A dated molecular perspective of eucalypt taxonomy, evolution and diversification

Andrew H. Thornhill[A,B,C,H], *Michael D. Crisp*[D], *Carsten Külheim*[D,E], *Kristy E. Lam*[A],
Leigh A. Nelson[F], *David K. Yeates*[F] *and Joseph T. Miller*[A,G]

[A]Centre for Australian National Biodiversity Research, National Research Collections, Black Mountain, CSIRO,
  Canberra, ACT 2601, Australia.
[B]Australian Tropical Herbarium, James Cook University, Cairns, Qld 4870, Australia.
[C]University and Jepson Herbaria, and Department of Integrative Biology, University of California, Berkeley,
  CA 94720-2465, USA.
[D]Division of Evolution and Ecology, Research School of Biology, Australian National University, Canberra,
  ACT 2601, Australia.
[E]School of Forest Resources and Environmental Science, Michigan Technological University, Houghton,
  MI 49931, USA.
[F]Australian National Insect Collection and Taxonomic Research and Information Network,
  National Research Collections, Black Mountain, CSIRO, ACT 2601, Australia.
[G]Office of International Science and Engineering, National Science Foundation, Alexandria, VA 22314, USA.
[H]Corresponding author. Email: andrew.thornhill@gmail.com

Abstract.   The eucalypts, which include *Eucalyptus*, *Angophora* and *Corymbia*, are native to Australia and Malesia and include over 800 named species in a mixture of diverse and depauperate lineages. We assessed the fit of the eucalypt taxonomic classification to a phylogeny of 711 species scored for DNA sequences of plastid *mat*K and *psb*A–*trn*H, as well as nuclear internal transcribed spacer and external transcribed spacer. Two broadly similar topologies emerge from both maximum likelihood and Bayesian analyses, showing *Angophora* nested within *Corymbia* or *Angophora*, sister to *Corymbia*. The position of certain species-poor groups on long branches fluctuated relative to the three major *Eucalyptus* subgenera, and positions of several closely related species within those subgenera were unstable and lacked statistical support. Most sections and series of *Eucalyptus* were not recovered as monophyletic. We calibrated these phylogenies against time, using penalised likelihood and constraints obtained from fossil ages. On the basis of these trees, most major eucalypt subgenera arose in the Late Eocene and Early Oligocene. All *Eucalyptus* clades with taxa occurring in south-eastern Australia have crown ages <20 million years. Several eucalypt clades display a strong present-day geographic disjunction, although these clades did not have strong phylogenetic statistical support. In particular, the estimated age of the separation between the eudesmids (*Eucalyptus* subgenus *Eudesmia*) and monocalypts (*Eucalyptus* subgenus *Eucalyptus*) was consistent with extensive inland water bodies in the Eocene. Bayesian analysis of macroevolutionary mixture rates of net species diversification accelerated in five sections of *Eucalyptus* subgenus *Symphyomyrtus*, all beginning 2–3 million years ago and associated with semi-arid habitats dominated by mallee and mallet growth forms, and with open woodlands and forests in eastern Australia. This is the first time that a calibrated molecular study has shown support for the rapid diversification of eucalypts in the recent past, most likely driven by changing climate and diverse soil geochemical conditions.

Additional keywords: Bayesian analysis of macroevolutionary mixtures (BAMM), eucalypts, molecular dating, Myrtaceae, phylogenetics.

Received 25 March 2018, accepted 14 February 2019, published online dd mmm yyyy

Dedicated to Ian Brooker (1934–2016) – 'Don't use the 'C'[orymbia] word with me Andrew'.
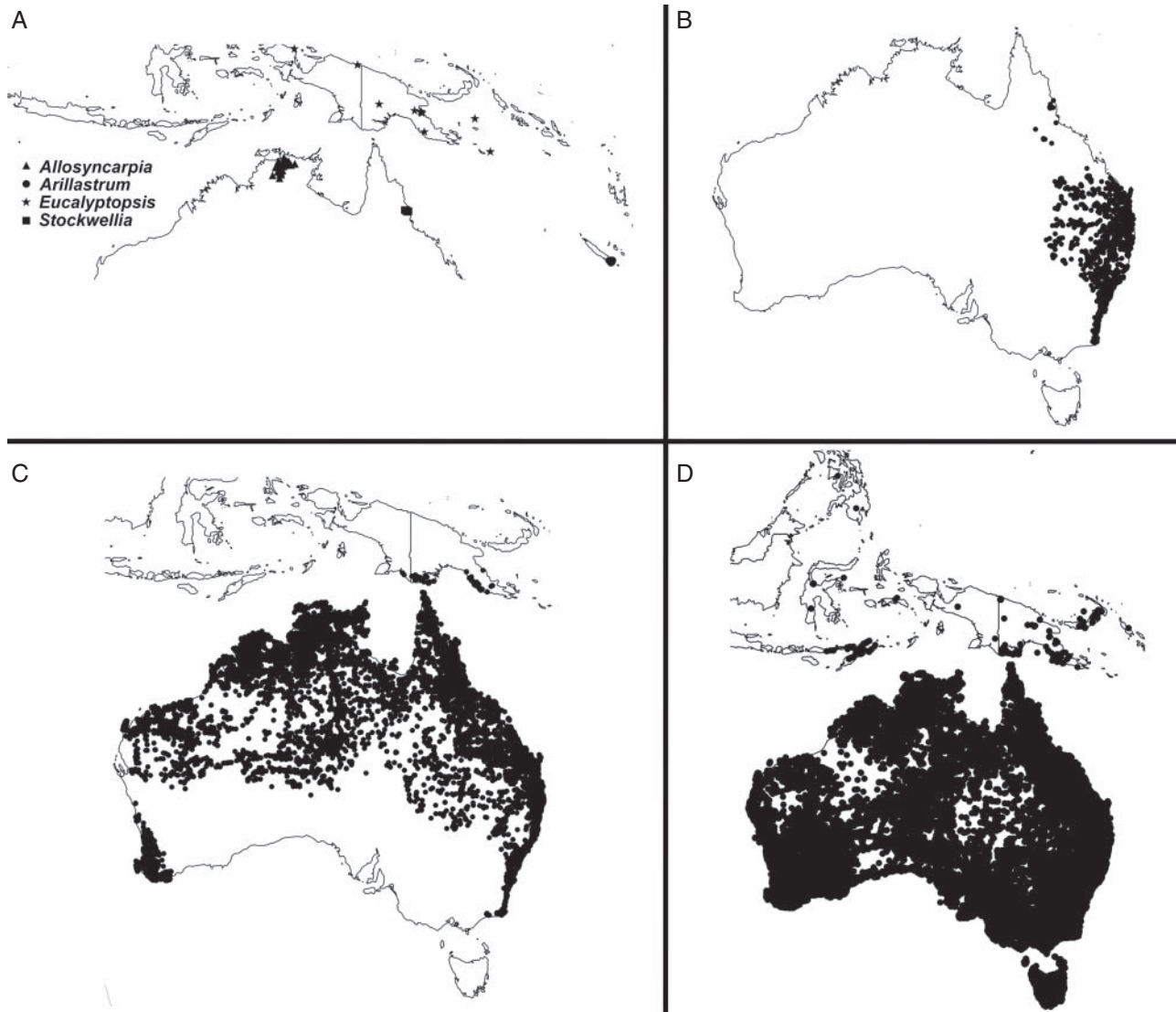
Fig. 1. The modern distribution of the seven genera included in Eucalypteae, highlighting the widespread *Eucalyptus*, the tropical, subtropical, and monsoonal distribution of *Angophora* and *Corymbia*, and the ranged-restricted rainforest genera. Distribution is represented by herbarium records from the Australasian Virtual Herbarium and Global Biodiversity Information Facility (GBIF). A. The rainforest genera *Allosyncarpia*, *Arillastrum*, *Eucalyptopsis* and *Stockwellia* occur only in northern Australia, New Guinea and New Caledonia. B. *Angophora* is found only on the eastern coast of Australia. C. *Corymbia* occurs in Australia and New Guinea. D. *Eucalyptus* occurs in Australia, New Guinea, Timor, Indonesia and the Philippines.

Introduction

Eucalypts grow as trees, mallees (multi-stemmed large shrubs), mallets (single-stemmed small trees lacking lignotubers) and, occasionally, small shrubs, in most vegetation communities of Australia. Eucalypts dominate most treed communities, except the subtropical and tropical rainforests of eastern Australia, the coastal fringe mangroves and the *Acacia*-dominated mulga woodlands of Australia's dry interior. The name 'eucalypt' has been applied to multiple genera within tribe Eucalypteae (Wilson 2011), but most commonly (including here) to the following three genera: *Eucalyptus* L'Her. (700+ species),

Corymbia K.D.Hill & L.A.S.Johnson (100+ species) and *Angophora* Cav. (12 species). *Eucalyptus* is primarily distributed in Australia with a small number of taxa, including endemics, in Indonesia, the Philippines and Papua New Guinea (Ladiges 1997; Fig. 1). *Corymbia* occurs in warm-temperate mainland Australia and New Guinea (Hill and Johnson 1995) and *Angophora* is restricted to eastern mainland Australia (Slee *et al.* 2006). Other Eucalypteae genera are *Arillastrum* Pancher & Baill. from New Caledonia (1 taxon), *Allosyncarpia* Blake from the Northern Territory (1 species), *Stockwellia* Carr, Carr & Hyland from far-northern Queensland (1 taxon), and *Eucalyptopsis* White from New Guinea (2 species).

Eucalypt classification

Eucalyptus obliqua was the first described eucalypt by the French botanist, Charles-Louis L'Héritier de Brutelle, in 1789 (L'Héritier de Brutelle 1789). Eight years later, *Angophora* was described by the Spanish botanist Antonio Cavanilles (Cavanilles 1797). Then more than 200 years later, Ken Hill and Lawrie Johnson (Hill and Johnson 1995) transferred the bloodwoods and ghost gums from *Eucalyptus* into a new genus named *Corymbia*. The eucalypts have had numerous systematic treatments and multiple classifications have been proposed to group eucalypt taxa on the basis of morphological characters; an excellent summary of most of these is provided by Ladiges (1997).

The current classification of the eucalypts is formed around two major works. Hill and Johnson (1995) segregated the genus *Corymbia* from *Eucalyptus sensu lato* (*s.l.*). This work formed a part of a larger classification scheme (Johnson and Hill 1999, 1990, 1991; Hill and Johnson 1991a, 1991b, 1992, 1994, 1995, 1998, 2000; Hill et al. 2001) that maintained *Angophora* as a separate genus and arranged *Eucalyptus* species into seven sections, similar to the informal classifications of Pryor and Johnson (1971). This classification was published in its entirety as the New South Wales Herbarium-based website Eucalink (K. D. Hill, see http://plantnet.rbgsyd.nsw.gov.au/PlantNet/Euc/, accessed 24 February 2017). This website is no longer maintained because of the untimely death of Ken Hill, meaning that all species described since 2004 are not included on the website.

In an alternative classification, Ian Brooker (2000) also followed the concepts of Pryor and Johnson, but with one fundamental difference, namely, all species of *Eucalyptus*, *Angophora* and *Corymbia* were placed into only one genus, *Eucalyptus*, although this was contrary to phylogenetic evidence (Ladiges and Udovicic 2000). Brooker arranged his classification in a hierarchical system of subgenera, sections, subsections, series and subseries. The data underlying this work were published in the Euclid interactive key but differed from the 2000 treatment by accepting the three-genus classification (Slee et al. 2006).

The most recent eucalypt classifications have been by Nicolle (2015a, 2015b, 2018), who retained the three-genus classification of Johnson and Hill, synonymised some of Brooker's subgenera, reclassified some sections and series, and interpolated many species named since Brooker's 2000 treatment.

Phylogenetic studies

Since the late 1990s, molecular phylogenetic approaches have been employed to deduce the relationship of groups among and within the eucalypts (Johnson 1972; Ladiges et al. 1995, 2003; Udovicic et al. 1995; Steane et al. 1999, 2002; Whittock et al. 2003; Crisp et al. 2004, 2011; Parra-O et al. 2006; Gibbs et al. 2009; Bayly et al. 2013). Much focus has been on resolving the relationships among the three genera, as well as the many subgenera, and results have been variable depending on the molecular marker used and taxa included in the analysis (Udovicic et al. 1995; Steane et al. 1999, 2002; Parra-O et al. 2006, 2009; Ochieng et al. 2007; Ladiges et al. 2010; Bayly et al. 2013). Many of the phylogenies from these studies show a monophyletic *Angophora* that is sister to a monophyletic *Corymbia,* with these two genera in turn being sister to a monophyletic *Eucalyptus*. However, other work, both recent and past, suggests that *Angophora* could be nested within a paraphyletic *Corymbia*, with this clade being sister to *Eucalyptus* (Steane et al. 1999; Bayly et al. 2013; González-Orozco et al. 2016; Schuster et al. 2018).

Evolutionary history

The eucalypts are an old Gondwanan lineage. Two reviews of the fossil record (Hill et al. 2016; Macphail and Thornhill 2016) found that evidence of eucalypts extends back to the Early Eocene. Both the fossil pollen record in Australia (Thornhill and Macphail 2012), and macrofossils in Patagonia of South America (Gandolfo et al. 2011; Hermsen et al. 2012) and Australia (Lange 1978; Ambrose et al. 1979) have recognisable eucalypt material. The oldest fossil-pollen morphotype, *Myrtaceidites tenuis*, is most similar in appearance to the extant *Corymbia* and *Angophora* pollen types (Thornhill et al. 2012a), whereas the oldest macrofossils have similarities with *Eucalyptus* and *Corymbia* (Lange 1978; Rozefelds 1996), and *E*. subg. *Symphyomyrtus* (Gandolfo et al. 2011). Further, molecular dating of a small representative sample of eucalypt taxa also arrived at an estimated Eocene age (Crisp et al. 2011; Thornhill et al. 2012a, 2015). At a higher level, it is estimated that the Eucalypteae tribe diverged from other Myrtaceae sometime in the Palaeocene (~60 million years ago; Thornhill et al. 2012a, 2015; Berger et al. 2016).

Eucalypt fossils have been recovered from most East Gondwanan land-masses, and it is assumed from the microfossil record that they did not become a dominant vegetation element in Australia until after the Oligocene (fig. 6.06 in Partridge 1999), a trend that is also shown more generally by Myrtaceae (Macphail and Truswell 1989; Martin 1991; Truswell 1993; Hill et al. 1999). A drying climate and an increase in fire frequency from the Oligocene to the Middle Miocene is considered to be the trigger of multiple pulses of expansion and contractions of *Eucalyptus* species (Wardell-Johnson et al. 1997), with further aridification in the Plio-Pleistocene (Potts and Pederick 2000; Macphail and Thornhill 2016), culminating in the modern vegetational domination over most of the Australian landscape.

In the present study, we examine the congruence of phylogenetics and taxonomy, by using a sample set of 711 taxa from *Angophora*, *Corymbia* and *Eucalyptus*. Specifically, we use the most recent eucalypt classification to see whether recognised infrageneric taxa can be recovered as monophyletic. Further, we explore the broad biogeography and evolutionary history of the eucalypts by using molecular dating and diversification-rate analyses.

Materials and methods

Terminology and taxonomy

The classification of Nicolle (2015b) was selected as the taxonomy of the present study. In addition, in the hope of forming a consensus for future works, we define three common names for groups in the Eucalypteae. Mesicalypt (mesic [eu]calypt) applies to the three rainforest genera

Allosyncarpia, *Eucalyptopsis* and *Stockwellia*, all of which occur in northern Australian or New Guinea. Newcalypt (New Caledonia [eu]calypt) applies to the sole genus, *Arillastrum*, that grows only in New Caledonia. Eucalypt applies to the three Australia-centric genera of *Angophora*, *Corymbia* and *Eucalyptus* and contains the overwhelming majority of species in the Eucalypteae. Common names of the three big *Eucalyptus* subgenera have also been previously used and we use them in the discussion of the present paper. Eudesmid applies to *Eucalyptus* subgenus *Eudesmia*. Monocalypt applies to *Eucalyptus* subgenus *Eucalyptus* and Symphyomyrt applies to *Eucalyptus* subgenus *Symphyomyrtus*. A scientific and common name and classification table is included (Table 1), to help those who are more familiar with the common names of smaller existing eucalypt groups (e.g. *Maidenaria* = red gums) interpret this paper.

Sampling

Genetic sequences were gathered for 732 taxa, including 711 eucalypt species (some of which are now considered subspecies), the single *Arillastrum* species, all four mesicalypt species and a diverse range of Myrtaceae outgroups used in previous studies (Wilson *et al*. 2005; Biffin *et al*. 2006; Thornhill *et al*. 2012a, 2015). Fresh leaves for DNA extraction were collected in the field or from cultivated plants of known origin growing in various arboreta or botanic gardens of south-eastern Australia. When fresh leaf material was not available, leaves were obtained from herbarium specimens housed in the Australian National Herbarium (CANBR). A table of the taxonomic status of each taxon used in the study, along with GenBank accessions and their sources or vouchers, is included in Table S1, available as Supplementary material to this paper.

Sequencing

Nuclear internal transcribed spacer (ITS) and external transcribed spacer (ETS) sequences from many previous eucalypt studies were downloaded from GenBank and used to determine which species needed to be newly sequenced. We aimed to sample every described eucalypt species for both nuclear and plastid markers. Species with an existing ITS or ETS sequence were not resampled. Four plastid regions were trialled to determine which amplified most easily. Of these, *mat*K, which has been used effectively to address systematic questions in Myrtaceae (Gadek *et al*. 1996), and the *psb*A–*trn*H intergenic spacer, were selected and sequenced using protocols and primers described by Gadek *et al*. (1996) for *mat*K and Lucas *et al*. (2005) for *psb*A–*trn*H. It was hoped that the use of plastid markers in concatenation with the previously utilised nuclear markers would better resolve relationships. ITS and ETS sequences were generated using the protocols and primers of Sun *et al*. (1994) and White *et al*. (1990) for ITS, and Lucas *et al*. (2007) and Wright *et al*. (2001) for ETS. Screening of ITS and ETS was performed by making single-locus phylogenies and excluding from any further analyses any newly generated sequence that formed a clade with sequences that were suggested to be pseudogenes by Ochieng *et al*. (2007). In total, 2317 new sequences from 686 species were generated

(see Table S1). Contiguous sequences were assembled and edited using Sequencher (ver. 3.0, Gene Codes Corporation, see http://www.genecodes.com/, accessed 24 February 2017). Alignments were constructed using MUSCLE (see https://www.drive5.com/muscle/, accessed 24 February 2017; Edgar 2004,) and manually corrected in Se-Al (ver. 2.0, A. Rambaut, see http://tree.bio.ed.ac.uk/software/seal/, accessed 24 February 2017).

Phylogenetic analyses

The four loci were concatenated into one alignment using SequenceMatrix (ver. 1.79, see http://www.ggvaidya.com/taxondna/, accessed 24 February 2017; Vaidya *et al*. 2011) that contained 3158 base pairs (available from the CSIRO Data Access Portal at https://data.csiro.au/dap/landingpage?pid=csiro:33546, accessed 22 February 2019). This is the same alignment that was used by González-Orozco *et al*. (2016). Maximum-likelihood (ML) analysis was performed on concatenated, plastid marker only, and nuclear marker-only alignments in the CIPRES Portal (M. A. Miller, M. T. Holder, R. Vos, P. E. Midford, T. Liebowitz, L. Chan, P. Hoover and T. Warnow, see http://www.phylo.org/, accessed 24 February 2017), using the RAxML (ver. 7.2.7, xxx; Stamatakis 2014) HPC BlackBox tool with a GTR + Gamma + I partition model on each locus. A further 10 analyses were run on the concatenated dataset to test for uncertain tree topology by using the HPC2 on XSEDE in CIPRES, in which each run started with a different seed number. The resulting bipartition trees were saved and FigTree (ver. 1.4.2, A. Rambaut, see http://tree.bio.ed.ac.uk/software/figtree/, accessed 24 February 2017) was used to view trees and generate figures.

To further explore tree topology, we ran a Bayesian analysis using MrBayes (ver. 3.2, xxx; Huelsenbeck and Ronquist 2001) on the CSIRO Burnett supercomputer cluster. All Myrtaceae outgroups except *Heteropyxis* and terminals that formed polytomies in the ML analyses were removed from the alignment to save computational time, leaving 675 species to analyse. The analysis was run continuously over a period of 4 months for 220 million generations. To achieve full convergence in MrBayes, it is recommended that the coefficient value of split frequencies of tree searches reaches less than 0.05. However, despite the extraordinary number of generations that we ran in the analysis, the coefficient value failed to go below 0.060544. The tree file from each analysis was summarised using sum-t in MrBayes and a 50% majority-rule consensus tree was generated after omitting a 20% burn-in.

Taxonomy mapping

Eucalypt classifications were traced onto three resulting phylogenies (two ML and one Bayesian) by using parsimony mapping in Mesquite (W. P. Maddison, see https://www.mesquiteproject.org/, accessed 24 February 2017). The classification of Nicolle (2015b) was traced at the subgenus, section and series levels to test the monophyly of these groupings. The eucalypt classifications of Brooker (2000) and Hill's Eucalink (see http://plantnet.rbgsyd.nsw.gov.au/PlantNet/Euc/) were also traced onto the phylogeny, but are not discussed in the present paper because they lack scoring for

Table 1. The taxonomic classification of Nicolle (2015*b*) detailed with the current number of recognised species for each taxonomic group

Eucalypt subgenera are ordered by how closely they are related to each other in the maximum likelihood-1 (ML-1) phylogeny

Group common name and genus	Subgenus	Section	Series	Species number	Common name
Mesicalypt					
Allosyncarpia				1	
Eucalyptopsis				2	
Stockwellia				1	
Newcalypt					
Arillastrum				1	
Eucalypt					
Angophora				12	Apples
	Angophora			12	
		Angophora		12	
			Angophora	10	Rough-barked apples
			Costatitae	2	Smooth-barked apples
Corymbia				90	Bloodwoods
	Blakella			34	
		Abbreviatae		19	Paper-fruited bloodwoods
			Tessellatae	3	
			Scutiformes	16	
		Maculatae		3	Spotted gums
		Naviculares		11	Yellow bloodwoods
		Torellianae		1	Cadaghi
	Corymbia			56	
		Calophyllae		4	
		Corymbia		52	
			Terminalipterae	11	
			Dorsiventrales	16	
			Isobilaterales	22	
			Cymbiformes	1	
			Jacobsianae	1	
			Trachyphloiae	1	
Eucalyptus				717	
	Acerosae			1	Plunkett mallee
	Eudesmia			25	Eudesmids
		Complanatae		12	Tropical eudesmids
			Scutelliformes	1	Bailey's stringybark
			Miniatae	7	Orange-flowered gums
			Similes	2	Tropical yellowjackets
			Tetrodontae	2	Tropical stringybarks
		Limbatae		13	
			Heteropterae	12	
			Ebbanoenses	1	
	Cuboidea			1	Narrow-leaved white mahogany
	Idiogenes			1	Gympie messmate
	Eucalyptus			124	Monocalpyts
		Frutices		19	Monocalypt mallees
			Proximae	1	
			Preissianae	4	
			Diversiformae	6	
			Calcicolae	2	
			Muricatae	3	
			Insulares	1	
			Subereae	2	
		Longistylus		5	
			Pedaria	1	Rate's tingle
			Jacksoniae	1	Red tingle
			Occidentales	2	
			Patentes	1	Blackbutt
		Eucalyptus		99	
			White mahoganies	10	White mahoganies

Table 1.   (*continued*)

Group common name and genus	Subgenus	Section	Series	Species number	Common name
			Pachyphloiae	29	Stringybarks
			Radiatae	15	Peppermints
			Psathyroxyla	10	
			Strictae	13	Mallee ashes
			Olsenianae	3	
			Fraxinales	5	
			Regnantes	2	
			Pauciflorae	3	Snow gums
			Longitudinales	3	Sallees
			Eucalyptus	1	Messmate stringybark
			Pseudostringybarks	2	
			Piperitales	1	
			Planchonianae	1	Needlebark
			Stenostomae	1	
		Primitiva		1	
	Cruciformes			1	Yellow tingle
	Alveolata			1	Tallowwood
	Symphyomyrtus			563	Symphyomyrts
		Bisectae		123	Mallees and mallets
			Halophilae	1	
			Heterostemones	8	
			Angustissimae	5	
			Balladonienses	2	
			Brockwayanae	1	
			Caesiae	1	
			Curviptera	30	
			Decurvae	2	
			Falcatae	18	
			Micrantherae	4	
			Porantherae	21	
			Salmonophloiae	1	Salmon gum
			Squamosae	2	
			Subulatae	27	
		Latoangulatae		22	
			Inclusae	1	Karri
			Connexentes	4	
			Lepidotae-Fimbriatia	4	Grey gums
			Pumilae	1	Pokolbin mallee
			Transversae	12	Mahoganies
		Domesticae		3	
		Equatoria		2	Rainbow gum
		Incognitae		3	
		Exsertaria		45	Red and white gums
			Erythroxylon	22	Eastern red gums
			Exsertae	6	Queensland red gums
			Rostratae	2	River red gums
			Subexsertae	15	Tropical white gums
		Maidenaria		79	Blue gums
			Kitsonianae	1	Gippsland mallee
			Sturgissianae	1	Ettrema mallee
			Acaciiformes	9	
			Argyrophyllae	5	
			Bridgesianae	4	
			Crenulatae	1	Buxton silver gum
			Foveolatae	10	Swamp gums
			Globulares	14	
			Microcarpae	5	Small-fruited white gums
			Neglectae	1	Omeo gum
			Orbiculares	9	
			Saxicola	4	Rock gums
			Semiunicolores	4	Yellow gums

Table 1.   (*continued*)

Group common name and genus	Subgenus	Section	Series	Species number	Common name
			Viminales	11	White gums
		Platysperma		7	Snappy gums
		Racemus		1	Hillgrove gum
		Adnataria		106	Boxes and ironbarks
			Buxeales	15	
			Heterophloiae	9	
			Lucasianae	7	Western boxes
			Melliodorae	7	Boxes
			Rhodoxyla	11	Ironbarks
			Siderophloiae	24	Ironbarks
			Striolatae	17	Tropical boxes
			Subbuxeales	15	Southern boxes
			Submelliodorae	1	
		Sejunctae		1	Sugar gum
		Bolites		1	Tuart
		Dumaria		76	Mallets
			Dissonae	1	
			Furfuraceae	3	
			Merrickianae	3	
			Ovulares	10	
			Rufispermae	34	
			Tetrapterae	16	
			Torquatae	9	
		Glandulosae		94	WA Mallee and gimlets
			Accedentes	7	
			Clinatae	4	
			Contortae	8	Gimlets
			Cornutae	13	Yates
			Dundasianae	1	Dundas blackbutt
			Elongatae	5	
			Erectae	25	
			Kruseanae	1	Bookleaf mallee
			Levispermae	22	Wandoo group
			Loxophlebae	4	
			Obliquae	1	
			Stricklandianae	3	

recently described species (they can be viewed in Mesquite, using the nexus files stored in the CSIRO DAP).

Molecular dating

Despite numerous efforts, the concatenated four-locus dataset was too large to be successfully started in a Bayesian (BEAST) analysis. Therefore, molecular dating using penalised likelihood as implemented in r8s (Sanderson 2003) was performed on the two ML and single Bayes phylogenies generated from the concatenated dataset. The most current eucalypt calibrations are those used and justified by Thornhill *et al*. (2012*a*, 2015), by which calibrations were selected using phylogenetics (Thornhill and Crisp 2012) after assessing extant pollen (Thornhill *et al*. 2012*b*) and fossil pollen (Thornhill and Macphail 2012). Information from the oldest macrofossils of *Eucalyptus* was also incorporated (Gandolfo *et al*. 2011; Hermsen *et al*. 2012). The following two eucalypt fossils can be utilised as calibrations on the basis of their best morphological fit to extant Myrtaceae: a Patagonian *Eucalyptus* macrofossil (Gandolfo *et al*. 2011) from the Early Eocene (51.7–52.1

million years ago) as a crown *Eucalyptus* calibration, and an Australian *Myrtaceidites tenuis* fossil pollen (Thornhill and Macphail 2012), also from the Eocene (45–47 million years ago) as a crown *Angophora* and *Corymbia* calibration. A Paleocene pollen *Myrtaceidites mesonesus* from New Zealand (61.7–65 million years ago) was used in previous studies to calibrate all Myrtaceae except three tribes, namely, Psiloxyloideae, Xanthostemoneae and Lophostemoneae (Thornhill *et al*. 2012*a*; Thornhill *et al*. 2015). We used the same fossil in the present study, meaning that the outgroups *Heteropyxis*, *Kjellbergiodendron* and two *Lophostemon* species were not constrained by a calibration in our dating analyses.

Analysis of diversification rates

Bayesian analyses of macroevolutionary mixtures (BAMM; Rabosky 2014) were performed on the two ML and single Bayes phylogenies to estimate and visualise diversification rates and their changes through time. We did not specify taxon numbers for terminals because the phylogeny was well sampled at the species level. Because BAMM does not run if

Maximum Likelihood 1
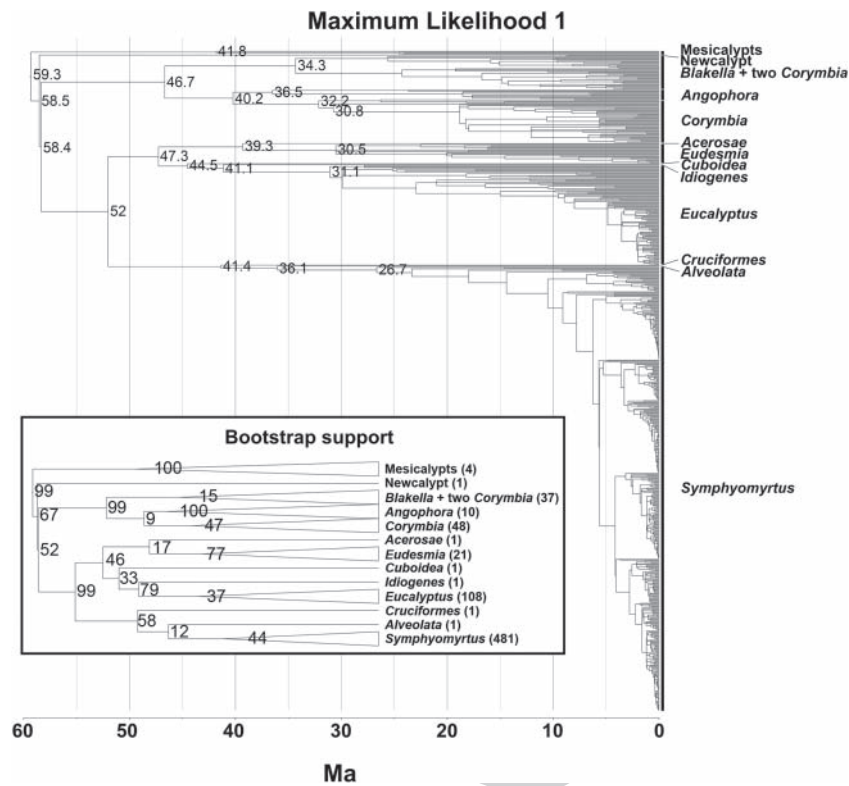
Fig. 2. Maximum likelihood-1 (ML-1) analysis using the concatenated dataset of internal transcribed spacer (ITS), external transcribed spacer (ETS), *mat*K and *psb*A–*trn*H, labelled with the informal higher-level groups mesicalypts (3 genera) and newcalypt (1 genus), the eucalypt genus *Angophora*, and all eucalypt subgenera as classified by Nicolle (2015*b*). Numbers at nodes in the larger phylogeny represent the penalised-likelihood estimated age. The numbers after each name in the inset box represent the number of terminals in the clade and numbers at nodes represent the bootstrap value in the ML analysis. Ma represents millions of years as returned for each penalised-likelihood dating analysis (a summary of estimated ages is provided in Table 2).

polytomies exist in the tree topology, an R script was run to add 0.0001 to every branch length of the phylogeny, which, in turn, randomly converted each phylogeny into a bifurcated tree. Diversification analyses were run using the BAMM stand-alone program, with the speciation and extinction model run for four chains of 5 000 000 generations. The BAMM tools R script was used to remove a burn-in of 10% and calculate the 95% credibility rate shift value for every branch in the tree, with the Bayes-factor criterion for including nodes as core shifts set to 5 (see BAMM instructions for further details). Bayes factors were mapped onto the phylogeny and colour-coded to indicate high values, with the notion that larger values (>2) are an indication of rate shifts.

Results

Tree topology

The 10 ML concatenated-dataset phylogenies all showed slight differences among the relationships of terminal taxa. More significantly, half of the ML trees showed *Corymbia s.l.* to be monophyletic and the other half showed *Corymbia s.l.* to be paraphyletic. Further, there were differences between the ML and Bayesian analyses in the relationships formed between the

species-poor *Eucalyptus* subgenera and the species-rich eudesmid, monocalypt and symphyomyrt subgenera. Given the two possibilities for *Corymbia s.l.*, we selected two ML (ML 1 is the same tree as used by González-Orozco *et al.* 2016), as well as the Bayesian analysis, for use in the taxonomic comparison, dating and diversification analyses (Fig. 2, 3). Interactive versions of the phylogenies with distribution and taxonomies mapped are available at the Phylolink at the Atlas of Living Australia (ML1: https://phylolink.ala.org.au/phylo/show/4827# node/eb2568e49e23ea923b35718982ae4d4f, ML2: https://phylo link.ala.org.au/phylo/show/4830#node/2c634a5f1ca7c22e527 b1e07e4f5631b, and Bayes: https://phylolink.ala.org.au/phylo/ show/4824#node/1ccefcaba46b93c454d42bf7753245dd). Groups outside of Eucalypteae were removed from figures, but the analyses indicated that tribes Leptospermeae and Chamelaucieae were the closest to Eucalypteae, followed by Syncarpieae. Bootstrap support was not high for this topological arrangement, and many more representatives from Leptospermeae and Chamelaucieae and broader Myrtaceae should be included to reach a conclusive result.
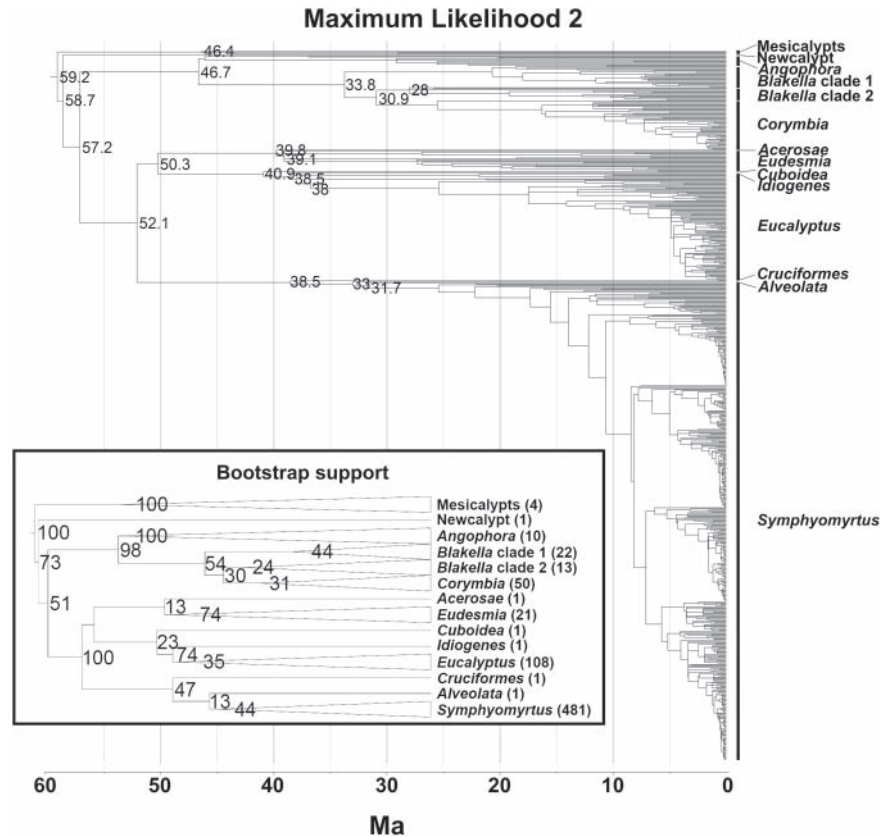
Topology: genera

In all concatenated analyses, a highly supported clade (i.e. greater than 95% bootstrap or posterior-probability support) was

Fig. 3. Maximum likelihood-2 (ML-2) analysis using the concatenated dataset of internal transcribed spacer (ITS), external transcribed spacer (ETS), *mat*K and *psb*A–*trn*H, labelled with the informal higher-level groups mesicalypts (3 genera) and newcalypt (1 genus), the eucalypt genus *Angophora*, and all eucalypt subgenera as classified by Nicolle (2015*b*). Numbers at nodes in the larger phylogeny represent the penalised-likelihood estimated age. Numbers after each name in the inset box represent the number of terminals in the clade and numbers at nodes represent the bootstrap value in the ML analysis. Ma represents millions of years as returned for each penalised-likelihood dating analysis (a summary of estimated ages is provided in Table 2).

recovered for crown mesicalypts (Fig. 2–4, Table 2). Within the mesicalypts, *Stockwellia* and *Eucalyptopsis* were sister, with *Allosyncarpia* being sister to them. The mesicalypts were recovered as sister to the other Eucalypteae genera. Newcalypt (*Arillastrum*) was consistently placed as sister to the eucalypts, although there was only medium statistical support for this placement (ML1 = 67%; ML2 = 73%). The three eucalypt genera formed two clades, namely, *Angophora* + *Corymbia* and a monophyletic *Eucalyptus*, and crowns of these two clades were highly supported.

Topology: subgenera

Half of the ML analyses and the Bayesian analysis placed *Angophora* as a clade within *Corymbia*, whereas the remaining ML analyses placed *Angophora* as sister to *Corymbia*. In all instances, the crown of *Angophora* was given 100% bootstrap or a posterior probability of 1, but the crowns of *Corymbia* subgenera were not as highly supported. *Corymbia* subgenera showed a mixed placement of poorly supported clades (Fig. 2–4, and Fig. S1, available as Supplementary material to this paper). In

ML 1, the *Corymbia* subgenera were not monophyletic because two species of the subgenus *Corymbia* were nested within subgenus *Blakella* (Fig. 2). In ML 2 (Fig. 3), subgenus *Corymbia* formed a clade that nested within subgenus *Blakella* (but had weak support). In the Bayesian analysis, both subgenus *Corymbia* and subgenus *Blakella* were monophyletic. In two instances (ML1 and Bayes, Fig. 2, 4), subgenus *Corymbia* was sister to *Angophora*, but never with high statistical support for the relationship.

The three species-rich *Eucalyptus* subgenera, namely, eudesmids, monocalypts and symphyomyrts, were all monophyletic (Fig. 2–4, and Fig. S1–S3, available as Supplementary material to this paper) and all crowns were moderately supported in ML and highly supported in the Bayesian analysis. The monotypic subgenera *Alveolata* (*E. microcorys*), *Acerosae* (*E. curtisii*), *Cruciformes* (*E. guilfoylei*), *Cuboidea* (*E. tenuipes*) and *Idiogenes* (*E. cloeziana*) consistently occurred as a long-branched sister to one of the three species-rich *Eucalyptus* subgenera, but their placement was not highly supported and moved between the analyses. More broadly, it was consistently found
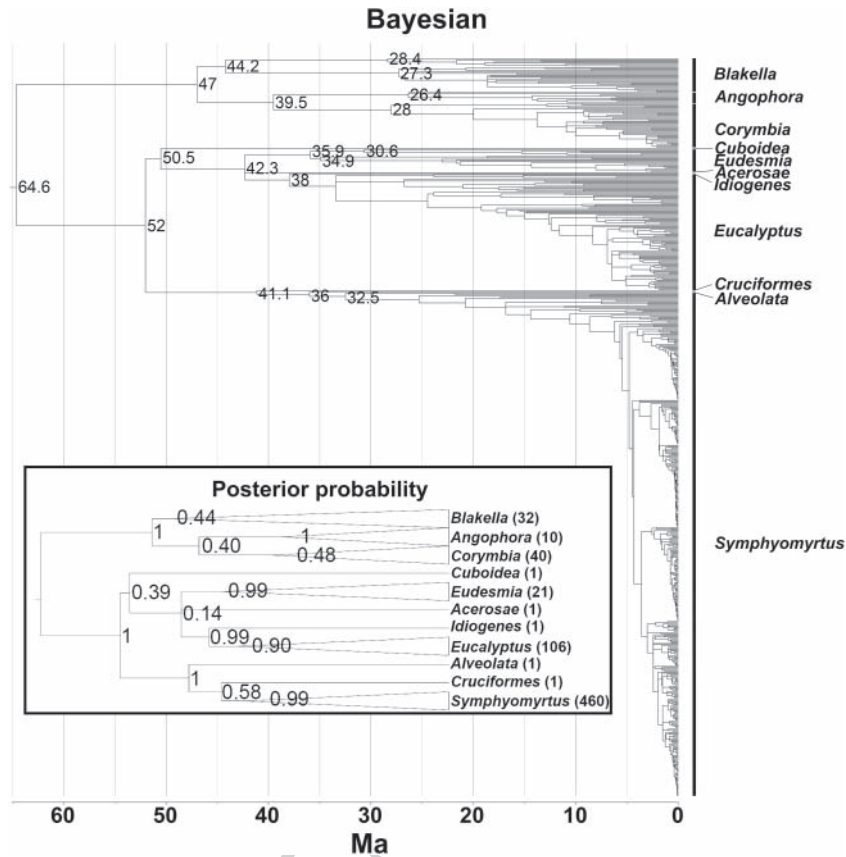
Bayesian

Fig. 4. Bayesian analysis using the concatenated dataset of internal transcribed spacer (ITS), external transcribed spacer (ETS), *mat*K and *psb*A–*trn*H, labelled with the informal higher-level groups mesicalypts (3 genera) and newcalypt (1 genus), the eucalypt genus *Angophora*, and all eucalypt subgenera as classified by Nicolle (2015*b*). Numbers at nodes in the larger phylogeny represent the penalised-likelihood estimated age. Numbers after each name in the inset box represent the number of terminals in the clade and numbers at nodes represent the posterior probability in the Bayesian analysis. Ma represents millions of years as returned for each penalised-likelihood dating analysis (a summary of estimated ages is provided in Table 2).

(although only with moderate support) that subgenera *Acerosae*, *Eudesmia*, *Cuboidea*, *Idiogenes* and *Eucalyptus* composed one large clade that was sister to a clade containing subgenera *Cruciformes*, *Alveolata* and *Symphyomyrtus*.

Topology: sections and series of Eucalyptus

Most sections and series were not recovered as monophyletic (Fig. S2, S3). Sections that were almost monophyletic were *Limbatae* and *Complanatae* in the eudesmids, and *Longistylus*, *Frutices* and *Eucalyptus* in the monocalypts. In the symphyomyrts, sections *Bisectae*, *Dumaria*, *Glandulosae*, *Adnataria*, *Exsertaria* and *Maidenaria* were almost monophyletic. The segregated section *Glandulosae* was widely separated from *Bisectae s.s.* in our phylogenies, supporting the decision to split *Bisectae*.

Some series were clearly resolved as monophyletic, mainly those with low species numbers (e.g. *Calophyllae*, *Miniatae*, *Decurvae* and *Squamosae*). Most of the series were not monophyletic; however, many clades contained a mix of species from two or more series. There were notable outliers in

many series and these might be attributed to any of several methodological issues, including incorrectly identified vouchers, poorly described species or laboratory mistakes. Alternatively, some might be the result of genetic phenomena such as incomplete lineage sorting, hybridisation, pseudogenes, paralogy or polyploidy. Taxa that are suspiciously out of place are marked with an asterisk in Table S1. Because many of these species were sequenced for the first time, we feel that it is best to flag them for future replication. Suspiciously placed taxa all occurred at shallow nodes and had no bearing on the crown age estimates of the deeper nodes.
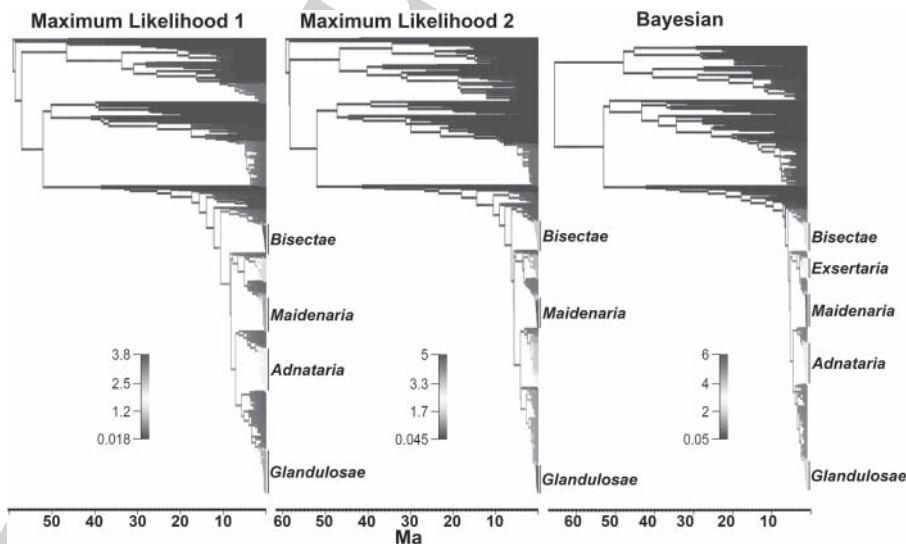
Nuclear v. plastid analyses

Separate ML analyses of nuclear and plastid datasets produced incongruent topologies (Fig. 5). Noticeably different arrangements between the two phylogenies occurred for the monotypic taxa newcalypt (*Arillastrum*), *Acerosae*, *Cuboidea*, *Idiogenes* and *Alveolata*. The nuclear analysis gave higher support for internal nodes, and both analyses returned some nodes with bootstraps values of >75. However, some nodes of

Table 2.   A summary of the estimated ages of select eucalypt groups

Estimated ages are represented by the three penalised-likelihood analyses performed on the two maximum-likelihood (ML) trees and Bayesian consensus tree
(values are listed in that order). NA, not assessed; N, no; Y, yes

| Clade	Monophyletic	Estimated ages (million years) for ML 1, ML 2 or Bayes	Bootstrap or posterior support of clade
Eucalypteae	Y	59.3, 59.2, NA	99, 100, NA
Mesicalypts (*Allosyncarpia, Eucalyptopsis, Stockwellia*)	Y	41.8, 46.4, NA	100, 100, NA
Eucalyptopsis	Y	24.1, 29.1, NA	100, 100, NA
Newcalypt + eucalypts	Y	58.6, 59.2, NA	67, 73, NA
Corymbia + Angophora	Y	46.7, 46.7, 47	99, 98, 1
Crown *Angophora*	Y	36.5, 46.1, 26.4	100, 100, 1
Crown *Corymbia*	Y or N (not monophyletic in all analyses)	NA, 33.8, NA	NA, 54, NA
Crown *Eucalyptus*	Y	52, 52.1, 52	99, 100, 1
Subgenus *Eudesmia* (eudesmids)	Y	30.5, 39.1, 35.9	77, 74, 0.99
Subgenus *Eucalyptus* (monocalypts)	Y	31.1, 38, 33.4	37, 35, 1
Subgenus *Symphyomytus*	Y	26.7, 31.7, 32.5	44, 44, 0.98
Subgenera *Acerosae (E. curtisii)* + Eudesmia + Cuboidea (*E. tenuipes*) + Idiogenes (*E. cloeziana*) + Eucalyptus	Y	47.3, 50.3, 50.5	46, 40, 0.38
Subgenera *Acerosae (E. curtisii)* + Eudesmia (not recovered in Bayes analysis)	Y or N	39.3, 39.7, 42.3	17, 13, NA
Subgenera *Cuboidea (E. tenuipes)* + Idiogenes (*E. cloeziana*) + Eucalyptus	Y	44.53, 40.9, 50.5	33, 23, 0.41
Subgenera *Idiogenes (E. cloeziana)* + Eucalyptus	Y	41.1, 38.5, 38	79, 74, 1
Subgenera *Alveolata (E. microcorys)* + Symphyomytus (not recovered in Bayes analysis)	Y or N	36.1, 33, 41.1	12, 13, NA
Subgenera *Cruciformes (E. guilfoylei)* + Alveolata (*E. microcorys*) + Symphyomytus	Y	41.4, 38.5, 36	58, 47, 1

Fig. 5.   A tanglegram showing the various position of eucalypt groups in the nuclear and plastid analyses.
Numbers at nodes represent the bootstrap value returned in the maximum-likelihood analysis of the two
datasets.

the plastid phylogeny did not have any bootstrap support, namely
the symphyomyrts, *Alveolata, Idiogenes* and *Cuboidea*. This
suggests that the plastid loci used in the present study are
extremely similar, even among deeper-level groups. *Corymbia*
was paraphyletic in both nuclear and plastid analyses.

Molecular dating

Three eucalypt chronograms are shown in Fig. 2–4 and the
estimated age range from the three chronograms of select
groups is summarised in Table 2. The crown age estimates of
both *Eucalyptus* and *Angophora + Corymbia* matched the age of
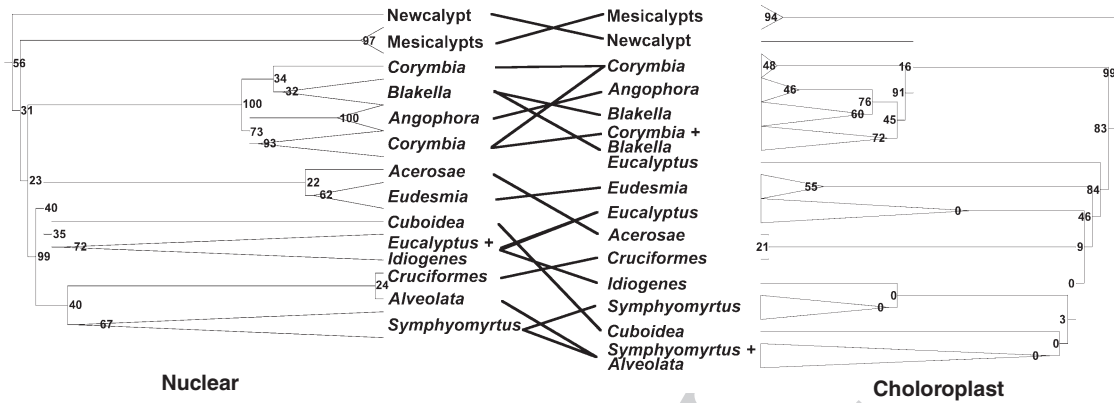
Fig. 6. Bayesian analyses of macroevolutionary mixtures (BAMM) using the maximum likelihood (ML) and Bayesian phylogenies. Significant diversification shifts occurred in *Eucalyptus* sections *Bisectae*, *Exsetaria*, *Maidenaria*, *Adnataria* and *Glandulosae* and they are labelled where significant on each phylogeny. Ma, million years.

the calibration that was applied to them. The estimated crown age of Eucalypteae was Palaeocene (~59 million years). The crown age of the mesicalypts was estimated as Middle Eocene (~41–46 million years), whereas the estimated crown age of *Eucalyptopsis* was Oligocene (~24–29 million years). Because of the very short internodes near the root of the tree, the estimated age of the divergence between newcalypts and the eucalypts was similar to the Eucalypteae crown (~59 million years). The crown age of the combined *Eucalyptus* subgenera *Acerosae*, *Eudesmia*, *Cuboidea*, *Idiogenes* and *Eucalyptus* was estimated as early Eocene (~47–50 million years). The crown age of combined *Eucalyptus* subgenera *Cruciformes*, *Alveolata* and *Symphyomyrtus* was estimated as middle Eocene (~36–41 million years). The crown ages of large subgenera *Eucalyptus* and *Eudesmia* were both estimated as late Eocene (~31–39 million years). The crown age of *Symphyomyrtus* was estimated as early Oligocene (~26–32 million years).

Diversification rates

Bayesian analyses of macroevolutionary mixtures (BAMM) identified five accelerations of net species diversification in *Eucalyptus* sections *Bisectae*, *Exsetaria*, *Maidenaria*, *Adnataria* and *Glandulosae* (Fig. 6). The accelerations all began 2–3 million years ago. The estimated rates of diversification differed among our three phylogenies, but consistently occurred in the same sections.

Discussion

The phylogenies of our study all resolved four main Eucalypteae clades in accordance with most broad eucalypt and Myrtaceae studies (Steane *et al*. 1999; Udovicic and Ladiges 2000; Bayly *et al*. 2013; Thornhill *et al*. 2015). All but the nuclear phylogenetic analysis returned the mesicalypts as sister to the rest of Eucalypteae. Similar to previous Myrtaceae studies (Thornhill *et al*. 2012*b*, 2015), newcalypt (*Arillastrum*) was consistently recovered as sister to the eucalypts, although there was never high statistical support for this placement. Our newcalypt divergence-age estimate was Paleocene (59 million years ago) rather than Eocene (Thornhill *et al*. 2015). Ladiges and Cantrill

(2007) suggested that some New Caledonia plant lineages formed through vicariance and this new age estimate is extremely close to the final vicariance split between Australia and Zealandia between 65 and 60 million years ago (Sdrolias *et al*. 2001). What remains enigmatic, even with this new older date, is how newcalypt(s) persisted on land in the Pacific during the time that New Caledonia was subducted between 60 and 37 million years ago (McLoughlin 2001; Ladiges and Cantrill 2007; Grandcolas *et al*. 2008; Thornhill *et al*. 2015), before arriving to where they are extant.

Classifying Corymbia

Two consistent but moderately supported clade arrangements were resolved within *Angophora* + *Corymbia*. Previous eucalypt analyses have produced alternative placements, the most recent using plastid genomes suggested *Angophora* within *Corymbia* (Bayly *et al*. 2013). If *Angophora* is shown to be undisputedly nested within *Corymbia* (e.g. ML 1 (Fig. 2) and Bayes (Fig. 4) in our study), then a taxonomic change is needed to resolve paraphyly. A simple solution should be sought and we suggest three alternatives (listen below) from least to most complex to implement.

(1) The simple solution is to sink the species of both *Angophora* and *Corymbia* into *Eucalyptus*. Most of the conflict between *Angophora* or *Corymbia*, and *Eucalyptus* binomials (e.g. *Angophora melanoxylon* and *Eucalyptus melanoxylon*) have already been resolved by Brooker (2000). This would make a single Australia-centric genus of just over 800 species. The main hurdle would be the acceptance of *Angophora*, a name that has been in use for over 200 years, being sunk into *Eucalyptus*.

(2) A slightly more complex resolution would be to transfer all *Corymbia* species to *Angophora*. This name change would align the group with the fossil pollen record that has *Eucalyptus* and *Angophora* or *Corymbia* morphotypes (Thornhill and Macphail 2012; Macphail and Thornhill 2016). *Angophora* has fewer species than *Corymbia* (12 *v*. >100), but is the older of the two names (1797 *v*. 1995). By the laws of the International Code of

Nomenclature for algae, fungi and plants, *Angophora* has priority over *Corymbia*. That is, unless conservation of *Corymbia* against *Angophora* could be applied for successfully under Art. 14.12 of the code (McNeill *et al*. 2012). There are no duplicated specific epithets in the two genera.

(3) Another option, if *Angophora* really is nested within *Corymbia*, would be to split *Corymbia* into smaller genera to correct the paraphyly of the latter (Schuster *et al*. 2018). *Corymbia* is classified into subgenera and these could be given genus status. Analyses of nuclear data have suggested that *Corymbia* subgenera are monophyletic (Schuster *et al*. 2018). In contrast, plastid and concatenated analyses have shown that *Corymbia* subgenera may not be monophyletic (Schuster *et al*. 2018 and the present study). If *Corymbia* was to be further split into smaller genera, then justification should be made as to why the *Eucalyptus* subgenera should not also be elevated to genus level. There is obvious morphological variation to warrant the decision (Andrews 1913; Pryor and Johnson 1971; Johnson 1976; Johnson and Briggs 1983), but it has never been accepted by eucalypt specialists.

Eucalypt subgeneric resolution

The phylogenies produced from a concatenated dataset showed that range-restricted monotypic subgenera *Alveolata*, *Acerosae*, *Cruciformes*, *Cuboidea* and *Idiogenes* occur on long branches that are sister to one of the three species-rich subgenera (*Eudesmia*, *Eucalyptus* or *Symphyomyrtus* (Fig. 2–4). They are not an 'idiosyncratic quirk of eucalypt classifiers' (Pryor and Johnson 1981). Further, the subgeneric classification of Nicolle (2015*b*), which made previous subgenera *Minutifructus* and *Primitiva* redundant, is phylogenetically justified because one is polyphyletic and nested in the symphyomyrts (*Minutifructus*), and the other is nested within the monocalypts (*Primitiva*).

Resolution of lower-level classification

The phylogenies of this study showed that most *Corymbia* and *Eucalyptus* sections and series were not monophyletic. Most sections and series represented by more than one species in the phylogeny that were shown to be monophyletic had poor statistical support, no significant branch length, or were nested among members of larger groups. Failure to find fine-scale resolution in the eucalypts can be attributed to several causes, from incomplete lineage sorting, introgression or hybridisation, taxonomic over-splitting, and lack of sufficient data to resolve relationships. Schuster *et al*. (2018) highlighted many of these possible causes for incongruence between taxonomy and phylogenetics in the *Corymbia* and *Angophora* clade and suggested that the nuclear genome may be key to resolving this incongruence, rather than the plastid genome. Our analyses of independent nuclear and plastid datasets (Fig. 5) displayed the same incongruence as in Schuster *et al*. (2018), and there was poor support for most resolved clades of the plastid analysis.

Much more data than the four loci used here will be necessary to achieve higher resolution of the eucalypt relationships. Previous eucalypt studies using a greater proportion of the nuclear or plastid genome than in the present study have still led to non-monophyly of lower-level groups (Nevill *et al*. 2014; Rutherford *et al*. 2015; Jones *et al*. 2016). It may be possible to

resolve relationships in cases of incomplete lineage sorting using a coalescent approach, which is being improved in programs such as STAR BEAST (ver. 2, xxx; Bouckaert *et al*. 2014). The rapidly developing field of phylogenomics also offers a way ahead (reviewed in McKain *et al*. 2018) and the major challenge of identifying orthologous genes is being addressed (Forslund *et al*. 2018), which will help with nuclear genome analyses. Over-splitting of wide-ranging groups may be another factor of non-monophyly in phylogenies, in which taxonomists have segregated small allopatric populations that have developed phenotypic differences, but are still not genetically distinct. We note that Hill's Eucalink (see http://plantnet.rbgsyd.nsw.gov.au/PlantNet/Euc/) classified eucalypt species into informal 'superspecies' and it may be these groups that return monophyletic genetic clades that contain taxa with varying degrees of morphological difference. Alternatively, there is likely to be widespread introgression across the eucalypts as has been inferred in *Corymbia* (Schuster *et al*. 2018). Further, the hybridising ability of what are considered to be more distantly related groups, as also illustrated by Schuster *et al*. (2018), means that a resolved bifurcated tree might not be appropriate, and that nodes closer to the tips of the phylogeny might be better represented as evolutionary networks (e.g. appendix 4 of Rutherford *et al*. 2015).

Eucalypteae age estimates

Estimating ages using penalised likelihood is dependent on having a phylogeny that has internal structure. We made a trade-off in this study by using trees that are well resolved but poorly supported at some nodes. Our discussion relies on the fact that recovered clades and their position are a reasonable reflection of reality; additional nuclear or plastid data will almost surely provide a more certain basis for dating eucalypt evolution and re-visiting the hypotheses that are presented here.

Most of the estimated subgeneric crown ages were younger than Eocene (Table 2). This aligns with both fossil pollen and macrofossil evidence that at least two Australian eucalypt clades existed by the Eocene (Lange 1978; Ambrose *et al*. 1979; Hill *et al*. 2016). Interestingly, the estimated crown ages of tropical mesicalypts (41.8–46.4 million years) were younger than ages of subtropical and temperate *Corymbia* + *Angophora* and *Eucalyptus*. Estimated divergence dates between extra-Australian *Corymbia* (e.g. *C. papuana*) and *Eucalyptus* (e.g. *E. deglupta*) species suggest that the spread of the eucalypts is a recent event and not the result of old Gondwanan lineages persisting in New Guinea or Indonesia. Ladiges *et al*. (2003) also proposed a more recent spread to northern landmasses, when the distance between these areas and Australia shrank in the Pliocene–Pleistocene, greatly enabling the ease of dispersal and rafting. The young age of the extra-Australian eucalypts is in direct contrast to the estimated Oligocene divergence age (24.1–29.1 million years ago) of the mesicalypt genus *Eucalyptopsis* from PNG and its sister genus *Stockwellia* from far-northern Queensland. The species of these two genera are either old lineages, or, the remaining members of a once more widespread group.

The dense and broad sampling of this study allowed crown-age estimation of all major eucalypt subgenera. Results suggest
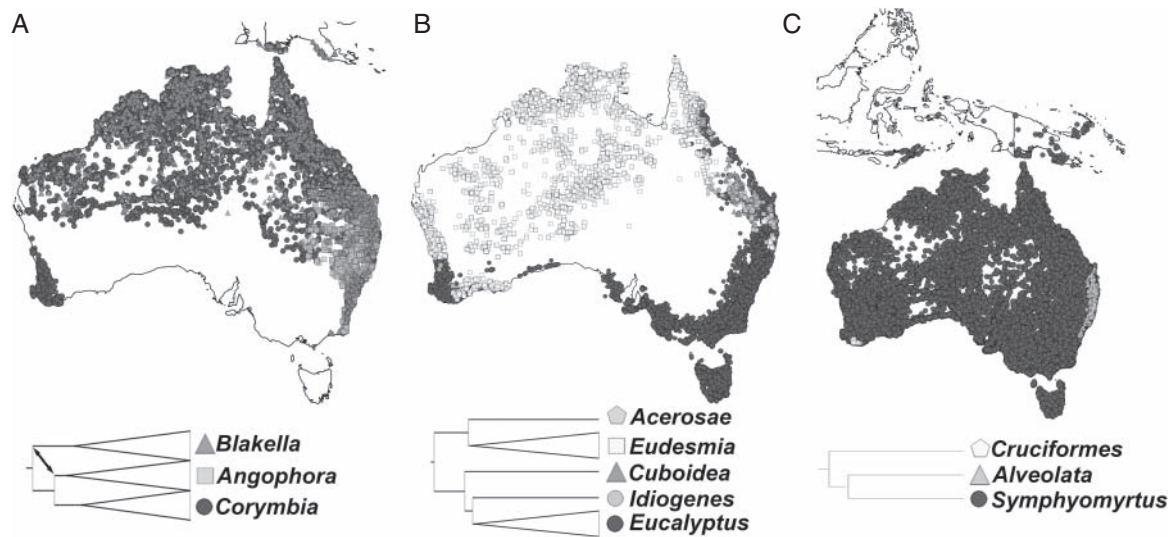
Fig. 7. Distributions of related eucalypt genera and subgenera as returned in the maximum likelihood-1 analysis, showing modern disjunctions. A. *Angophora* is restricted to the eastern coast of Australia. *Blakella* occurs across northern Australia above the Tropic of Capricorn and extends south down the entire eastern coast of Australia, whereas *Corymbia* has the same pattern as *Blakella*, with the addition of a south-western temperate distribution. B. The almost diagonal disjunct split between the clade of *Acerosae* + *Eudesmia* and the clade of *Cuboidea*, *Idiogenes* and *Eucalyptus*. C. The widespread *Symphyomyrtus* is sister to the range-restricted *Cruciformes* in the west and *Alveolata* in the east.

that the crowns of eudesmid, monocalypt and symphyomyrt clades are Late Eocene to Early Oligocene in age. It should be noted that the oldest unequivocal eucalypt macrofossils (Gandolfo *et al*. 2011; Hermsen *et al*. 2012) most closely resemble symphyomyrts, but were used as the calibration of crown *Eucalyptus* in our study. If symplesiomorphy can be disregarded for the Patagonian fossils, and their age had been applied to the symphyomyrt crown, then the age estimate of that subgenus would have increased by ~20 million years.

Modern distributions: Angophora + Corymbia

Angophora + *Corymbia* have a predominantly northern, tropical and monsoonal climatic distribution, with small extensions down the south-western and south-eastern coastal regions of Australia (see Fig. 1B, C, 7A). Hill *et al*. (2016) suggested that a potential environment for the earliest habitat of eucalypts was open vegetation in a monsoonal climate and the extant distribution of *Angophora* + *Corymbia* may be evidence that they have retained the ancestral plesiomorphic preference to such conditions. Fossil pollen evidence from the south of Australia indicates that these two genera occurred there in the Eocene (Thornhill and Macphail 2012; Macphail and Thornhill 2016) and purported macrofossil evidence of *Angophora* in Tasmania (Pryor and Johnson 1981) suggests that *Angophora* and *Corymbia* have become more restricted or dispersed to the north as temperate and arid conditions have arisen in southern Australia.

Within *Corymbia* + *Angophora*, there were some phylogenetic groupings that exhibit disjunctions (Fig. S4, available as Supplementary material to this paper). The red bloodwoods *C. gummifera* + *C. trachyphloia* of temperate eastern coast of Australia and *C. ficicolia* + *C. calophylla* + *C. haematoxylon* of south-western Australia formed sister clades, which was also shown by Ladiges *et al*. (2011). The

placement of *C. trachyphloia* is dubious because it has been placed in other sections in previous studies (Parra-O *et al*. 2006; Schuster *et al*. 2018) and is likely to be a result of the plastid signal in the joint analyses. Age estimates of this divergence ranged from 7.3 to 26.3 million years and, with such wide uncertainty in the divergence-time estimates, any of several events could be invoked to explain the east–west disjunction. Sister clades of *Corymbia* subgenus *Blakella* formed a monsoonal east–west split in their distribution (Fig. S4), which was estimated at between 30.9 and 44.2 million years in age.

Modern distributions: eudesmids + monocalypts

One of the most striking disjunctions is between the eudesmids and the monocalypts (Fig. 7B), but it must be noted that support for the clade containing these two subgenera was low (posterior probability 0.14; bootstrap 46). An almost diagonal linear split from the south-west to the north-east of Australia occurs between the eudesmids and the monocalypts. *Eudesmia* + *Acerosae* occur in an arc that stretches from the south-west of Western Australia to the north-east of Queensland and mainly encompasses arid to monsoonal climatic regions. *Eucalyptus* + *Idiogenes* + *Cuboidea* occur in an opposite arc from the north-east of Australia to south-west of Western Australia that fringes the eastern part of the continent and is associated with cooler, mountainous and temperate climatic conditions. Further, there is almost no overlap between these two groups, yet they cover most of Australia apart from the eastern arid interior. We estimated the divergence age between the eudesmids and monocalypts as early Eocene and Ladiges *et al*. (2010) illustrated how subgenus *Eucalyptus* may have become separated by an Eocene inland water body within Australia. Strikingly, the extant distributions of eudesmids and monocalypts trace this shoreline and now occupy what would have been two main landmasses that existed during this time

(Fig. S5, available as Supplementary material to this paper). The closeness in the age of the estimated divergence between these two subgenera, and of the inland sea, point to the disjunction being the result of vicariance caused by the inland sea separating their ancestor. After the vicariance, the eudesmids radiated into the north-western arc of Australia, and the monocalypts radiated around the south-eastern arc. Bui *et al*. (2017) noted that soil geochemistry was influential in the turnover of monocalypt species around the south-east. Further, the eudesmids and moncalypts did not adapt to grow on the soil created by the inland sea, and only some symphyomyrts adapted and radiated into that part of Australia.

Four phylogenetic groups within the eudesmids, congruent with those found by Gibbs *et al*. (2009), displayed a strong disjunct distribution (Fig. S6, available as Supplementary material to this paper). Age estimates were possible for the crown of three of these four groups, and it was estimated that the western group was 20.1–27.4 million years old, the northern monsoonal 22.5–30.6 million years old, and the north-eastern 18–18.6 million years old. Unfortunately, the backbone relationship of the four clades was not well resolved, appearing as a large polytomy, which is in contrast to the better resolution that Gibbs *et al*. (2009) were able to obtain using a parsimony analysis. Without resolution of how these eudesmid clades are related to each other, it is impossible to infer any dispersal direction or centre of evolution.

There was a strong west–east separation within the monocalypts (Fig. S7, available as Supplementary material to this paper). The western Australian monocalypts were paraphyletic with respect to the eastern monocalypts, and the phylogeny suggests that multiple western Australian monocalpyt lineages had formed before an east–west monocalypt split. Ladiges *et al*. (2010) showed the same result using ITS and ETS sequences. A noticeable difference in the present study was that *E. planchoniana* was recovered in the western monocalypt clade, which confounded an otherwise perfect east–west split in the monocalypt phylogeny. *Eucalyptus planchoniana* has an uncertain history with respect to its taxonomic placement. Brooker (2000) placed it in its own section after a long period of indecision as to whether or not the anthers are reniform with confluent slits (features that are shared by the eastern Australian monocalypts and some of the Western Australian monocalypts, including *E. marginata*; Ladiges *et al*. 2010). Further replicate DNA sampling of *E. planchoniana* would test whether this result is anomalous. The crown of eastern monocalypts was estimated to be between 14.1 and 19.2 million years old, approximately coinciding with the uplift of the Nullarbor. It is likely that the eastern monocalypts have diversified and spread the length of the eastern coast of Australia, including all of Tasmania, within the last 20 million years. This aligns with the eucalypt fossil record that has failed to recover *Eucalyptus* macro- or microfossils in south-eastern Australia any older than Late Oligocene in age (Hill *et al*. 2016; Macphail and Thornhill 2016)

Modern distributions: symphyomyrts + Cruciformes + Alveolata

The widespread and diverse symphyomyrt clade was placed either sister to the monotypic *Cruciformes* that is restricted to south-western Western Australia, or *Alveolata* that is restricted to south-eastern Queensland (Fig. 7C). The ML and Bayesian phylogenies suggested that both *Cruciformes* and *Alveolata* are separate from each other and the symphyomyrts. The Bayesian analysis gave strong support for a clade containing the symphyomyrts, *Cruciformes,* and *Alveolata*, whereas ML produced only moderate to weak support. Our age estimates for the crown of *Cruciformes*, *Alveolata* and the symphyomyrts place it in the Middle to Late Eocene (36–41.4 million years ago). This points to two things; first, the divergence between the two range-restricted monotypic lineages, that currently occur on opposite sides of Australia, is extremely old. The Middle to Late Eocene age again suggests that a vicariant event caused this disjunction. In the *Cruciformes* and *Alveolata* instance, it is likely that there has been extinction to restrict the two lineages to widely isolated areas on the eastern and western coast of Australia, as has previously been suggested by Ladiges *et al*. (2011) and Bayly *et al*. (2013). The second is that the symphyomyrts did not start diversifying until after the Middle to Late Eocene. This age estimate needs further examination. There is more than one morphotype of *Eucalyptus* fossil from the Eocene of South America (Gandolfo *et al*. 2011; Hermsen *et al*. 2012), and perhaps as many as five, and they are all likely to be symphyomyrts. Therefore, the symphyomyrt diversification either occurred after it diverged from the South American eucalypts, or the symphyomyrt crown age is older than the age estimates of the present study.

Similar to the monocalypts, no symphyomyrt clade containing a south-eastern species returned an estimated crown age older than 20 million years, and most crown-age estimates were younger than 10 million years old (Fig. S8, available as Supplementary material to this paper). It was possible to identify some sister clades that displayed a disjunct distribution between each other within the symphyomyrts, albeit with some overlap. The sheer number of taxa in the symphyomyrts, and the recent diversification of some groups makes it difficult to easily identify groups that may be geographically distinct. A better resolved phylogeny, in combination with a detailed biogeographic analysis that scores each taxon for the biome or finer-scale bioregion on the basis of their native distribution would possibly have more success in identifying recently derived disjunct groups.

Diversification

Bayesian analyses of macroevolutionary mixtures (BAMM) suggested that diversification increase has occurred in some *Eucalyptus* lineages during the Plio-Pleistocene (2–3 million years ago). *Eucalyptus* sections that display significant diversification shifts do not all have a common biome, habit or fire strategy. Three sections, namely *Bisectae*, *Dumaria* and *Glandulosae*, occur solely in semi-arid areas of Australia and are mallee and mallet in habit. However, other sections that displayed high diversification rates, namely *Adnataria*, *Maidenaria*, *Exsertaria* and *Latoangulatae*, grow in open woodlands and open forests of eastern Australia. Given that such a varied group of *Eucalyptus* lineages accelerated their diversification during the same period, it suggests that there might be a common driver that affected all lineages simultaneously. It has previously been suggested that climate

change is the obvious candidate for the eucalypt dominance and diversity of eucalypts (Ladiges *et al.* 2003; Crisp *et al.* 2004). Our phylogenetic estimates suggested that diversification occurred rapidly and recently in the Plio-Pleistocene, a time of continent-wide aridification (Hill 1998). The various eucalypt lineages with accelerated diversification were already pre-adapted to the new, changed climate (e.g. drier, more seasonal, and so more conducive to bushfires), which gave them an advantage over competitor trees from other groups (Crisp *et al.* 2011) and, once a change in climate took place, they took full advantage.

The fossil record suggested that eucalypts were not the dominant element of the Australian vegetation until sometime after the commencement of the Miocene (Hill *et al.* 2016). Further, *Eucalyptus* fossil pollen increases in abundance closer to the present, and is associated with an increase in charcoal (Sniderman and Haberle 2012). It has been suggested that, *c.* 1.5 million years ago, an increase in fire frequency, possibly created by changed rainfall, caused the diversification of the fire-tolerant groups such as the eucalypts (Sniderman and Haberle 2012). However, it is important to note the difference between abundance or domination and diversification. Paleo-pollen can be used to infer past vegetation composition, but because of the conservative nature of pollen morphology, records are limited to identifying the dominance of plant groups, rather than historical diversity within groups. Fossil eucalypt pollen can only currently be separated into three main morphotypes (Macphail and Thornhill 2016), and on the basis of modern pollen morphology, these three types can be linked to two large eucalypt groups, namely *Angophora + Corymbia*, and *Eucalyptus* (Thornhill and Macphail 2012). The *Eucalyptus* fossil morphotype (*Myrtaceidites eucalyptoides*) became abundant in records after the Miocene. However, we currently have no way of discerning how many species contributed to the rise in *Eucalyptus* pollen in the fossil record.

It is possible that the pollen record and the diversification estimations are both correct. The eucalypt dominance (inferred from the pollen record) could have begun in the Miocene because a small number of ancestral taxa expanded their range across the continent and colonised newer climatic biomes. As the continent further dried, it began to form smaller pockets of unique areas that then caused a rapid delineation of these widespread ancestral taxa into morphologically distinct species. Ladiges *et al.* (2003) suggested that eucalypt species are breaking into fragmented ranges, which could easily be a driver of diversifying large widespread species into isolated populations that become unique entities. Bui *et al.* (2017) suggested that climate and geochemistry are major factors in controlling the current distribution of eucalypt species. By combining different climate with a mosaic of soil types, it is possible to create a great number of niches into which the ancestral taxa could spread and then diversify.

Quickly adapting to newer geochemical environments could be another driver of diversification. Pryor (1976) noted that *Adnataria* occurs in eastern Australia on inland-slope soils that are younger and more fertile, but not on older land with infertile soils that host other eucalypt groups instead. The diversification estimates of the present study support the hypothesis that *Adnataria* has undergone a recent acceleration. Pryor (1976) also noted that in the west of Australia, sections *Bisectae* and *Dumaria* occur on the inland slopes that consist of newer soil, and both of these clades show recent increased diversification in our study. Bui *et al.* (2017) described the importance of geochemistry (and climate) to the distribution of some eucalypt lineages. However, the study of Bui *et al.* (2017) was only a taxonomic assessment, lacking any phylogenetic or evolutionary timing information. With the addition of the dated phylogenies of the present paper, it is now possible to assess whether there is correlation between the age of eucalypt lineages and the soils that they grow on, and this work is currently in progress (Elisabeth Bui, pers. comm.).

We must note that the BAMM results should come with caveats. Taxonomic artefacts cannot be discounted as the cause of the apparent diversification rate upturn. If a group has been over-split because it is morphologically variable, despite being genetically similar, then a high diversification rate may be inferred for that clade (Rabosky *et al.* 2013; Wiens *et al.* 2015). In our study, high diversification estimates occurred in clades containing many species. However, in support of our rate estimates, section *Eucalyptus*, the eudesmids and subgenus *Corymbia* all contained a number of species comparable to those of the clades with increased diversification rates, but did not display any significant diversification rate shift. The placement of internal calibrations could also influence the inference of higher or lower diversification rates. Our phylogenies showed clustering of nodes closer to the tips of the tree. Only three calibrations could justifiably be used in our study. Without calibrations to control internal nodes, many young ages were returned. Identifying new fossils to calibrate the internal nodes of the eucalypts may overcome some of the younger age estimates returned. Alternatively, they may also confirm the young age estimates if older fossils cannot be identified, or do not actually exist. Unfortunately, fossil discovery is random, but it may be possible to re-evaluate known fossils to see whether they can be applied to nodes (e.g. Lange 1978; Holmes *et al.* 1982).

Future research

Bayly (2016) proposed several ways in which future research on eucalypt systematics could progress. The underlying questions come down to the following two elements: how are the eucalypts related to each other and how old are they? The first question could be resolved with phylogenetics and then taxonomic relationships updated on the basis of a well resolved and well supported tree; solving what to do with *Corymbia* and *Angophora* should be one clear goal. The relationships among lower-level taxa may forever remain muddied. We have shown that many of the sections are almost monophyletic; this may be because of one of many reasons, and not all of them are genetic or taxonomic. Further replicate sampling of eucalypt species will test the validity of these non-monophyletic sections, and we hope that the present study has made progress towards this goal.

The fossil record shows that the eucalypts are old. The two recent reviews on the eucalypt micro- and macrofossil record summarised current knowledge of the group and how it can be improved (Hill *et al.* 2016; Macphail and Thornhill 2016). High priority should be given to finding macrofossils that can correspond with some of the major eucalypt clades that have easily identifiable apomorphic features of distinctive clades,

such as *Angophora* or the eudesmids. Identifying the modern-day relatives of eucalypt macrofossils from other regions such as Antarctica and New Zealand will also help in determining the deeper evolutionary history of the group. Unfortunately, we are somewhat reliant on serendipitous finds; however, as the Patagonian eucalypt fossils illustrate, when they are found they significantly change our perspective.

In addition to Bayly's two elements of importance, we add a third, namely, when and how often did eucalypts diversify around Australia? Our results suggest that some lineages of eucalypts have diversified very recently. It may be possible to test this by using the fossil record. Finding distinct pollen morphotypes that show phylogenetic signals within the eucalypts, if possible, would be a significant advancement. Although Thornhill *et al.* (2012*b*) surveyed the pollen of several extant eucalypt species, the study was not purely focused on discovering unique pollen types within the eucalypts. With the availability of a well sampled phylogeny, we are better placed to search for unique extant eucalypt pollen types that could be associated with fossil pollen. If such morphotypes could be identified, it would pave the way to revisit past paleo-vegetation studies to not only investigate rise in dominance, but also the diversification of the eucalypts through time. Likewise, the macrofossil record of eucalypts could be re-assessed to categorise how many distinct groups were present in each epoch and whether there are distinct morphological characters in the fossils that could link them to phylogenetic clades for molecular dating. A properly coded and detailed biogeographic analysis of the eucalypts, incorporating well resolved and substantiated phylogenies that allows for more robust dating estimates, and newer methods such as BioGeoBears (Matzke 2013) could assist with interpreting the eucalypt spread.

Conclusions

Eucalypteae is an old lineage with a mix of genera that are depauperate and diverse. Depauperate genera are likely to have suffered extinction and retreated to environments more similar to their ancestral area. All extant south-eastern Australian species of *Eucalyptus* have estimated ages that are younger than 20 million years, supporting fossil evidence that the dominant vegetation component of Australian forests has been this way only since the Late Oligocene to Early Miocene. Further, diverse clades such as some of those in *Eucalyptus s.l.* have adapted to a drying environment and have both radiated and diversified over the landscape of Australia in the recent past, evidenced by significant upward shifts in diversification rates in the past two million years in five sections of *Eucalyptus*.

It is becoming more apparent that a name change is needed for at least part of *Corymbia*; despite the phylogenies in the present study broadly sampling almost all Eucalypteae species, we could not fully resolve the relationships of the major subgeneric groups. We have outlined three options for changes that could be made to the taxonomy, so as to reflect the phylogeny in the eucalypts but encourage that any change should consider the fossil record, especially pollen. The next research steps involve sampling as many described eucalypts as possible by using high-throughput sequencing methods. By sequencing substantially more loci, it is expected that significant progress can be made on how the major

eucalypt groups are related and how best to deal with any taxonomic issues that may be identified by a better-resolved phylogeny. The effect of hybridisation and introgression in the evolutionary history of the eucalypts may make this a challenging task.

References

Ambrose GJ, Callen R, Flint RB, Lange RT (1979) *Eucalyptus* fruits in stratigraphic context in Australia. *Nature* **280**, 387–389. doi:10.1038/280387a0

Andrews EC (1913) The development of the natural order Myrtaceae. *Proceedings of the Linnean Society of New South Wales* **38**, 529–568. doi:10.5962/bhl.part.13567

Bayly MJ (2016) Phylogenetic studies of eucalypts: fossils, morphology, and genomes. *Proceedings of the Royal Society of Victoria* **128**, 12–24. doi:10.1071/RS16002

Bayly MJ, Rigault P, Spokevicius A, Ladiges PY, Ades PK, Anderson C, Bossinger G, Merchant A, Udovicic F, Woodrow IE, Tibbits J (2013) Chloroplast genome analysis of Australian eucalypts – *Eucalyptus, Corymbia, Angophora, Allosyncarpia* and *Stockwellia* (Myrtaceae). *Molecular Phylogenetics and Evolution* **69**, 704–716. doi:10.1016/j.ympev.2013.07.006

Berger BA, Kriebel R, Spalink D, Sytsma KJ (2016) Divergence times, historical biogeography, and shifts in speciation rates of Myrtales. *Molecular Phylogenetics and Evolution* **95**, 116–136. doi:10.1016/j.ympev.2015.10.001

Biffin E, Craven LA, Crisp MD, Gadek PA (2006) Molecular systematics of *Syzygium* and allied genera (Myrtaceae): evidence from the chloroplast genome. *Taxon* **55**, 79–94. doi:10.2307/25065530

Bouckaert R, Heled J, Kühnert D, Vaughan T, Wu C-H, Xie D, Suchard MA, Rambaut A, Drummond AJ (2014) BEAST 2: a software platform for Bayesian evolutionary analysis. *PLoS Computational Biology* **10**, e1003537doi:10.1371/journal.pcbi.1003537

Brooker MIH (2000) A new classification of the genus *Eucalyptus* L'Hér. (Myrtaceae). *Australian Systematic Botany* **13**, 79–148. doi:10.1071/SB98008

Bui EN, Thornhill AH, González-Orozco CE, Knerr NJ, Miller JT (2017) Climate and geochemistry as drivers of eucalypt diversification in Australia. *Geobiology* **15**, 427–440. doi:10.1111/gbi.12235

Cavanilles AJ (1797) *Icones et Descriptiones Plantarum* **4**(1), 21[tab. 338]

Crisp MD, Cook LG, Steane DA (2004) Radiation of the Australian flora: what can comparisons of molecular phylogenies across multiple taxa tell us about the evolution of diversity in present-day communities? *Philosophical Transactions of the Royal Society of London – B. Biological Sciences* **359**, 1551–1571. doi:10.1098/rstb.2004.1528

Crisp MD, Burrows GE, Cook LG, Thornhill AH, Bowman DMJS (2011) Flammable biomes dominated by eucalypts originated at the Cretaceous–Palaeogene boundary. *Nature Communications* **2**, 193 doi:10.1038/ncomms1191

Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research* **32**, 1792–1797. doi:10.1093/nar/gkh340

Forslund K, Pereira C, Capella-Gutierrez S, Sousa da Silva A, Altenhoff A, Huerta-Cepas J, Muffato M, Patricio M, Vandepoele K, Ebersberger I, Blake J, Fernández Breis J, Boeckmann B, Gabaldón T, Sonnhammer E, Dessimoz C, Lewis S (2018) Gearing up to handle the mosaic nature of life in the quest for orthologs. *Bioinformatics* **34**, 323–329. doi:10.1093/bioinformatics/btx542

Gadek PA, Wilson PG, Quinn CJ (1996) Phylogenetic reconstruction in Myrtaceae using *mat*K, with particular reference to the position of *Psiloxylon* and *Heteropyxis*. *Australian Systematic Botany* **9**, 283–290. doi:10.1071/SB9960283

Gandolfo MA, Hermsen EJ, Zamaloa MC, Nixon KC, González CC, Wilf P, Cúneo NR, Johnson KR (2011) Oldest known *Eucalyptus* macrofossils are from South America. *PLoS One* **6**, e21084 doi:10.1371/journal.pone.0021084

Gibbs AK, Udovicic F, Drinnan AN, Ladiges PY (2009) Phylogeny and classification of *Eucalyptus* subgenus *Eudesmia* (Myrtaceae) based on nuclear ribosomal DNA, chloroplast DNA and morphology. *Australian Systematic Botany* **22**, 158–179. doi:10.1071/SB08043

González-Orozco CE, Pollock LJ, Thornhill AH, Mishler BD, Knerr NJ, Laffan SW, Miller JT, Rosauer DF, Faith DP, Nipperess DA, Kujala H, Linke S, Butt N, Külheim C, Crisp MD, Gruber B (2016) Phylogenetic approaches reveal biodiversity threats under climate change. *Nature Climate Change* **6**, 1110–1114. doi:10.1038/nclimate3126

Grandcolas P, Murienne J, Robillard T, Desutter-Grandcolas L, Jourdan H, Guilbert E, Deharveng L (2008) New Caledonia: a very old Darwinian island? *Philosophical Transactions of the Royal Society of London – B. Biological Sciences* **363**, 3309–3317. doi:10.1098/rstb.2008.0122

Hermsen EJ, Gandolfo MA, Zamaloa MDC (2012) The fossil record of *Eucalyptus* in Patagonia. *American Journal of Botany* **99**, 1356–1374. doi:10.3732/ajb.1200025

Hill RS (1998) Poor soils and a dry climate: the evolution of the Australian scleromorphic and xeromorphic vegetation. *Australian Biologist* **11**, 26–29.

Hill KD, Johnson LAS (1991*a*) Systematic studies in the eucalypts. 4. New taxa in *Eucalyptus* (Myrtaceae). *Telopea* **4**, 321–349. doi:10.7751/telopea19914932

Hill KD, Johnson LAS (1991*b*) Systematic studies in the eucalypts. 3. New taxa and combinations in *Eucalyptus* (Myrtaceae). *Telopea* **4**, 223–267. doi:10.7751/telopea19914928

Hill KD, Johnson LAS (1992) Systematic studies in the eucalypts. 5. New taxa and combinations in *Eucalyptus* (Myrtaceae) in Western Australia. *Telopea* **4**, 561–634. doi:10.7751/telopea19814948

Hill KD, Johnson LAS (1994) Systematic studies in the eucalypts. 6. A revision of the coolibahs, *Eucalyptus* subgenus *Symphyomyrtus* section Adnataria series Oligantheae subseries Microthecosae (Myrtaceae) K.D. *Telopea* **5**, 743–771. doi:10.7751/telopea19943000

Hill KD, Johnson LAS (1995) Systematic studies in the eucalypts 7. A revision of the bloodwoods, genus *Corymbia* (Myrtaceae). *Telopea* **6**, 185–504. doi:10.7751/telopea19953017

Hill KD, Johnson LAS (1998) Systematic studies in the eucalypts. 8. A review of the eudesmioid eucalypts, *Eucalyptus* subgenus *Eudesmia* (Myrtaceae). *Telopea* **7**, 375–414. doi:10.7751/telopea19982006

Hill KD, Johnson LAS (2000) Systematic studies in the eucalypts. 10. New tropical and subtropical eucalypts from Australia and New Guinea (*Eucalyptus*, Myrtaceae). *Telopea* **8**, 503–539. doi:10.7751/telopea20002007

Hill RS, Truswell EM, McLoughlin S, Dettman ME (1999) The evolution of the Australian flora: fossil evidence. 'Flora of Australia Volume 1 – Introduction'. (Ed. AE Orchard) pp. 251–320. (CSIRO Publishing: Melbourne, Vic., Australia)

Hill KD, Johnson LAS, Blaxell DF (2001) Systematic studies in the eucalypts. 11. New taxa and combinations in *Eucalyptus* section Dumaria (Myrtaceae). *Telopea* **9**, 259–318. doi:10.7751/telopea20013003

Hill RS, Beer YK, Hill KE, Maciunas EC, Tarran MA, Wainman CC (2016) Evolution of the eucalypts: an interpretation from the macrofossil record. *Australian Journal of Botany* **64**, 600–608. doi:10.1071/BT16117

Holmes WBK, Holmes FM, Martin HA (1982) Fossil Eucalyptus remains from the Middle Miocene Chalk Mountain Formation, Warrumbungle Mountains, New South Wales. *Proceedings of the Linnean Society of New South Wales* **106**, 299–310.

Huelsenbeck JP, Ronquist FR (2001) MrBayes: Bayesian inference of phylogenetic trees. *Bioinformatics* **17**, 754–755.

Johnson LAS (1972) Evolution and classification in *Eucalyptus*. *Proceedings of the Linnean Society of New South Wales* **97**, 11–29.

Johnson LAS (1976) Problems of species and genera in *Eucalyptus* (Myrtaceae). *Plant Systematics and Evolution* **125**, 155–167. doi:10.1007/BF00986148

Johnson LAS, Briggs BG (1983) Myrtaceae. Comments on comments. *Taxon* **32**, 103–105. doi:10.2307/1219859

Johnson LAS, Hill KD (1990) New taxa and combinations in *Eucalyptus* and *Angophora* (Myrtaceae). *Telopea* **4**, 37–108. doi:10.7751/telopea19904916

Johnson LAS, Hill KD (1991) Systematic studies in the eucalypts. 2. A revision of the gimlets and related species: *Eucalyptus* extracodical series Salubres and Annulatae (Myrtaceae). *Telopea* **4**, 201–222. doi:10.7751/telopea19914927

Johnson LAS, Hill KD (1999) Systematic studies in the eucalypts. 9. A review of series Sociales (*Eucalyptus* subgenus *Symphyomyrtus*, section Bisectaria, Myrtaceae). *Telopea* **8**, 165–218. doi:10.7751/telopea19993001

Jones RC, Nicolle D, Steane DA, Vaillancourt RE, Potts BM (2016) High density, genome-wide markers and intra-specific replication yield an unprecedented phylogenetic reconstruction of a globally significant, speciose lineage of *Eucalyptus*. *Molecular Phylogenetics and Evolution* **105**, 63–85. doi:10.1016/j.ympev.2016.08.009

L'Héritier de Brutelle CL (1789) 'Sertum Anglicum.' (Didot: Paris, France)

Ladiges PY (1997) Phylogenetic history and classification of eucalypts. In 'Eucalypt Ecology: Individuals to Ecosystems'. (Ed. WJRJ Woniarski) pp. 16–29. (Cambridge University Press: Cambridge, MA, USA)

Ladiges PY, Cantrill D (2007) New Caledonia–Australian connections: biogeographic patterns and geology. *Australian Systematic Botany* **20**, 383–389. doi:10.1071/SB07018

Ladiges P, Udovicic F (2000) Comment on a new classification of the eucalypts. *Australian Systematic Botany* **13**, 149–152. http://www.publish.csiro.au/?paper=SB99011 doi:10.1071/SB99011

Ladiges PY, Udovicic F, Drinnan AN (1995) Eucalypt phylogeny: molecules and morphology. *Australian Systematic Botany* **8**, 483–497. doi:10.1071/SB9950483

Ladiges PY, Udovicic F, Nelson G (2003) Australian biogeographical connections and the phylogeny of large genera in the plant family Myrtaceae. *Journal of Biogeography* **30**, 989–998. doi:10.1046/j.1365-2699.2003.00881.x

Ladiges PY, Bayly MJ, Nelson GJ (2010). East–west continental vicariance in *Eucalyptus* subgenus *Eucalyptus*. In 'Beyond Cladistics: the Branching of a Paradigm The Branching of a Paradigm'. (Eds DM

Williams, S Knapp) pp. 267–302. (University of California Press: Berkeley, CA, USA) doi: doi:10.1525/california/9780520267725.003.0014

Ladiges P, Parra-O C, Gibbs A, Udovicic F, Nelson G, Bayly M (2011) Historical biogeographical patterns in continental Australia: congruence among areas of endemism of two major clades of eucalypts. *Cladistics* **27**, 29–41. doi:10.1111/j.1096-0031.2010.00315.x

Lange RT (1978) Carpological evidence for fossil *Eucalyptus* and other Leptospermeae (subfamily Leptospermoideae of Myrtaceae) from a tertiary deposit in the South Australian arid zone. *Australian Journal of Botany* **26**, 221–233. doi:10.1071/BT9780221

Lucas EJ, Belsham SR, Lughadha EMN, Orlovich DA, Sakuragui CM, Chase MW, Wilson PG (2005) Phylogenetic patterns in the fleshy-fruited Myrtaceae: preliminary molecular evidence. *Plant Systematics and Evolution* **251**, 35–51. doi:10.1007/s00606-004-0164-9

Lucas EJ, Harris SA, Mazine FF, Bellsham SR, Lughadha EMN, Telford A, Gasson PE, Chase MW (2007) Suprageneric phylogenetics of Myrteae, the generically richest tribe in Myrtaceae (Myrtales). *Taxon* **56**, 1105–1128. doi:10.2307/25065906

Macphail MK, Thornhill AH (2016) How old are the eucalypts? A review of the microfossil and phylogenetic evidence. *Australian Journal of Botany* **64**, 579–599. doi:10.1071/BT16124

Macphail MK, Truswell E (1989) Palynostratigraphy of the central west Murray Basin. *BMR Journal of Australian Geology and Geophysics* **11**, 301–331.

Martin H (1991) Tertiary stratigraphic palynology and palaeoclimate of the inland river systems in New South Wales. In 'The Cainozoic in Australia: a Re-appraisal Evidence'. (Eds M Williams, P De Dekker, A Kershaw) pp. 181–194. (Geological Society of Australia: Sydney, NSW, Australia)

Matzke NJ (2013) Probabilistic historical biogeography: new models for founder-event speciation, imperfect detection, and fossils allow improved accuracy and model-testing. *Frontiers of Biogeography* **4**, 242–247.

McKain MR, Johnson MG, Uribe-Convers S, Eaton D, Yang Y (2018) Practical considerations for plant phylogenomics. *Applications in Plant Sciences* **6**, e1038doi:10.1002/aps3.1038

McLoughlin S (2001) The breakup history of Gondwana and its impact on pre-Cenozoic floristic provincialism. *Australian Journal of Botany* **49**, 271–300. doi:10.1071/BT00023

McNeill J, Barrie FR, Buck WR, Demoulin V, Greuter W, Hawksworth DL, Herendeen PS, Knapp S, Marhold K, Prado J, Prud'homme van Reine WF, Smith GF, Wiersema JH, Turland NJ (2012) 'International Code of Nomenclature for Algae, Fungi, and Plants (Melbourne Code)', adopted by the Eighteenth International Botanical Congress, July 2011, Melbourne, Australia. *Regnum Vegetabile*, vol. 154. (Koeltz Scientific Books: Königstein, Germany)

Nevill PG, Després T, Bayly MJ, Bossinger G, Ades PK (2014) Shared phylogeographic patterns and widespread chloroplast haplotype sharing in *Eucalyptus* species with different ecological tolerances. *Tree Genetics & Genomes* **10**, 1079–1092. doi:10.1007/s11295-014-0744-y

Nicolle D (2015a) Classification of the eucalypts (*Angophora, Corymbia* and *Eucalyptus*) Version 1. Available at http://www.dn.com.au/Classification-Of-The-Eucalypts-V1-Feb2015.pdf [Verified 24 February 2017]

Nicolle D (2015b) Classification of the eucalypts (*Angophora, Corymbia* and *Eucalyptus*) Version 2. Available at http://www.dn.com.au/Classification-Of-The-Eucalypts-V2-Apr2015.pdf [Verified 24 February 2017]

Nicolle D (2018) Classification of the eucalypts (*Angophora, Corymbia* and *Eucalyptus*) Version 3. Available at http://www.dn.com.au/Classification-Of-The-Eucalypts.pdf [Verified 14 September 2018]

Ochieng JW, Henry RJ, Baverstock PR, Steane DA, Shepherd M (2007) Nuclear ribosomal pseudogenes resolve a corroborated monophyly of the eucalypt genus *Corymbia* despite misleading hypotheses at functional ITS paralogs. *Molecular Phylogenetics and Evolution* **44**, 752–764. doi:10.1016/j.ympev.2007.04.017

Parra-O C, Bayly MJ, Udovicic F, Ladiges PY (2006) ETS sequences support the monophyly of the eucalypt genus *Corymbia* (Myrtaceae). *Taxon* **55**, 653–663. doi:10.2307/25065641

Parra-O C, Bayly MJ, Drinnan A, Udovicic F, Ladiges P (2009) Phylogeny, major clades and infrageneric classification of *Corymbia* (Myrtaceae), based on nuclear ribosomal DNA and morphology. *Australian Systematic Botany* **22**, 384–399. doi:10.1071/SB09028

Partridge AD (1999) Late Cretaceous to Tertiary geological evolution of the Gippsland Basin, Victoria. PhD thesis, Latrobe University, Bundoora, Melbourne, Vic., Australia.

Potts BM, Pederick LA (2000) Morphology, phylogeny, origin, distribution and genetic diversity of eucalypts. In 'Diseases and Pathogens of Eucalypts'. (Eds PJ Keane, GA Kile, FD Podger, BN Brown) pp. 11–34. (CSIRO Publishing: Melbourne, Vic., Australia)

Pryor LD (1976) Geographic distribution of *Eucalyptus*. In 'The Biology of Eucalypts'. pp. 1–4. (Edward Arnold: London, UK)

Pryor LD, Johnson LAS (1971) 'A classification of the eucalypts.' (Australian National University: Canberra, ACT, Australia)

Pryor LD, Johnson LAS (1981) *Eucalyptus*, the universal Australian. In 'Ecological Biogeography of Australia.' (Ed. A Keast) pp. 501–536. (Dr W Junk Publishers: The Hague, Netherlands)

Rabosky DL (2014) Automatic detection of key innovations, rate shifts, and diversity-dependence on phylogenetic trees. *PLoS One* **9**, e89543 doi:10.1371/journal.pone.0089543

Rabosky DL, Santini F, Eastman J, Smith SA, Sidlauskas B, Chang J, Alfaro ME (2013) Rates of speciation and morphological evolution are correlated across the largest vertebrate radiation. *Nature Communications* **4**, 1958 doi:10.1038/ncomms2958

Rozefelds AC (1996) Eucalyptus phylogeny and history: a brief summary. *Tasforests* **8**, 15–26.

Rutherford S, Wilson PG, Rossetto M, Bonser SP (2015) Phylogenomics of the green ash eucalypts (Myrtaceae): a tale of reticulate evolution and misidentification. *Australian Systematic Botany* **28**, 326–354. doi:10.1071/SB15038

Sanderson MJ (2003) r8s: inferring absolute rates of molecular evolution and divergence times in the absence of a molecular clock. *Bioinformatics* **19**, 301–302. doi:10.1093/bioinformatics/19.2.301

Schuster TM, Setaro SD, Tibbits JFG, Batty EL, Fowler RM, McLay TGB, Wilcox S, Ades PK, Bayly MJ (2018) Chloroplast variation is incongruent with classification of the Australian bloodwood eucalypts (genus *Corymbia*, family Myrtaceae). *PLoS One* **13**, e0195034 doi:10.1371/journal.pone.0195034

Sdrolias M, Muller R, Gaina C (2001) Plate tectonic evolution of eastern Australian marginal ocean basins. In 'Eastern Australasian Basins Symposium', 25–28 November 2001, Melbourne, Vic., Australia (Eds K Hill, T Bernecker) pp. 227–237. (Petroleum Exploration Society of Australia: Melbourne, Vic., Australia)

Slee AV, Brooker MIH, Connors JR, Duffy SM (2006) EUCLID eucalypts of southern Australia (third edition) (on CD-ROM) (CSIRO Publishing: Melbourne, Vic., Australia)

Sniderman JMK, Haberle SG (2012) Fire and vegetation change during the Early Pleistocene in southeastern Australia. *Journal of Quaternary Science* **27**, 307–317. doi:10.1002/jqs.1547

Stamatakis A (2014) RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30**(9), 1312–1313. doi:10.1093/bioinformatics/btu033

Steane DA, McKinnon GE, Vaillancourt E, Potts BM, Vaillancourt RE, Potts BM (1999) ITS sequence data resolve higher level relationships among the eucalypts. *Molecular Phylogenetics and Evolution* **12**, 215–223. doi:10.1006/mpev.1999.0612

Steane DA, Nicolle D, McKinnon GE, Vaillancourt RE, Potts BM (2002) Higher-level relationships among the eucalypts are resolved by ITS-sequence data. *Australian Systematic Botany* **15**, 49–62. doi:10.1071/SB00039

Sun Y, Skinner DZ, Liang GH, Hulbert SH (1994) Phylogenetic analysis of *Sorghum* and related taxa using internal transcribed spacers of nuclear ribosomal DNA. *Theoretical and Applied Genetics* **89**, 26–32. doi:10.1007/BF00226978

Thornhill AH, Crisp MD (2012) Phylogenetic assessment of pollen characters in Myrtaceae. *Australian Systematic Botany* **25**, 171–187. doi:10.1071/SB11019

Thornhill AH, Macphail MK (2012) Fossil myrtaceous pollen as evidence for the evolutionary history of the Myrtaceae: a review of fossil *Myrtaceidites* species. *Review of Palaeobotany and Palynology* **176–177**, 1–23. doi:10.1016/j.revpalbo.2012.03.003

Thornhill AH, Popple LW, Carter RJ, Ho SYW, Crisp MD (2012*a*) Are pollen fossils useful for calibrating relaxed molecular clock dating of phylogenies? A comparative study using Myrtaceae. *Molecular Phylogenetics and Evolution* **63**, 15–27. doi:10.1016/j.ympev.2011.12.003

Thornhill AH, Hope GS, Craven LA, Crisp MD (2012*b*) Pollen morphology of the Myrtaceae. Part 1: tribes Eucalypteae, Lophostemoneae, Syncarpieae, Xanthostemoneae and subfamily Psiloxyloideae. *Australian Journal of Botany* **60**, 165–199. doi:10.1071/BT11174

Thornhill AH, Ho SYW, Külheim C, Crisp MD (2015) Interpreting the modern distribution of Myrtaceae using a dated molecular phylogeny. *Molecular Phylogenetics and Evolution* **93**, 29–43. doi:10.1016/j.ympev.2015.07.007

Truswell E (1993) Vegetation changes in the Australian Tertiary in response to climatic and phytogeographic forcing factors. *Australian Systematic Botany* **6**, 533–558. doi:10.1071/SB9930533

Udovicic F, Ladiges PY (2000) Informativeness of nuclear and chloroplast DNA regions and the phylogeny of the eucalypts and related genera (Myrtaceae). *Kew Bulletin* **55**, 633–645. doi:10.2307/4118780

Udovicic F, McFadden GI, Ladiges PY (1995) Phylogeny of *Eucalyptus* and *Angophora* based on 5S rDNA spacer sequence data. *Molecular Phylogenetics and Evolution* **4**, 247–256. doi:10.1006/mpev.1995.1023

Vaidya G, Lohman DJ, Meier R (2011) SequenceMatrix: concatenation software for the fast assembly of multi-gene datasets with character set and codon information. *Cladistics* **27**, 171–180. doi:10.1111/j.1096-0031.2010.00329.x

Wardell-Johnson GW, Williams JE, Hill KD, Cumming R (1997) Evolutionary biogeography and contemporary distribution of eucalypts. In 'Eucalypt Ecology: Individuals to Ecosystems'. (Eds J Williams, J Woinarski) pp. 92–128. (Cambridge University Press: Cambridge, UK)

White TJ, Bruns TD, Lee S, Taylor J (1990) Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. In 'PCR Protocols: a Guide to Methods and Applications'. (Eds MA Innis, DH Gelfand, JJ Sninsky, TJ White) pp. 315–322. (Academic Press: San Diego, CA, USA)

Whittock S, Steane DA, Vaillancourt RE, Potts BMBM (2003) Molecular evidence shows that the tropical boxes (*Eucalyptus* subgenus *Minutifructus*) are over-ranked. *Transactions of the Royal Society of South Australia* **127**, 27–32.

Wiens JJ, Lapoint RT, Whiteman NK (2015) Herbivory increases diversification across insect clades. *Nature Communications* **6**, doi:10.1038/ncomms9370

Wilson PG (2011) Myrtaceae. In 'The Families and Genera of Vascular Plants. Volume X. Flowering Plants Eudicots: Sapindales, Cucurbitales, Myrtaceae'. (Ed K Kubitzki) pp. 212–271. (Springer-Verlag: Heidelberg, Germany)

Wilson PG, O'Brien MM, Heslewood MM, Quinn CJ (2005) Relationships within Myrtaceae *sensu lato* based on a *mat*K phylogeny. *Plant Systematics and Evolution* **251**, 3–19. doi:10.1007/s00606-004-0162-y

Wright SD, Yong CG, Wichman SR, Dawson JW, Gardner RC (2001) Stepping stones to Hawaii: a trans-equatorial dispersal pathway for *Metrosideros* (Myrtaceae) inferred from nrDNA (ITS plus ETS). *Journal of Biogeography* **28**, 769–774. doi:10.1046/j.1365-2699.2001.00605.x

Handling editor: David Baum

AUTHOR QUERIES

1. Please also provide a supplier and supplier location (or URL) for the program, as well as the citation of the paper that introduced RAxML. Ditto for MrBayes