# Phylogenetic placement of the enigmatic Western Australian genus *Emblingia* based on *rbc*L sequences

#### GREGORY T. CHANDLER and RANDALL J. BAYER

CSIRO, Centre for Plant Biodiversity Research, Australian National Herbarium, GPO Box 1600, Canberra City, ACT 2601, Australia (Email: gregory.chandler@pi.csiro.au)

## Abstract

Assignment of the enigmatic Australian genus *Emblingia* to a particular family or order has been difficult. Informative morphological characters have not as yet been found to place *Emblingia* conclusively into a family, though it does share a number of attributes with the Capparaceae and Resedaceae. As a result, in the past it has been put in various families (Capparaceae, Sapindaceae, Goodeniaceae and Polygalaceae), representing a number of orders, as well as in its own family, the Emblingiaceae. The current molecular study, using *rbcL*, shows strong support for the placement of *Emblingia* within the Brassicales, and possibly sister to the Resedaceae. Further morphological and molecular studies within the Brassicales are needed before finalizing the familial placement of this genus. At this time, we consider treatment of *Emblingia* as a monotypic family, Emblingiaceae, within the order Brassicales the most satisfactory solution.

Keywords: chloroplast DNA, Emblingia, Emblingiaceae, phylogeny, rbcL.

Received 17 May 1999; revision received 26 July 1999; accepted 18 October 1999

## Introduction

As part of our systematic studies of the Polygalaceae of Australia for the "Flora of Australia" series, we need to address the familial placement of *Emblingia*. The 'Flora of Australia' follows Cronquist (1981), who places *Emblingia* in the Polygalaceae, a placement that is not widely accepted. It has been located in other families, representing a number of orders, namely the Capparaceae (von Mueller 1860), Sapindaceae (Thorne 1992), Goodeniaceae (Erdtman *et al.* 1969) and Polygalaceae (Cronquist 1981), and has also been placed in its own family, the Emblingiaceae (Airy Shaw 1965; Dahlgren 1980; Takhtajan 1980; APG 1998).

*Emblingia calceoliflora* was described by von Mueller in 1860 and was placed into the Capparaceae (tribe Capparideae), using the presence of a gynandrophore (a stalk supporting the androecium and gynoecium above the

insertion of the corolla) as the uniting character. Assignment of *Emblingia* to a family has been difficult because of conflicting morphological, anatomical, palynological and chemical data (Erdtman *et al.* 1969). Pollen morphology shows affinities with the Polygalaceae, floral morphology is closest to the Sapindaceae, while leaf and stem anatomy show similarities to the Goodeniaceae and the Polygalaceae. Affinities to the Capparaceae had largely been dismissed, mainly because previous authors felt that the other families shared more characters with *Emblingia* than does the Capparaceae (for example, Erdtman *et al.* 1969).

As morphology has failed to provide a clear answer to the phylogenetic position of *Emblingia*, this study used *rbcL* sequence data to determine its affinities.

## Materials and methods

#### DNA isolation, amplification and sequencing

DNA was extracted from herbarium material (CANB 251 220, A. S. George 9756) and purified according to methods outlined in Gilmore *et al.* (1993), except the amounts of components were scaled down for our pur-

Correspondence: Dr Gregory T. Chandler.

poses. The *rbc*L region was amplified via the polymerase chain reaction (PCR) using Taq DNA polymerase. The PCR reaction mixture consisted of 10 µL of 10X reaction buffer, 6 µL of 25 mmol/L magnesium chloride solution, 4 µL of a 10-mmol/L dNTP solution in equimolar ratio, 25 pmol of each primer, 10–50 ng of template DNA and 1.0 unit of polymerase in a total volume of 100 µL. The PCR samples were heated to 94°C for 3 min prior to the addition of DNA polymerase to denature unwanted proteases and nucleases. The double-stranded PCR products were produced via 30 cycles of denaturation (94°C for 1 min), primer annealing (48°C for 1 min) and extension (72°C for 2 min). A 7-min final extension cycle at 72°C followed the 30th cycle to ensure the completion of all novel strands.

The region was usually amplified as a single piece using primers 'Z1' and 'Z1351R' to amplify across the *rbcL* region. Double-stranded PCR products were cleaned by column purification using the Wizard PCR Preps DNA Purification System (Promega, Madison, WI, USA) prior to sequencing.

The double-stranded PCR products were then used as templates in cycle sequencing reactions which employed five primers across the region (kindly provided by G. Zurawski, Palo Alto, CA, USA), including the terminal primers 'Z-1' and 'Z-1351R' and, in addition, the internal primers 'Z-234', 'S-523' and 'Z-1020R'. The doublestranded PCR products were sequenced using the dideoxy chain termination method (Sanger et al. 1977) with the use of the Big Dye Terminator RR Kit (Perkin-Elmer Applied Biosystems, Norwalk, CT, USA) and an ABI automated sequencer in the Division of Plant Industry (CSIRO, Canberra, Australia). An annealing temperature of 60°C was used for all primers. The cycle sequencing protocol followed the manufacturer's instructions. Sequences were assembled and aligned using Sequencher 3.0 (Gene Codes Corporation, Ann Arbor, MI, USA).

# Taxon selection

In order to determine a preliminary position for *Emblingia*, we began with the *rbcL* matrix of 499 taxa from Chase *et al.* (1993) in an heuristic search using PAUP 4.0.0d65 (Swofford 1997) on a Macintosh G3. We then reduced the number of taxa to include the large clade including *Emblingia*. Family representation was expanded using 26 sequences from GenBank, as we included genera from within families that have included *Emblingia* in the past, namely Capparaceae, Goodeniaceae, Polygalaceae and Sapindaceae. For outgroups, we used lower dicots from the Berberidales, Hamamelidales and Nelumbonales, which are sister to the higher dicots in the analysis of Chase *et al.* (1993). We used the classification of Thorne (1992) for all family and ordinal circumscriptions.

#### Sequence data analysis

Final phylogenetic reconstruction was performed on unweighted characters by heuristic searches with simple, closest and furthest addition of taxa. Heuristic searches employing a random addition sequence of 100 replicates were also conducted to search for other islands of most parsimonious trees (Maddison 1991). A strict consensus tree (Margush & McMorris 1981) was constructed for the set of equally most parsimonious cladograms.

Bootstrap (Felsenstein 1985) analyses were used to estimate the robustness of clades. The bootstrap analysis employed 100 replicates of heuristic (SIMPLE addition sequence) searching. The amount of phylogenetic information in the parsimony analysis was assessed by use of the consistency index (CI; Kluge & Farris 1969) and the retention index (RI; Farris 1989).

## Results

#### Sequence characteristics

Length of the nearly complete *rbc*L gene that was sequenced was 1313 base pairs (out of ~1428 bp), beginning at base pair number 56902 of the *Zea* chloroplast genome map (Maier *et al.* 1995) and ending at base pair number 58218. No indels were encountered when aligning the sequence with the main matrix of Chase *et al.* (1993). The G/C content is 44%. The sequence has been submitted to GenBank (accession number AF146014).

# Phylogenetic reconstruction

The initial search performed using the matrix of 499 taxa from Chase *et al.* (1993) showed *Emblingia* clearly allied to the Brassicales. The other taxa we included in our analyses were then added to this matrix, where *Emblingia* continued to come out in the Brassicales. The number of taxa were then reduced to focus the search on the Brassicales and other proposed relatives, as well as all taxa that were to be found as sister to these groups in the initial analysis, plus outgroups. This was done in order to perform a more exhaustive search. This left a total of 86 taxa in the analysis.

The phylogenetic analysis of the sequence data of the reduced data set yielded 42 equally parsimonious trees of 2455 steps (CI=0.29; RI=0.65; Fig. 1). One of the most parsimonious trees that is topologically identical to the strict consensus tree is presented in Fig. 1 (with the three minor branches that collapse in the strict tree marked with a



**Fig. 1** One shortest tree of 42 equally parsimonious trees of length 2455 (steps) resulting from phylogenetic analysis of sequence data of the *rbcL* gene that is topologically identical to the strict consensus tree. Dashed lines indicate branches that did not appear in the strict consensus tree. Clades that are discussed in the text are labeled with capital letters. The number of synapomorphies are indicated above the branch, with bootstrap values shown below. Bootstrap values less than 50% not shown. The consistency index (CI) was 0.29, and the retention index (RI) was 0.65. Also shown are the families to which the various taxa belong, and the order to which these families belong, according to Thorne (1992), except for *Emblingia*, which he places in the Sapindaceae, of the order Rutales. *Emblingia* has been emboldened, and the families Capparaceae, Goodeniaceae, Polygalaceae and Sapindaceae have been emboldened and italicized.

dashed line). Island searches (Maddison 1991) on the data sets did not reveal any islands of trees of shorter length.

## Topology of major clades

Beginning at the base of the tree (clade A) and proceeding systematically toward the top of the tree (clade K), a number of well supported groups are evident. The basal group of this analysis is clade A (synapomorphies (syn) = 43; bootstrap value (bsv) = 100), which contains members of the family Goodeniaceae (Campanulales). Clade B (syn =6; bsv=99) contains members of the Geraniales (Geraniaceae, Oxalidaceae) as well as the Crossosomataceae (Rosales), while clade C (syn=8; bsv=56) also contains members of the family Geraniaceae, plus members of the Greyiaceae and Francoaceae (Saxifragales).

Members of the Rutales (Fabaceae) and Polygalales (Polygalaceae) make up clade D (syn=22; bsv=97), whereas clade E (syn = 17; bsv = 100) contains genera from the order Myrtales, including the families Combretaceae, Lythraceae, Myrtaceae, Melastomataceae, Onagraceae and Punicaceae. The order Malvales makes up clade F (syn = 14; bsv = 79) and comprises the families Bombacaceae, Dipterocarpaceae, Malvaceae, Sterculiaceae and Tiliaceae, while clade G (syn=23; bsv=100) contains members of the Rutales and comprises the families Aceraceae, Anacardiaceae, Burseraceae, Hippocastanaceae, Leitneriaceae, Rutaceae, Sapindaceae and Simaroubaceae. Clade H (syn = 13; bsv = 100) contains members of the Rutales (Akaniaceae and Bretschneideraceae) and clade I (syn = 7; bsv = 88) contains members of the orders Rutales (Moringaceae) and Violales (Caricaceae).

The polyordinal clade J (syn=20; bsv=99) contains members of the orders Batales, Brassicales and Geraniales, with clade K (syn=7; bsv=56) consisting solely of members of the Brassicales (Brassicaceae, Capparaceae, Gyrostemonaceae and Resedaceae). The topology of this clade is comparable to that same clade seen in fig. 10a in Chase *et al.* (1993).

# Position of Emblingia

A consensus of all trees indicates that *Emblingia* belongs within the Brassicales (Capparales) (Fig. 1; clade J; syn = 20; bsv=99), and sister to Reseda (Resedaceae). The *Reseda*/*Emblingia* clade is sister to *Tovaria* (Capparaceae), and this larger clade is in turn sister to Pentadiplandra (Capparaceae). The *Emblingia*/*Reseda* clade has weak support (syn=9; bsv= < 50). The addition of *Emblingia* does not change the basic topology of this clade as shown in fig. 10a of Chase *et al.* (1993).

Table 1 shows comparisons between *Emblingia* and supposedly related families, based on data obtained primarily from Erdtman *et al.* (1969), Cronquist (1981), Rodman (1991) and Ronse Decraene & Smets (1997). Table 2 lists the 27 taxa and their GenBank accession numbers that were added to this analysis, 26 which are not found in the appendix that follows the Chase *et al.* (1993) paper, plus the sequence obtained for *Emblingia* in this analysis.

## Discussion

The analysis has shown that *Emblingia* has strong brassicalean affinities and has little relationship with the families Polygalaceae, Sapindaceae or Goodeniaceae.

**Table 1** Distribution of selected diagnostic morphological characters used to assess the relationship of *Emblingia* with the families Capparaceae, Goodeniaceae, Polygalaceae, Resedaceae and Sapindaceae. Character states that are bold and underlined are those that are shared with *Emblingia* 

	Emblingia	Capparaceae	Resedaceae	Sapindaceae	Polygalaceae	Goodeniaceae
Floral symmetry	zygomorphic	actinomorphic, some <u>zygomorphic</u>	zygomorphic	± <u>zygomorphic</u> , occ. actinomorphic	zygomorphic	zygomorphic
Petal number	<u>2</u>	( <u>2</u> ), 4 (6)	(0, <u>2</u> , 4–8), 6	4, 5	3, (5)	5
Gynandrophore	present	mostly present	present	absent, <u>ex. <i>Serjania</i></u>	absent	absent
Stamen number	<u>4</u>	4, sometimes 36	3-50+	8 ( <u>4</u> –10)	8	5
Locule number	<u>2–3</u>	<u>1–4</u>	1	<u>2–3</u>	(1), <b><u>2</u></b> , 5, 7, 8	(1), <b>2</b> , (4)
Placentation	axile	parietal	<b>axile</b> , occ parietal	axile	mostly <b>axile</b>	<u>axile,</u> basal axile
Replum	absent?	present or <u>absent</u>	absent	absent	absent	<u>absent</u>
Embryo orientation	<u>curved</u>	curved	<u>curved</u>	curved	straight	straight
Mustard oils	unknown	present	present	absent	absent	absent

© 2000 The Society for the Study of Species Biology Plant Species Biology, 15, 67–72

**Table 2** List of species (with family placement) with GenBank accession numbers for those taxa not found in the appendix of the *Annals of the Missouri Botanical Gardens* Vol. 80, no. 3

Taxon	GenBank accession number		
Aesculus pavia (Hippocastanaceae)	U39277		
Anthotium rubiflorum (Goodeniaceae)	X87378		
Azima tetracantha (Salvadoraceae)	U36782		
Brunonia australis (Goodeniaceae)	X87380		
Capsella bursa-pastoris (Brassicaceae)	D88904		
Comesperma ericinum (Polygalaceae)	L29492		
Coopernookia strophiolata (Goodeniaceae)	X87381		
Diaspasis filifolia (Goodeniaceae)	X87384		
Dipteronia sinensis (Sapindaceae)	U39268		
Dodonaea triquetra (Sapindaceae)	U39822		
Emblingia calceoliflora (Emblingiaceae)	AF146014		
Francoa sonchifolia (Francoaceae)	L11184		
Goodenia ovata (Goodeniaceae)	X87386		
Gyrostemon sp. Cranfield 02068672	L22439		
(Gvrostemonaceae)			
Koelreuteria paniculata (Sapindaceae)	U39283		
Lechenaultia heteromera (Goodeniaceae)	X87388		
Nasturtium officinale (Brassicaceae)	AF020325		
Punica granatum (Lythraceae)	L10223		
Pentadiplandra brazzeana (Capparaceae)	U38533		
Salvadora angustifolia (Salvadoraceae)	U38532		
Selliera radicans(Goodeniaceae)	X87395		
Spondias cytherea (Anacardiaceae)	U39274		
Tapirira mexicana (Anacardiaceae)	U39273		
Tersonia cyathiflora (Gyrostemonaceae)	L22441		
Tetracentron sinense (Trochodendraceae)	L12668		
Velleia paradoxa (Goodeniaceae)	X87396		
Verreauxia reinwardtii (Goodeniaceae)	X87397		

Conflicting morphological data from numerous sources (see von Mueller 1861; Bentham 1863; Pax & Hoffman 1936; Erdtman et al. 1969; Ronse Decraene & Smets 1997) have failed to place Emblingia unequivocally into a family. Indeed, the reports seem only to have added to the confusion. An example of this can be found in Erdtman et al. (1969), where the authors compared *Emblingia* with the families Capparaceae, Goodeniaceae, Polygalaceae and Sapindaceae. Erdtman suggested that pollen morphology showed that Emblingia is most similar to the Polygalaceae, while Leins (in Erdtman et al. 1969) concluded that Emblingia was most similar to the Sapindaceae with respect to floral morphology. Using floral anatomy, Melville (in Erdtman et al. 1969) suggested Emblingia resembles Scaevola (Goodeniaceae), whereas Metcalfe (in Erdtman et al. 1969), based on stem and leaf anatomy, concluded that it was closest to the Goodeniaceae, with the next closest family being the Polygalaceae. None of these authors found a particularly close affinity between *Emblingia* and the Capparaceae.

There were suggestions in Erdtmann *et al.* (1969) that *Emblingia* may be best placed within its own family, the Emblingiaceae. The Emblingiaceae had in fact already been published by Airy Shaw (1965), raised from what had been considered a subfamily of the Capparaceae (the Emblingioideae, Pax 1891), a fact which seems to have been overlooked by Erdtman *et al.* (1969).

Several morphological characters have been used to argue the placement of *Emblingia* in particular families (Table 1). It can be seen that *Emblingia* has zygomorphic flowers, as do all the relevant families in this analysis, although some of them have only a few taxa with this state (as in the Capparaceae). Hence, this is not a good character for determining the placement of *Emblingia*. In regard to petal number, the only other families that have any members with only two petals, as in *Emblingia*, are the Capparaceae and the Resedaceae, and the only other family with 4-merous corolla is the Sapindaceae, with the Goodeniaceae and Polygalaceae having three or five (Table 1).

A major character is the presence of a gynandrophore in Emblingia, which it shares with the Capparaceae, Resedaceae and rarely the Sapindaceae, but which is absent in the Goodeniaceae and Polygalaceae (Table 1). The presence of four stamens also allies Emblingia with the Capparaceae and to a lesser extent the Resedaceae (three to 50-plus stamens), and the Sapindaceae (mostly eight but sometimes four stamens; Table 1). The two-tothree-loculate ovary with axile placentation is most consistent with Sapindaceae (Table 1) but combinations of these characters also occur in the other families. The presence of a replum in Emblingia appears to be controversial, with Ronse Decraene and Smets indicating that it is present, while Pax & Hoffmann (1936) and Erdtman et al. (1969) say that it is absent. If present, it would ally Emblingia with the Capparaceae, which is the only family to have this character (Table 1), while if absent, no affinity would be indicated with any particular family.

Curved embryos, as occur in *Emblingia*, are also found in the Capparaceae, Resedaceae and Sapindaceae but not in the Goodeniaceae or Polygalaceae (Table 1). Although the presence of glucosinolates (mustard oils) have apparently not been investigated in *Emblingia* (it is absent in the analysis by Rodman 1991), the presence of such oils would further strengthen the alliance of *Emblingia* with the Capparaceae and Resedaceae (Table 1). Glucosinolates are also present in the brassicalean genera *Reseda* and *Tovaria* (Rodman 1991), the sister taxa of *Emblingia* in this analysis (Fig. 1). Rodman (1996) does not show Reseda and *Tovaria* to be sister taxa in his cladistic analysis, though they are near to each other; otherwise the topology of his tree is the same as the one obtained here.

In summary, past morphology has been unable to place *Emblingia* conclusively into a family. It does share a

number of characters with the Capparaceae and Resedaceae (Table 1) and perhaps the familial affinity of Emblingia lies with these families. There is strong molecular support for the placement of Emblingia within the brassicalean clade (Fig. 1; clade J; syn = 20, bsv = 99). This analysis also shows somewhat inconclusive support for a sister relationship between Emblingia and the Resedaceae, with only nine synapomorphies supporting the clade (Fig.1). Further morphological and molecular studies within the Brassicales, especially within the Capparaceae, which is being shown to be paraphyletic, and the Resedaceae, will be needed before a final statement can be made. For example, the Angiosperm Phylogeny Group (1998) sink the Capparaceae into the Brassicaceae, indicating that further work is required within this group.

We consider the placement of *Emblingia* within its own family within the Brassicales as the most satisfactory solution until such studies have been undertaken. Hopefully, our future morphological and phytochemical studies on *Emblingia* will further elucidate the relationships in this clade.

#### Acknowledgements

We thank G. Zurawski (Palo Alto, CA, USA) for providing the sequence primers, and Laurie Adams, Lindy Cayzer, Rogier de Kok and Joe Miller (CSIRO, Canberra) for providing critical reviews of our manuscript.

## References

- Airy Shaw H. K. (1965) Diagnoses of new families, new names, etc., for the seventh edition of Willis's Dictionary. *Kew Bulletin* 18: 249–273.
- Angiosperm Phylogeny Group (1998) An ordinal classification for the families of flowering plants. Annals of the Missouri Botanical Garden 85: 531–553.
- Bentham G. (1863). Polygaleae In: *Flora Australiensis* Vol. 1, pp. 137–149. Lovell Reeve, London.
- Chase M. W., Soltis D. E., Olmstead R. R. G., Morgan D., Les D.
  H., Mishler B. D., Duvall M. R., Price R. A., Hills H. G., Qiu Y. -L., Kron K. A., Rettig J. H., Conti E., Palmer J. D., Manhart J. R., Sytsma K. J., Michaels H. J., Kress W. J., Karol K. G., Clark W. D., Hedren M., Gaut B. S., Jansen R. K., Kim K.-J., Wimpee C. F., Smith J. F., Furnier G. R., Strauss S. H., Xiang Q.-Y., Plunkett G. M., Soltis P. S., Swensen S. M., Williams S. E., Gadek P. A., Quinn C. J., Eguiarte L. E., Golenberg E., Learn G. H., Graham S. W. Jr, Barrett S. C. H., Dayanandan S., Albert

V. A. (1993) Phylogenetics of seed plants: an analysis of nucleotide sequences form the plastid gene *rbcL. Annals of the Missouri Botanical Garden* **80**: 528–580.

- Cronquist A. J. (1981) An Integrated System of Classification of Flowering Plants. Columbia University Press, New York.
- Dahlgren R. (1980) A revised system of classification of the angiosperms. Botanical Journal of the Linnean Society 80: 91–124.
- Erdtman G., Leins P., Melville R., Metcalfe C. R. (1969) On the relationships of *Emblingia*. *Botanical Journal of the Linnean Society* **62**: 169–186.
- Farris J. S. (1989) The retention index and the rescaled consistency index. *Cladistics* 5: 417–419.
- Felsenstein J. (1985) Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* **39**: 783–791.
- Gilmore S., Weston P. H., Thomson J. A. (1993) A simple, rapid, inexpensive, and widely applicable technique for purifying plant DNA. *Australian Systematic Botany* 6: 139–148.
- Kluge A. G., Farris J. S. (1969) Quantitative phyletics and the evolution of Anurans. *Systematic Zoology* 18: 1–32.
- Maddison D. R. (1991) The discovery and importance of multiple islands of most parsimonious trees. *Systematic Zoology* 40: 315–328.
- Maier R. M., Neckermann K., Igloi G. L. (1995) Complete sequence of the maize chloroplast genome: gene content, hotspots of divergence and fine tuning of genetic information by transcript editing. *Journal of Molecular Biology* 251: 614–628.
- Margush T., McMorris F. R. (1981) Consensus n-trees. Bulletin of Mathematical Biology 43: 239–244.
- Pax F. (1891) Capparidaceae. In: *Die Natürlichen Pflanzenfamilien III* (eds E.Engler, K.Prantl), Vol. 2, pp. 209–236. Engelmann, Leipzig.
- Pax F., Hoffmann K. (1936) Capparidaceae. In: Die Natürlichen Pflanzenfamilien (eds E.Engler, K.Prantl), Vol. 17b, pp. 146–223. Engelmann, Leipzig.
- Rodman J. E. (1991) A taxonomic analysis of glucosinolateproducing plants, part 2: cladistics. *Systematic Botany* 16: 619–629.
- Rodman J. E. (1996) Molecules, morphology, and Dahlgrenís expanded order Capparales. *Systematic Botany* 21: 289–308.
- Ronse Decraene L. P., Smets E. F. (1997) Evidence for carpel multiplications in the Capparaceae. *Belgian Journal of Botany* 130: 59–67.
- Sanger F., Nicklen S., Coulson A. R. (1977) DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Science USA* 74: 5463–5467.
- Swofford D. (1997) PAUP: Phylogenetic Analysis Using Parsimony, Pre-Release, Version 4.0.0d64. Laboratory of Molecular Systematics, Smithsonian Institution, Washington, DC & Sinauer, Sunderland, MA.
- Takhtajan A. L. (1980) Outline of the classification of flowering plants Magnoliophyta. *Botanical Review* 46: 225–359.
- Thorne R. T. (1992) Classification and geography of the flowering plants. *Botanical Review* 58: 225–348.
- von Mueller F. J. H. (1860) Fragmenta Phytographiae Australiae, Vol. 2, pp. 2–4. Gubern, Coloniae, Victoria.