subgroups. However, the rate of correct case assignment was lower in the smaller groups than for the ratio data. © 2003 The Linnean Society of London, *Botanical Journal of the Linnean Society*, 2003, 143, 231–242.

ADDITIONAL KEYWORDS: classification – key – liana – Magnoliales – palaeotropical – systematics – taxonomy – variation.

INTRODUCTION

Leaf morphology has always played an important role in plant taxonomy, particularly for identifying taxa in which variation in floral structures is uninformative (e.g. *Quercus* spp.) or in which flowering specimens are infrequent owing to, for example, a limited flowering season (Stace, 1989). However, the relative paucity of informative leaf characters in many taxa means that often the most effective way to characterize taxonomic variation is to sample extensively a set of landmarks or other measures of leaf shape and so generate an accurate numerical representation of the leaf types in each sample (Dickinson, Parker & Strauss, 1987; Stace, 1989; Marcus, 1990; Marcus, Bello & García-Valdecasas, 1993; McLellan & Endler, 1998; Jensen,

Ciofani & Miramontes, 2002). Statistical analysis of this type of data has revealed that it can be a powerful key for taxonomic differentiation among species, populations and hybrids across a broad systematic and taxonomic spectrum, e.g. Aceraceae (Jensen *et al.*, 2002), Araceae (Ray, 1992), Asteraceae (Hodálová & Marhold, 1998), Fagaceae (Jensen *et al.*, 1993), Nothofagaceae (Premoli, 1996), and Rosaceae (Parnell & Needham, 1998; Rumpunen & Bartish, 2002).

Although the analysis of landmark data has become a popular choice for leaf morphology studies (see Jensen *et al.*, 2002, for a recent review), analysis of leaf measurement data using multivariate techniques such as Principal Components Analysis (PCA; Sneath & Sokal, 1973) is also employed successfully for the investigation of taxonomic relationships (e.g. Premoli, 1996; Hodálová & Marhold, 1998; Parnell & Needham, 1998; Clausen, Meyer & Renner, 2000). Canonical Discriminant Analysis (DA) is another powerful mul-

*Corresponding author. E-mail: conor.v.meade@may.ie

tivariate method, which uses patterns of variation in multidimensional space as a basis for defining and distinguishing among taxonomic groups, providing measures of group coherence and probability of group membership for individual cases. Raw data are unsuitable for these parametric analyses when it is not normally distributed, and a transformation such as $\ln x$ or $\log_{10} x$ is generally used to give the data a normal spread – although any transformation that normalizes the data is in theory acceptable (Sokal & Rohlf, 1981).

One such transformation is the conversion of raw data into ratios. Some authors have argued that ratios in themselves are not reliable summaries of variation because they can compound recording error (Sokal & Rohlf, 1981) and do not completely factor out size (Bookstein *et al.*, 1985). However, ratios can conveniently summarize data where, for example, the relationships between a set of characters and one other character are of interest over and above the relationship among all the characters simultaneously, such as is the case when analysing shape patterns that might be of use in taxonomic keys. Furthermore, when the analysis is complete, ratio characters with a high correlation for the major summary axes in PCA can be used to plot cases in a scatter diagram, thereby separating cases along axes that are not too abstract from reality, as can be the case with summary axes and, to an extent, logarithmic transformations.

There have been some investigations of leaf shape in the Annonaceae aside from the general descriptions of leaf morphology presented in taxonomic accounts. Jovet-Ast (1942) and van Setten & Koek-Noorman (1986) studied the general anatomy and venation patterns of annonaceous leaf types, and although taxon identification using leaf morphology is possible based on the findings of this work, both studies were largely descriptive and did not include any comparative or statistical analysis of leaf shape. More recent systematic work by van Zulen, Koek-Noorman & Maas (1995) and Chatrou (1997) have included a more rigorous analysis of leaf morphology – the first covering a phylogenetic investigation of 60 species of the neotropical genus *Duguetia*, the latter focusing on four species of *Malmea* in Central America. Although both of these latter studies used clustering algorithms to differentiate among samples, the categorical scoring used for all characters in the *Duguetia* analysis was unsuccessful and yielded a poorly resolved phylogeny as a result (with a consistency index score of 0.112). Chatrou's (1997) study of *Malmea* included nine leaf shape characters from a total sample of 53 morphological characters, and these were recorded either as binary scores (six characters), raw length measures (two characters) or as a ratio (one character). Leaf shape characters occupied a central position in the

reclassification of the taxonomic group that emerged from the study: a new subspecies was identified largely using leaf shape characters; and overall four of the ten most important clustering characters were related to leaf shape, with length-to-width ratio and the shape of the leaf base being the two most important.

The contrasting results of these two studies are echoed in results presented by Meade *et al.* (2002), which suggest that leaf characters display a high degree of homoplasy and must be treated with care in any phylogenetic analysis (as is the case with a large proportion of floral and fruit characters in the Annonaceae: van Heusden, 1992; Kessler, 1993; Doyle & Le Thomas, 1996). At the same time, however, Chatrou (1997) demonstrates that the morphological variation present in leaf characters is extremely useful for elucidating taxonomic relationships, particularly among closely related taxa.

This current work stems from a revision of the *Uvaria* group in continental Asia (Meade, 2000). *Uvaria* L. is the second largest palaeotropical genus in the Annonaceae and includes approximately 150 species (Kessler, 1993; Meade, 2000). All *Uvaria* species are climbers and are typically found in lowland perhumid tropical forest, although certain taxa are common in riverine and montane forest habitats also. The climbing habit in *Uvaria* is facilitated by backward orientated twining branchlets (Fig. 1A), and the genus can be readily distinguished from most other Annonaceae by the combined presence of stellate hairs, many-seeded apocarpous fruits and an imbricate perianth. Several genera share these characters and together these form the *Uvaria* group: *Cyathostemma*, *Rauwenhoffia*, *Ellipeiopsis*, *Ellipeia* and *Anomianthus* (all in continental Asia), *Tetrapetalum* (in Malesia), and *Balonga* and *Afroguatteria* in Africa (van Heusden, 1992; van Setten & Koek-Noorman, 1992; Kessler, 1993; Meade, 2000). *Ellipeiopsis* is unique in the group in having an exclusively shrubby habit.

Presently there is a reliable array of discrete character differences that can be used to differentiate between taxa in the *Uvaria* group; however, a common problem (as with Annonaceae material in general) is that a large number of duplicate sterile sheets are created for each flowering or fruiting specimen, leading to a situation in which many herbaria have sterile duplicates received on exchange. In the absence of adequate labelling or an updated collectors' database, most of these specimens cannot be identified with any degree of certainty by the non-specialist, short of assigning them to *Uvaria* if they have stellate ferruginous hairs.

It was evident from early phases of this revision work that leaf shape variation was consistent enough to allow taxonomic identification of many species based on leaf shape alone: for example, *Uvaria narum*

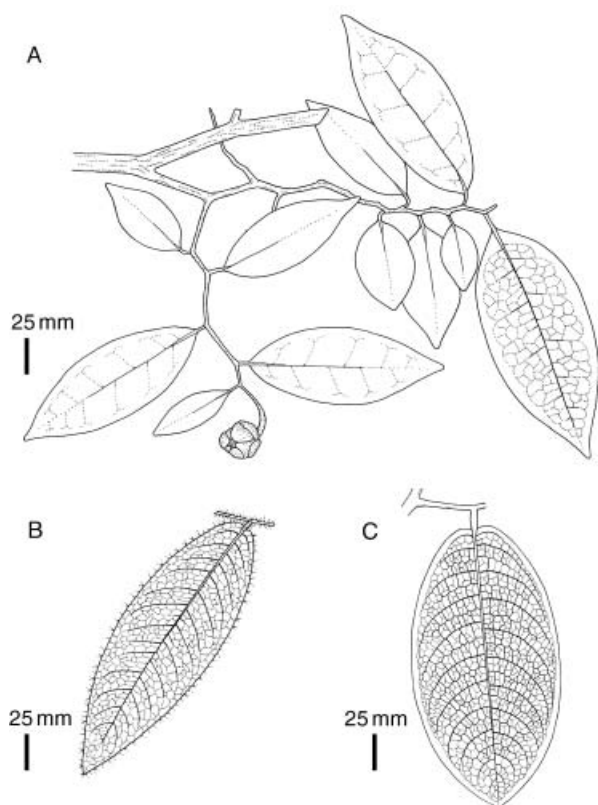


Figure 1. Selection of species illustrating the variation in shape in the *Uvaria* group. A. *Uvaria narum* (Dunal) Bl. Flowering shoot with leaf habit showing acute base and acuminate apex. B. *Uvaria* sp. 2. Lanceolate leaf habit with narrowly cordate base and acute apex. C. *Uvaria pauciovulata* Hook. F. & Thoms. Ovate-acute leaf habit showing broadly cordate base and shortly acute apex.

(Dunal) Bl. (Fig. 1A), *Uvaria* sp. 2 (Fig. 1B), and *Uvaria pauciovulata* Hook. F. & Thoms (Fig. 1C). However, much of this differentiation was based on loose descriptive terms rather than any statistically robust methodology, and the criteria used for identifying differences becomes largely redundant among closely related taxa. The goal of the work presented here was to develop a method for taxon identification using leaf shape, and three specific goals were set for the analysis: (i) identification of the most variable leaf shape characters; (ii) developing the most effective method for identifying sterile specimens using leaf shape data; and (iii) identification of leaf shape 'bauplan' for each taxon.

MATERIAL AND METHODS

SAMPLING AND DATA COLLECTION

Leaf shape measurements were made using an overlay protractor with 17 radians spanning 180° around a

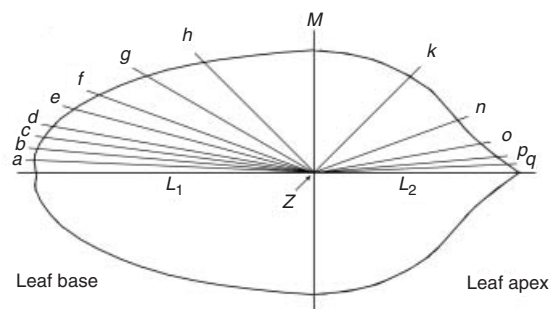


Figure 2. Leaf overlay protractor used to generate leaf shape data. Sixteen radians are indicated, with one further radian measurement (L) calculated by adding L_1 and L_2 . The leaf centroid (Z) is also indicated.

central point, a method used by Dickinson *et al.* (1987) and Premoli (1996). The central point corresponds to the leaf centroid at the intersection of the midrib with the diagonal across the widest point of the leaf (Fig. 2). The radians were designed to intersect the leaf margin at a frequency correlated with the extent of curvature at that point – therefore the leaf base and leaf apex are intersected by radians at closer intervals than the medial portion of the leaf margin. As the leaves of *Uvaria* group taxa are symmetrical, measurements were made on one half of the upper blade only.

A total of 349 sheets (one leaf per sheet) were measured for the analyses, representing five genera and 34 taxa in the *Uvaria* group, and an average of 10.15 cases per taxon (Table 1). For the lesser collected species all available material was sampled, which in some cases amounted to the type specimen only. For certain other taxa in the sample a much higher than average number of cases were studied, and this is an artefact of additional sampling carried out for a related study on morphological variation in floral and fruit structure in these species. All sample measurements were taken from flowering or fruiting sheets where taxon identification was unambiguous, and leaves were measured from either flowering or non-flowering branches.

DATA ANALYSIS

Data-normalization for multivariate analysis was achieved by conversion of measurements into ratios (Table 2) and using the \log_{10} transformation. For the ratio transformations, the radians for the basal and apical portions of the leaf were divided by the length of the basal (L_1) and apical (L_2) midrib radians, respectively, thereby providing an estimate of leaf margin curve at these points. Values greater than 1 indicate a cordate/obcordate curve, whereas values less than 1 indicate a cuneate/acuminate curve. Overall leaf length along the midrib (L) was divided by the basal

Table 1. Number of samples measured per species

Taxon	Sample size
<i>Anomianthus dulcis</i> (Dunal) Sinclair	13
<i>Cyathostemma argenteum</i> (Bl.) Sincl.	4
<i>C. micranthum</i> (Hook. f. & Thoms.) Sinclair	13
<i>C. urayii</i> King	1
<i>Ellipeiopsis cherrevensis</i> (Pierre ex Finet & Gagnep.) R. E. Fr.	6
<i>E. ferruginea</i> (Buch.-Ham. ex Hook. f. & Thoms.) R. E. Fr.	5
<i>Rauwenhoffia siamensis</i> Craib	9
<i>Uvaria boniana</i> Finet & Gagnep.	7
<i>U. calamistrata</i> Hance	6
<i>U. cordata</i> (Dunal) Alston	21
<i>U. curtisii</i> King	13
<i>U. dac</i> Pierre ex. Finet & Gagnep.	16
<i>U. flexuosa</i> Jovet-Ast	1
<i>U. grandiflora</i> Roxb. ex Hornem.	13
<i>U. hahnii</i> Finet & Gagnep.	10
<i>U. hamiltonii</i> Hook. f. & Thoms.	13
<i>U. hirsuta</i> Jack.	14
<i>U. lamponga</i> Scheff.	4
<i>U. lobbiana</i> Hook. f. & Thoms.	14
<i>U. lurida</i> Hook. f. & Thoms.	25
<i>U. maclurei</i> Diels	5
<i>U. macropoda</i> Hook. f. & Thoms.	16
<i>U. microcarpa</i> Champ. ex. Benth	18
<i>U. narum</i> (Dunal) Bl.	19
<i>U. pauciovulata</i> Hook. f. & Thoms.	12
<i>U. pierrei</i> Finet. & Gagnep.	11
<i>U. rufa</i> Bl.	18
<i>U. semi-carpifolia</i> Hook. f. & Thoms.	10
<i>Uvaria</i> sp. 1	1
<i>Uvaria</i> sp. 2	5
<i>U. sphenocarpa</i> Hook. f. & Thoms.	4
<i>U. timoriensis</i> Bl.	3
<i>U. tonkinensis</i> Finet & Gagnep.	3
<i>U. zeylanica</i> L.	16
Total	349
Mean sample size	10.3
Median sample size	10.5

midrib radian (L_1) to provide an estimate of the overall ovate/obovate leaf shape. L was also divided by the leaf width (M) to determine the relative breadth of the overall leaf shape.

PCA and DA were used to summarize the patterns evident in the dataset using SPSS 8.0 (SPSS Inc., 1997). For PCA, cases with missing values were excluded casewise, and components with a Kaiser–Guttman score of 1 were regarded as significant. For DA, prior probabilities for group membership were computed from group sizes and characters were added simultaneously.

Table 2. Leaf character measurements and ratio transformations used in multivariate analysis. a – q and L_1 and L_2 represent leaf radians

Abbreviation	Measurement	Ratio transformation
L	Overall length of leaf	L/L_1
L_1	Length from base of midrib to widest point	
L_2	Length from widest point to midrib apex	
M	Width of leaf at widest point	L/M
a	Length of radian a	a/L_1
b	Length of radian b	b/L_1
c	Length of radian c	c/L_1
d	Length of radian d	d/L_1
e	Length of radian e	e/L_1
f	Length of radian f	f/L_1
g	Length of radian g	g/L_1
h	Length of radian h	h/L_1
k	Length of radian k	k/L_2
n	Length of radian n	n/L_2
o	Length of radian o	o/L_2
p	Length of radian p	p/L_2
q	Length of radian q	q/L_2

ANALYSIS OF LEAF-SHAPE PATTERNS IN THE ENTIRE DATASET

Sampling and analysis of the entire dataset was approached in two ways for PCA. In the first approach, all cases in the complete dataset were entered with no a priori grouping, meaning that each taxon had between one and 25 entries, depending on sampling frequency. This was designed to investigate patterns based on the total variation present in all cases in the dataset. In the second approach mean ratio values were calculated a priori for each taxon so that only 34 values (i.e. for 34 taxa) were used. This was designed to analyse differences between taxa based on the leaf shape bauplan for each taxon.

For DA, all samples were entered for each taxon with a priori grouping so that each had between one and 25 entries, depending on sampling frequency, and this facilitated construction of a multivariate model for each taxon during the analysis.

ANALYSIS OF SUBGROUPS DEFINED BY DISCRETE MORPHOLOGICAL CHARACTERS

Two methods were tried to increase the rate of successful taxon classification using DA on the entire dataset. The first involved removing the smaller sam-

Table 3. Taxonomic sub-groups united by possession of discrete morphological characters

Group	Semi-glabrous	Hairs mixed	Cuneate base	Glabrous	Veins > 20	Margins hirsute	Scabrid
Taxa	<i>U. argentea</i>	<i>U. curtisii</i>	<i>U. boniana</i>	<i>U. boniana</i>	<i>U. cordata</i>	<i>U. curtisii</i>	<i>U. rufa</i>
	<i>U. cordata</i>	<i>U. dac</i>	<i>U. calamistrata</i>	<i>U. lurida</i>	<i>U. hamiltonii</i>	<i>U. hirsuta</i>	<i>U. timoriensis</i>
	<i>U. calamistrata</i>	<i>U. grandiflora</i>	<i>U. flexuosa</i>	<i>U. maclurei</i>	<i>U. lamponga</i>	<i>U. sp. 2</i>	<i>U. sp. 2</i>
	<i>U. flexuosa</i>	<i>U. hamiltonii</i>	<i>U. lurida</i>	<i>U. macropoda</i>	<i>U. semecarpifolia</i>		
	<i>U. lamponga</i>	<i>U. hirsuta</i>	<i>U. macropoda</i>	<i>U. tonkinensis</i>			
	<i>U. lobbiana</i>	<i>U. sphenocarpa</i>	<i>U. micrantha</i>				
	<i>U. micrantha</i>	<i>U. sp. 2</i>	<i>U. narum</i>				
	<i>U. microcarpa</i>	<i>A. dulcis</i>	<i>U. zeylanica</i>				
	<i>U. narum</i>	<i>E. cherrevensis</i>					
	<i>U. pierrei</i>	<i>E. ferruginea</i>					
	<i>U. semecarpifolia</i>						
	<i>U. wrayii</i>						
	<i>U. zeylanica</i>						
	<i>U. sp. 1</i>						
	<i>E. pauciovulata</i>						
	<i>R. siamensis</i>						

Taxa are grouped according to discrete leaf characteristics. Semi-glabrous: upper blade typically glabrous with some hairs over midrib; Hairs mixed: upper blade typically having simple and stellate hairs; Cuneate base: leaf base typically cuneate; Glabrous: no hairs present on upper blade or midrib; Veins > 20, some leaves with more than 20 primary veins; Margins hirsute: leaves with prominent stellate hairs on margin; Scabrid: hairs on upper blade scabrid to the touch. See text for discussion.

ples, including the unique samples, from the group so as to increase the average sample size for each taxon in the analysis; however, this reduced the success rates for case classification. The second method involved assigning taxa into subgroups based on discrete leaf morphology characters and performing DA on these smaller groups. This latter technique proved much more effective and was selected as the optimal method for subdividing the sample.

In total seven groups of taxa were identified from the entire sample group using the following discrete characters (Table 3): upper leaf blade semi-glabrous (hairs persistent on midrib, deciduous elsewhere) – ‘semi-glabrous’; upper blade containing a mix of simple and stellate hairs – ‘hairs mixed’; leaf base cuneate – ‘cuneate base’; upper surface entirely glabrous – ‘glabrous’; leaf veins greater than 20 in number – ‘veins > 20’; leaf margin hirsute – ‘margins hirsute’; and hairs on upper surface scabrid to touch – ‘scabrid’.

RESULTS

MULTIVARIATE ANALYSIS OF LEAF SHAPE VARIABILITY

Ratio-transformed data

Table 4 describes the results for PCA using sampling approach 1 (all cases in the dataset entered with no a priori grouping) and sampling approach 2 (mean taxon ratio values derived from a priori single-taxon groups), as well as the DA results (all cases in the dataset entered with a priori grouping).

In both the PCA analyses and the DA analysis of ratio values, leaf character c/L_1 emerges as highly correlated with the major summary axis. As a result, taxa with cordate leaf bases tend to score highly on Component 1 and Function 1 reflecting a c/L_1 score close to 1.0, whereas those with cuneate or narrowly retuse bases and c/L_1 scores much less than 1.0 tend to have low scores.

To illustrate the patterns identified in the multivariate analyses, two figures are included that plot taxon scores for the ratio measurements most strongly correlated with the PCA summary axes. The first, Figure 3, shows a scatter plot generated using mean taxon scores for c/L_1 and L/L_1 , the characters most closely associated with principal components 1 and 2, respectively, in Table 3. As c/L_1 is a summary of leaf base shape, Figure 4 has been included showing a scatter plot generated using mean taxon scores for p/L_2 , a measure of leaf apex shape, and L/L_1 . Both of these graphs also contain mean leaf outlines for species with extreme values on the character axes.

Log₁₀-transformed data

Component and Factor scores for PCA and DA of the \log_{10} -transformed data are given in Table 4. PCA yielded only one significant component for both the total and the mean taxon value datasets. e and g emerge as the characters most strongly associated with the principal component in the analysis of the entire dataset, whereas e and o are most strongly correlated with component 1 in the analysis of mean

Table 4. Multivariate analysis of variation in leaf shape in the *Uvaria* group – component/factor scores from Principal Components Analysis and Discriminant Analysis of total leaf radian dataset. See text for discussion

Sample	Cases		% cases correctly assigned (DA)	Component (PCA)* Factor (DA)	Eigenvalue	% of Variance	Correlated characters	Correlation
Principal Components Analysis								
Total dataset	349	Ratio values		1	8.83	55.18	<i>e/L</i> ₁ , <i>d/L</i> ₁	0.91
				2	3.45	21.57	<i>c/L</i> ₁ <i>L/L</i> ₁ <i>p/L</i> ₂	0.90 0.84 –0.67
	log ₁₀ values		1	15.64	92.01	<i>e</i> <i>L</i> ₁	0.98 0.97	
			1	10.09	67.25	<i>c/L</i> ₁ <i>d/L</i> ₁	0.96 0.94	
Mean taxon values	34	Ratio values		2	2.77	18.44	<i>L/L</i> ₁ <i>p/L</i> ₂	0.85 –0.58
				1	16.05	94.44	<i>e</i> <i>o</i>	0.99 0.98
	log ₁₀ values		1	3.20	35.60	<i>c/L</i> ₁ , <i>b/L</i> ₁ <i>k/L</i> ₂ , <i>n/L</i> ₁	0.62 0.59	
			2	1.78	19.90	<i>g/L</i> ₁ <i>f/L</i> ₁	0.47 0.38	
Discriminant Analysis								
Total dataset	349	Ratio values	52.0	1	3.39	38.90	<i>M</i> <i>h</i>	0.78 0.74
				2	1.60	18.30	<i>L</i> ₂ <i>k, M</i>	0.32 0.28
	log ₁₀ values	61.7	1	3.39	38.90	<i>M</i> <i>h</i>	0.78 0.74	
			2	1.60	18.30	<i>L</i> ₂ <i>k, M</i>	0.32 0.28	

*Significant components based on Kaiser–Gutman criteria.

taxon values. *M* and *L*₂ are most strongly correlated with Discriminant Functions 1 and 2, respectively, in the analysis of the entire dataset.

ANALYSIS OF TAXONOMIC SUBGROUPS

Ratio-transformed data

The highest classification rates using DA for the ratio-transformed data were achieved for the smaller groups, with the scabrid group (three taxa, 23 cases) and the veins > 20 group (four taxa, 49 cases), both facilitating a 100% correct classification rate for cases in the analysis (Table 5, Fig. 5). The largest group, semi-glabrous (16 taxa, 161 cases), had a success rate of 77%, significantly higher than the score for the entire dataset (52%, Table 3). The lowest success rate, 67.9%, was generated for the cuneate base group.

Log₁₀-transformed data

Table 5 includes a comparison of the classification rates for the log₁₀- and ratio-transformed data. Higher levels of correct classification are seen in the DA analysis of log₁₀ values for the two larger groups: semi-gla-

brous yielded 83.9% of cases correctly classified and hairs mixed yielded 72.6% (vs. 77 and 70.8%, respectively, for the ratio-transformed data). The lowest classification rate is seen for the glabrous group, with only 55.9% of cases correctly assigned.

LEAF SHAPE BAUPLAN FOR EACH TAXON

Appendix 1 shows mean leaf shape outlines (bauplan) for each taxon included in the analysis. The radian scores for each bauplan are calculated as the mean untransformed distance values for that radian in each taxon sample. Taxa are arranged according to their assigned leaf type subgroups.

DISCUSSION

LEAF-SHAPE PATTERNS IN THE *UVARIA* GROUP

In both the PCA and the DA analysis of ratio leaf characters for the total dataset, character *c/L*₁ emerges as highly correlated with the major summary axis (Table 4). The first six leaf ratios (*a–f/L*₁, including *c/L*₁) were designed to summarize the shape of the leaf

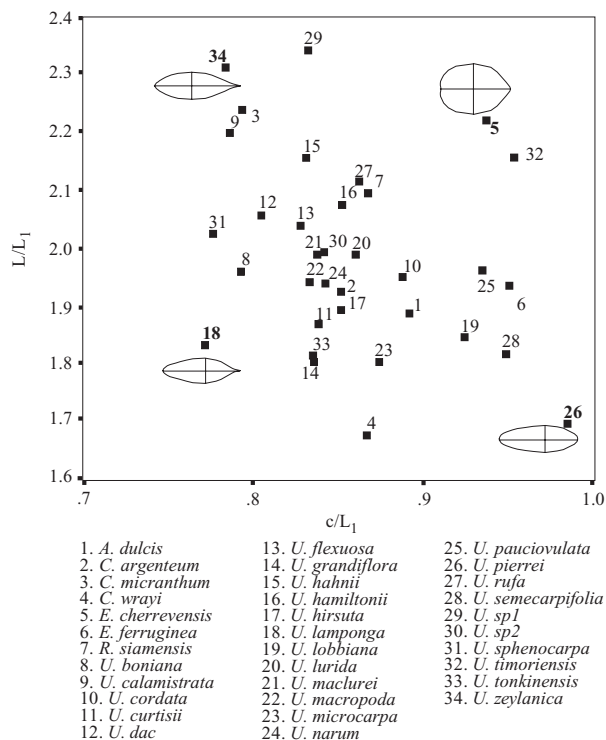


Figure 3. Taxon scores for the characters most closely associated with components 1 (c/L_1) and 2 (L/L_1) in Principal Components Analysis of mean leaf ratio values. L/L_1 represents the position of the widest point of the leaf along the midrib, distinguishing ovate leaves (towards the top of the plot) from obovate leaves (towards the bottom). c/L_1 represents the extent to which the leaf base is cuneate (to the left), retuse (centre), or cordate (to the right). Leaf outlines represent the mean shape (independent of size) for the adjacent taxon in bold type. Identification numbers for taxa are shown below.

base in each sample, and these results indicate that c/L_1 , and to a lesser extent b/L_1 , d/L_1 , and e/L_1 , contain the greatest variation of values and carry the greatest discriminating power of these leaf-base values.

PCA of the complete and mean value datasets identify L/L_1 as the character most strongly correlated with function 2 (Table 4), whereas DA identifies g/L_1 . The inference here is that whereas g/L_1 accounts for a significant amount of variation uncorrelated with c/L_1 , it also has a significant discriminating power. L/L_1 , by contrast, accounts for a larger proportion of the variation uncorrelated with c/L_1 , but much of that variation occurs within individual taxa and so it has a weaker discriminating power.

Plotting c/L_1 and L/L_1 values provides a clear separation of taxa according to leaf base and ovate/obovate shape, respectively (Fig. 3). However, leaf apex scores such as p/L_2 , which have a lower correlation score with the principal summary axes, also separate

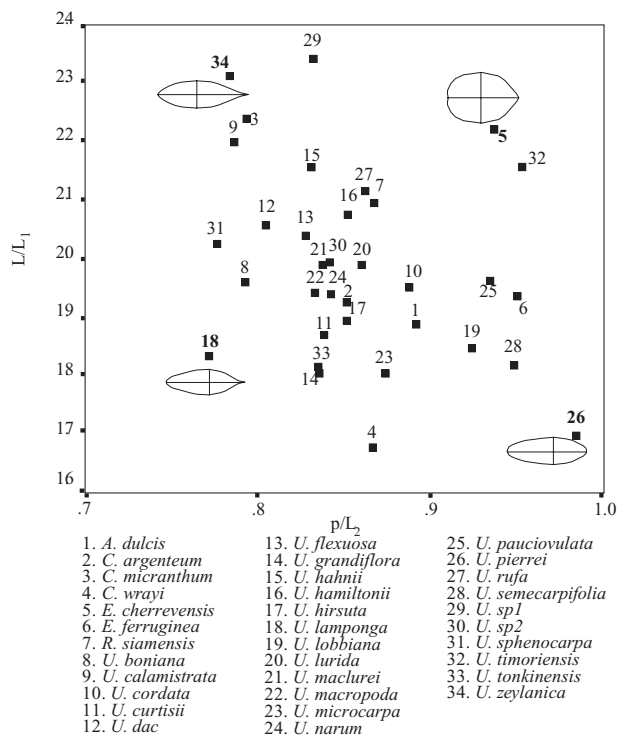


Figure 4. Taxon scores for characters L/L_1 and p/L_2 . L/L_1 represents the position of the widest point of the leaf along the midrib, distinguishing ovate leaves (towards the top of the plot) from obovate leaves (towards the bottom). p/L_2 represents the extent to which the leaf apex is acuminate (to the left), acute (centre), or obtuse (to the right). Leaf outlines represent the mean shape (independent of size) for the adjacent taxon in bold type. Identification numbers for taxa are shown below.

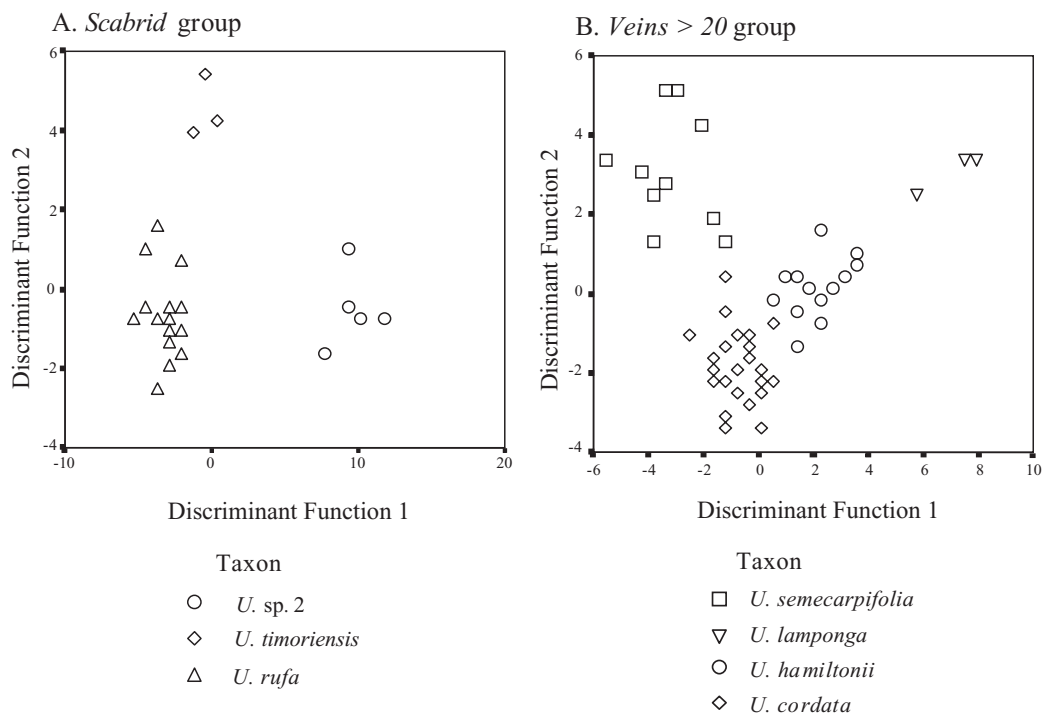
taxa when plotted against L/L_1 (Fig. 4). In summary, sampling of the entire dataset provides a useful guide to character variation within the sample taxon group, but it is less valuable as a means for discriminating between taxa.

ANALYSIS OF TAXONOMIC SUBGROUPS

The division of taxa into seven subgroups greatly increased the accuracy of DA classification against that seen for analysis of the entire dataset (Tables 4, 5) for both ratio and \log_{10} -transformed data. This trend might be expected given the higher character to case ratio present in smaller groups. However, a significant additional factor is the taxonomic composition of the groups. For example, the margins hirsute group contains three taxa and 33 cases and scores 84.8 and 93.9% correct classification for \log_{10} and ratio values, respectively, whereas the veins > 20 group, with four taxa and 49 cases, scores 100% for both data types. Equally, the largest group, semi-glabrous (16 taxa, 161

Table 5. Multivariate analysis of variation in leaf shape in the *Uvaria* group – Discriminant Analysis of taxonomic subgroups

Taxonomic group	No of taxa	Cases	% cases correctly classified		Factor	Eigenvalue	% of variance	Correlated characters	Correlation
			Ratio values	(log ₁₀ values)					
Semi-glabrous	16	161	77.0	(83.9)	1	5.69	45.10	<i>a/L</i> ₁	0.64
					2	3.20	25.30	<i>g/L</i> ₁	-0.63
Hairs mixed	11	113	70.8	(72.6)	1	3.55	51.90	<i>h/L</i> ₁	0.70
					2	0.88	12.80	<i>a/L</i> ₁	0.63
Cuneate base	8	106	67.9	(62.3)	1	1.58	40.00	<i>g/L</i> ₁	0.48
					2	1.11	28.10	<i>o/L</i> ₂	0.30
Glabrous	5	60	81.7	(55.9)	1	1.40	56.00	<i>p/L</i> ₂	0.33
					2	0.73	29.00	<i>h/L</i> ₁	0.49
Veins > 20	4	49	100.0	(100)	1	7.24	52.30	<i>e/L</i> ₁	-0.45
					2	4.48	32.40	<i>a/L</i> ₁	0.44
Margins hirsute	3	33	93.9	(84.8)	1	3.75	74.20	<i>b/L</i> ₁	0.37
					2	1.31	25.80	<i>a/L</i> ₁	0.50
Scabrid	3	23	100.0	(91.3)	1	8.83	83.90	<i>b/L</i> ₁	0.21
					2	1.70	16.10	<i>b/L</i> ₁	0.39

**Figure 5.** Discriminant analysis of taxonomic subgroups. Scatter plots with cases using Discriminant Function 1 and 2 scores. A. Scabrid. B. Veins > 20 group.

cases, 83.9 and 77% correct classification, respectively, for log₁₀ and ratio data), scores higher than the second largest group, hairs mixed (11 taxa, 113 cases, 72.6 and 70.8% correct classification).

Although taxon-specific clustering of cases on the DA scatterplots increases with higher correct classification rates, it is only when the correct classification level approaches 100% that unambiguous classifica-

tion of all cases becomes possible (Fig. 5). Below this level only a subset of taxa within each group are clearly defined when plotted together – limiting the extent to which new sample cases could be unambiguously assigned to one taxon or another.

Appendix 1 presents the leaf bauplan outlines for each taxon according to the taxonomic subgroup they are assigned to. Some of the groups include taxa with very different leaf shapes, e.g. veins > 20 and scabrid, and the DA classification scores obviously benefit from these clear differences. In other groups such as the cuneate base group the differences between taxa are less clear, and this similarity reflects the generally closer phylogenetic and morphological relationship among the taxa in this group (Meade, 2000). For example, *Uvaria narum*, *U. lurida* and *U. macropoda* show enough morphological intergradation to suggest that hybridization may take place along contact zones between the species in India and Sri Lanka, and they may even represent localized elements within a single species unit.

Other analyses of the Annonaceae indicate that morphological characters quickly become homoplastic (i.e. phylogenetically discordant) if sampled over too large a range of taxa (Doyle & Le Thomas, 1994, 1996; van Zuilén *et al.*, 1995). However, the present study has shown that this pattern can be exploited for taxonomic purposes where several taxa from different phylogenetic lineages are placed into a single taxonomic group based on a superficially similar character – such as vein number, hair density or texture. This groups together taxa that in other respects are quite dissimilar. As such, it has proven convenient in this investigation to define part of the leaf-shape key on the basis of characters that are similar by convergence rather than by common ancestry, thereby grouping together taxa that have quite different leaf shapes (e.g. Fig. 5B, Appendix 1 – veins > 20).

Whereas the overall probability of confidently classifying a specimen is largely dependent on whether it belongs to one of the smaller groups, or whether it has an unusual shape overall, using the combined approach of the subgroup key and the bauplan images maximizes the use of available leaf shape information and provides a new key for leaf types in the *Uvaria* group.

CLASSIFICATION SUCCESS OF RATIO TRANSFORMATIONS VS. LOG₁₀ TRANSFORMATIONS

As mentioned in the Introduction, the argument for using ratio transformations in this analysis is rooted in the desire to create taxonomically meaningful distinctions among taxa that can be easily visualized in a two-dimensional plot. Clearly it is also important that these transformations do not obscure the variation present in the untransformed data. In order to assess

the effect of ratio transformations in relation to a standard transformation method, DA was carried out on the log₁₀ values of all radian measurements for the taxa in the main dataset and in each subgroup (Tables 3, 4). As shown in Table 5, the trend in successful case identification using DA was quite similar in both transformed datasets, with successful case assignment increasing with smaller group sizes. The log₁₀ data generated the highest rate of correct classification in the larger groups, and the gap between the two transformation methods increases with larger group size. However, in the smaller groups the situation is reversed, with the ratio data facilitating higher rates of correct classification. One possible explanation for this trend lies in the normalizing effect of the two transformations.

In general, the log₁₀ transformation produced a more normalized distribution of data in all characters than did the ratio conversions (data not shown), and this resulted in a more even spread in the data, particularly in the larger datasets. Conversely, in the ratio-transformed data, in which taxa show a higher number of outliers, cases are less prone to being misclassified into adjacent groups when the dataset (and number of alternative groups) is much reduced in size.

An additional factor may be that the ratio transformations emphasize the differences among taxa in a very particular way. Where taxa show differences in terms of leaf base and leaf apex shape but not in the medial radians of the leaf (such as in the glabrous group, Appendix 1), analysis of the ratio-transformed data will emphasize this difference because of the relationship among all the radians and either L_1 or L_2 (Table 1, Fig. 2). However, where taxa show differences in the medial portions of the leaf they will be better separated in analysis of the log₁₀-transformed data because this variation is not obscured by any correlation between the radian measurements and L_1 or L_2 (e.g. the hairs mixed group).

TAXONOMIC FINDINGS

These results provide new data for resolving taxonomic relationships within the *Uvaria* group, particularly between taxa that have proven difficult to separate using phylogenetic techniques. The leaf types of *Ellipeiopsis ferruginea* and *E. cherrevensis* (Pierre ex Finet & Gagnep.) R. E. Fries are clearly separated in the plots produced in the analysis (Figs 3, 4). Taxonomic analysis of floral and fruit structures by Meade (2000) suggested that these may be the same taxon; however, the results presented here establish that the leaf type of *E. cherrevensis* is distinctive. Combined with the separate distributions of the two taxa, there seems sufficient support for the retention of *E. cherrevensis* as a variety of

E. ferruginea. A second significant result is the difference between the leaves of *Uvaria cordata* and *U. microcarpa* Champ. ex Benth. (Figs 3, 4). Whereas these two taxa can be easily distinguished from one another based on fruit morphology, the flowers of both species are quite similar and difficult to distinguish. The differences in leaf shape identified here, combined with the discrete difference in vein number (*U. cordata* is assigned to the veins > 20 group) should provide a more secure basis for distinction between the two species.

Another useful distinction is between *U. hamiltonii* Hook. f. & Thoms. and a new undescribed species from Indochina, *Uvaria* sp. 2. These taxa are clearly separated on the basis of fruit structure, although the leaves of the two initially appear quite similar. However, they are clearly separated from one another in Figures 3 and 4, and also differ through a discrete difference in vein number (*U. hamiltonii* is assigned to the veins > 20 group). Another useful finding is the difference between *U. rufa* Bl. and *U. timoriensis* Bl. Meade (2000) highlighted that differentiating between the type material of these two species is difficult, particularly given the many similarities in floral structure; however, the results presented here clearly separate the two on the basis of leaf shape and support their retention as separate species (Fig. 5A).

ECOLOGICAL CORRELATIONS

Although variation in leaf shape has proven useful for discriminating between taxa in the *Uvaria* group, ecological patterns from the same data are much less clear. Smaller leaf size and modification of the leaf cuticle and indument are correlated with adaptation for reduction of water stress in plants (Givnish, 1979; Sandquist & Ehleringer, 1997), and all of these strategies are utilized by *Uvaria* group taxa. In general, large leaves or (semi-) glabrous leaves (such as those seen in *U. semecarpifolia* and *U. lurida*, respectively) are most common in wet-habitat taxa, whereas densely hairy or, to a lesser extent, small leaves (such as those seen in *Ellipeiopsis ferruginea* and *U. zeylanica*, respectively) are common in taxa that tend to occupy more open, dry-soil habitats. Ecologically adaptable taxa such as *U. cordata* show large variation in leaf size and indument, with specimens from drier areas having smaller, more densely hairy leaves, whereas those from well-watered sites have much larger leaves with more sparse hairs.

However, aside from indument and leaf size, the patterns in leaf shape identified in this study show little correlation with the habitat preferences of particular taxa. Most taxa with cuneate glabrous leaves are restricted to well-watered soils, but *U. zeylanica* is

typically found in drier sites such as sand dunes. Taxa with cordate hairy leaves are also common in wet forest habitats, although many do better in drier habitats, e.g. *Ellipeiopsis ferruginea*, *E. cherrevensis*, *Anomianthis dulcis* and *Uvaria rufa*.

Although ecological factors have probably played a role in the evolution of different leaf types in the *Uvaria* group, the nature of this relationship is unclear from the data presented in this study.

CONCLUSION

Although most taxa could be clearly separated using multivariate analysis of data generated using the leaf radians technique, certain taxa were shown to have very similar leaf shape types. As a result, discriminant analysis of all the taxa together yielded a relatively low rate of correct classification, with unambiguous assignment of cases to taxa within these clusters of similar leaf shape types proving quite difficult. As shown in Table 3, only 52% of cases were assigned to their correct taxon when the entire ratio-transformed dataset was analysed simultaneously. However, the division of taxa into subgroups defined by discrete leaf morphology characters greatly improved the rate of correct classifications using DA, with some groups showing a correct classification rate of 100%. However, the overall benefit of using ratios for the purposes of visualizing taxonomic differences is to an extent offset by the sampling bias of the ratios themselves – variation in the medial portion of the leaf outline was much better analysed using the log₁₀-transformed data, for example. Nevertheless, the selected ratio values facilitated a high level of discrimination between taxa that differed in terms of leaf base and leaf apex shape, and this discrimination improved where taxonomic groups were composed of phylogenetically distant taxa. Combined with the use of leaf shape bauplan, this method proved effective at identifying sterile specimens, and we are confident that this combined method is the most accurate yet described for identifying sterile Annonaceae specimens.

Shape recognition is more or less an intuitive procedure for the botanist, so whereas many typical leaf shapes can be distinguished from one another, it is often unclear what is the basis for this differentiation. By identifying the radians with greatest discriminatory power (in the *Uvaria* group, at least), these results focus the attention of the botanist on a few specific leaf-shape characters. With additional work, in particular the creation of larger databases and the utilization of digital imaging technology, leaf-shape analysis may become an invaluable aid for identifying the many sterile Annonaceae specimens that exist in herbaria.

ACKNOWLEDGEMENTS

We thank curators and staff of the many herbaria who loaned material for the study. We also thank colleagues in Thailand and Vietnam who assisted in field collection of new material, and the three referees for their helpful suggestions. C.M. was supported in this work by fellowships from Trinity College and South Dublin County Council, with additional support from the Trinity Trust (Dublin), the DA Webb Trust (Dublin), Smurfit Corporation (Dublin) and the Oleg Polunin Trust (UK).

REFERENCES

- Bookstein FL, Chernoff B, Elder R, Humphries J, Smith G, Strauss R. 1985.** *Morphometrics in evolutionary biology*. Special Publication 15. Philadelphia: The Academy of Natural Sciences of Philadelphia.
- Chatrou L. 1997.** Studies in Annonaceae XXVII. Macromorphological variation of recent invaders in northern Central America: the case of *Malmea* (Annonaceae). *American Journal of Botany* **84**: 861–869.
- Clausing G, Meyer K, Renner SS. 2000.** Correlations among fruit traits and evolution of different fruits within Melastomataceae. *Botanical Journal of the Linnean Society* **133**: 303–326.
- Dickinson TA, Parker WH, Strauss RE. 1987.** Another approach to leaf shape comparisons. *Taxon* **36**: 1–20.
- Doyle JA, Le Thomas A. 1994.** Cladistic analysis and pollen evolution in Annonaceae. *Acta Botanica Gallica* **141**: 149–170.
- Doyle JA, Le Thomas A. 1996.** Phylogenetic analysis and character evolution in Annonaceae. *Adansonia, Bulletin du Muséum national d'Histoire naturelle, Paris, Sér 4*, **18**: 279–334.
- Givnish TJ. 1979.** On the adaptive significance of leaf form. In: Solbrig OT, Jain S, Johnson GB, Raven PH, eds. *Topics in plant population biology*. New York: Columbia University Press, 375–407.
- van Heusden ECH. 1992.** Flowers of Annonaceae: morphology, classification and evolution. *Blumea* Suppl. **7**.
- Hodálová I, Marhold K. 1998.** Morphometric comparison of *Senecio germanicus* and *S. nemorensis* (Compositae) with a new species from Romania. *Botanical Journal of the Linnean Society* **128**: 277–290.
- Jensen RJ, Ciofani KM, Miramontes LC. 2002.** Lines, outlines, and landmarks: morphometric analyses of *Acer rubrum*, *Acer saccharinum* (Aceraceae) and their hybrid. *Taxon* **51**: 475–492.
- Jensen RJ, Hokanes SC, Isebrands JG, Hancock JF. 1993.** Morphometric variation in oaks of the Apostle Islands in Wisconsin: evidence of hybridisation between *Quercus rubra* and *Q. ellipsoidalis* (Fagaceae). *American Journal of Botany* **80**: 1358–1366.
- Jovet-Ast S. 1942.** Recherches sur les Annonacées D'Indochine. Anatomie foliaire – Répartition géographique. *Mémoires Du Muséum National d'Histoire Naturelle* **16**: 1–308.
- Kessler PJA. 1993.** Annonaceae. In: Kubitski S, ed. *Families and genera of vascular plants: 2 flowering plants: dicotyledons magnoliid, hamamelid and caryophyllid families*. Berlin: Springer Verlag, 93–129.
- Marcus LF. 1990.** Traditional morphometrics. In: Rohlf FJ, Bookstein FL, eds. *Proceedings of the Michigan morphometric workshop* Special Publication no. 2. Ann Arbor, MI: The University of Michigan Museum of Zoology, 77–122.
- Marcus LF, Bello E, García-Valdecasas A, eds. 1993.** *Contributions to morphometrics*. Monografías del Museo Nacional de Ciencias Naturales. Madrid: Museo Nacional de Ciencias Naturales.
- McLellan T, Endler JA. 1998.** The relative success of some methods for measuring and describing the shape of complex objects. *Systematic Biology* **48**: 192–199.
- Meade CV. 2000.** A systematic revision of the *Uvaria* group (Annonaceae) in continental Asia. Unpublished PhD Thesis, University of Dublin, Trinity College.
- Meade CV, Hodkinson T, Chalermglin P, Chase M, Parnell JAN. 2002.** Revision of *Uvaria* L. in Continental Southeast Asia: 2. Floral character evolution in the *Uvaria* L. Group. *Annonaceae Newsletter* **13**: 29–32.
- Parnell J, Needham M. 1998.** Morphometric variation in Irish *Sorbus* L. (Rosaceae). *Watsonia* **22**: 153–161.
- Premoli AC. 1996.** Leaf architecture of South American *Nothofagus* (Nothofagaceae) using traditional and new methods in morphometrics. *Botanical Journal of the Linnean Society* **121**: 25–40.
- Ray T. 1992.** Landmark and eigenshape analysis: homologous contours: leaf shape in *Syngonium* (Araceae). *American Journal of Botany* **79**: 69–76.
- Rumpunen K, Bartish IV. 2002.** Comparison of different estimates based on morphometric and molecular data, exemplified by various leaf shape descriptors and RAPDs in the genus *Chaenomeles* (Rosaceae). *Taxon* **51**: 69–82.
- Sandquist DR, Ehleringer JR. 1997.** Intraspecific variation of leaf pubescence and drought response in *Encelia farinosa* associated with contrasting desert environments. *New Phytologist* **135**: 635–644.
- van Setten AK, Koek-Noorman J. 1986.** Studies in Annonaceae VI. A leaf anatomical survey of genera of the Annonaceae in the Neotropics. *Botanische Jahrbücher Systematik, Pflanzengeschichte und Pflanzengeographie* **108**: 17–50.
- van Setten AK, Koek-Noorman J. 1992.** Studies in Annonaceae XVII. Fruits and seeds of Annonaceae. Morphology and its significance for classification and identification. *Bibliotheca Botanica* **142**: 1–101.
- Sneath PHA, Sokal RR. 1973.** *Numerical taxonomy*. San Francisco: W.H. Freeman.
- Sokal RR, Rohlf FJ. 1981.** *Biometry: the principles and practice of statistics in biological research*, 2nd edn. New York: W.H. Freeman.
- SPSS. 1997.** *SPSS® for Windows Release 8.0.0*. Chicago: SPSS Inc.
- Stace CA. 1989.** *Plant taxonomy and biosystematics*, 2nd edn. Cambridge: Cambridge University Press.

van Zuilen CM, Koek-Noorman J, Maas PJM. 1995. A phylogenetic analysis of *Duguetia* (Annonaceae) based on mor-

phological data. *Plant Systematics and Evolution* **194**: 173–188.

APPENDIX 1

Mean leaf shape outlines for 34 *Uvaria* group taxa. A. Semi-glabrous group. B. Hairs mixed group. C. Cuneate base group. D. Glabrous group. E. Veins > 20 group. F. Margins hirsute group. G. Scabrid group

