Ancient Gondwana break-up explains the distribution of the mycoheterotrophic family Corsiaceae (Liliales)

Mennes, C.B.; Lam, V.K.Y.; Rudall, P.J.; Lyon, S.P.; Graham, S.W.; Smets, E.F.; Merckx, V.S.F.T.

Citation

Mennes, C. B., Lam, V. K. Y., Rudall, P. J., Lyon, S. P., Graham, S. W., Smets, E. F., & Merckx, V. S. F. T. (2015). Ancient Gondwana break-up explains the distribution of the mycoheterotrophic family Corsiaceae (Liliales). *Journal Of Biogeography*, *42*(6), 1123-1136. doi:10.1111/jbi.12486

Journal of Biogeography (*J. Biogeogr.*) (2015) **42**, 1123–1136

ORIGINAL ARTICLE

Ancient Gondwana break-up explains the distribution of the mycoheterotrophic family Corsiaceae (Liliales)

Constantijn B. Mennes[1,*], Vivienne K. Y. Lam[2], Paula J. Rudall[3], Stephanie P. Lyon[4], Sean W. Graham[2], Erik F. Smets[1,5] and Vincent S. F. T. Merckx[1]

[1]Naturalis Biodiversity Center, Leiden University, Leiden, The Netherlands, [2]Department of Botany, University of British Columbia, Vancouver, British Columbia V6T 1Z4, Canada, [3]Royal Botanic Gardens Kew, Richmond, Surrey, UK, [4]Department of Botany, University of Wisconsin Madison, Madison, WI 54706, USA, [5]Section Ecology, Evolution and Biodiversity Conservation, KU Leuven, BE-3001 Leuven, Belgium

*Correspondence: Constantijn B. Mennes, Naturalis Biodiversity Center, Leiden University, PO Box 9517, Leiden, The Netherlands.
E-mail: constantijn.mennes@naturalis.nl

ABSTRACT

Aim Many plant families have a disjunct distribution across the southern Pacific Ocean, including the mycoheterotrophic family Corsiaceae, which provides a prime example of this biogeographical pattern. A better grasp of the family's evolutionary relationships is needed to understand its historical biogeography. We therefore aimed to (1) test the uncertain monophyly of Corsiaceae, (2) define its phylogenetic position, and (3) estimate divergence times for the family, allowing us to assess whether the distribution of the family is the result of vicariance.

Location Southern South America and Australasia.

Methods We analysed various combinations of mitochondrial and nuclear data to address the monophyly, phylogenetic position and age of Corsiaceae. To test its monophyly, we used a three-locus data set including most monocot orders, and to infer its exact phylogenetic position, we used a five-locus extended data set. We corroborated these findings using an independent plastome dataset. We then used a two-locus dataset with taxa from all monocot orders, and a three-locus dataset containing only taxa of Liliales, to estimate divergence times using a fossil-calibrated uncorrelated lognormal relaxed-clock approach.

Results Corsiaceae is a monophyletic family and the sister group of Campynemataceae. This clade is the sister group of all other Liliales. The crown age of Corsiaceae is estimated to be 53 Ma (95% confidence interval 30–76 Ma).

Main conclusions Corsiaceae is an ancient family of mycoheterotrophic plants, whose crown age overlaps with the plate-tectonic split of Gondwana, consistent with a vicariance-based explanation for its current distribution.

Keywords

Australasia, BEAST, Corsiaceae, Gondwana, historical biogeography, Liliales, mycoheterotrophy, southern South America, vicariance.

INTRODUCTION

Fully mycoheterotrophic plants are non-photosynthetic and rely on their mycorrhizal symbionts for carbon uptake. They often differ from their autotrophic relatives as a result of reduced vegetative growth and modified floral structures (Leake, 1994), and show elevated rates of molecular evolution (Merckx *et al.*, 2006, 2009) that can obscure their systematic affinities. Many clades of mycoheterotrophic plants are characterized by remarkably disjunct distribution patterns (Leake, 1994; Merckx *et al.*, 2013a). For example, the genus *Seychellaria* (Triuridaceae) grows disjunctly in Madagascar, the Seychelles and Tanzania (Maas-van de Kamer & Weustenfeld, 1998), and the otherwise South American genus *Voyria* (Gentianaceae) has a single species (*V. primuloides*) known from West and Central Africa (Merckx *et al.*, 2013b). Such distribution patterns have often been interpreted as indicative of an ancient widespread distribution for these lineages (e.g. Jonker, 1938; Leake, 1994) but, at least for *Seychellaria* and *Voyria*, dispersal mechanisms have recently

doi:10.1111/jbi.12486

been inferred (Mennes *et al.*, 2013; Merckx *et al.*, 2013b). The divergence ages of many mycoheterotrophic clades remain to be investigated.

One family of particular interest is the mycoheterotrophic Corsiaceae Becc. (Fig. 1), which is found disjunctly in South America, Australasia and China. This family currently consists of approximately 30 species in three genera (Neinhuis & Ibisch, 1998; Zhang *et al.*, 1999; Jones & Gray, 2008). The tropical genus *Corsia* Becc. (*c.* 27 species) occurs in New Guinea, the Solomon Islands and northern Australia (van Royen, 1972; Neinhuis & Ibisch, 1998; Jones & Gray, 2008). The temperate genus *Arachnitis* Phil. (one or two species) is found in southern South America, including the Falkland Islands, and reaches the northern limit of its distribution in Bolivia (Ibisch *et al.*, 1996; Neinhuis & Ibisch, 1998). The poorly known and monotypic genus *Corsiopsis* D.X. Zhang, R.M.K. Saunders & C.M. Hu has only been recorded in southern China and its description is based on a single specimen collected in 1974 (Zhang *et al.*, 1999). The disjunct distribution of Corsiaceae was suggested to be the result of a Gondwanan link between South America and Australasia via Antarctica (Zhang *et al.*, 1999).

Members of the Corsiaceae are characterized by a prominent floral synapomorphy: the outer median tepal is enlarged, forming a distinct labellum (van Royen, 1972; Neinhuis & Ibisch, 1998; Rudall & Bateman, 2002). Before molecular data were used to infer the phylogenetic relationships of Corsiaceae, the group was classified as a family or tribe closely related to Burmanniaceae, Thismiaceae or sometimes Geosiridaceae or Orchidaceae (i.e. other monocot families that contain mycoheterotrophic species) (Bentham & Hooker, 1883; Engler, 1889; Jonker, 1938; Cronquist, 1981; Thorne, 1983; Dahlgren *et al.*, 1985; Rübsamen, 1986). In the morphological cladistic analysis of Stevenson & Loconte (1995), Corsiaceae and other families containing mycoheterotrophic plants were grouped in Orchidales, together with the bigeneric chlorophyllous family Campynemataceae (*Campynema* Labill. and *Campynemanthe* Baill.). Neyland & Hennigan (2003) inferred a close relationship between *Corsia* and *Campynema* based on 26S rDNA data from 45 monocot taxa (including two species of *Arachnitis* and one of *Corsia*). The resulting clade was strongly supported, but there was only weak support for it belonging to a larger clade with other species of Liliales. They also did not recover Corsiaceae as monophyletic: *Arachnitis* was instead inferred to be sister to *Thismia* (Thismiaceae, Dioscoreales), but with poor support for this relationship. Based on several morphological differences (e.g. number of stigmas, size of seeds, presence of a rhizome), Ibisch *et al.* (1996) also raised the question of whether *Corsia* and *Arachnitis* are closely related. Later analyses that lacked *Corsia* found *Arachnitis* to be poorly supported as the sister group of Liliales based on mitochondrial *atp*A (Davis *et al.*, 2004), *atp*A and nuclear 18S rDNA (Fay *et al.*, 2006), or the mitochondrial genes *atp*A, *cob* and *nad*5 (Petersen *et al.*, 2013). Kim *et al.* (2013) recovered *Arachnitis* as a poorly supported sister group of a large clade comprising commelinids, Asparagales and Liliales, based on a single plastid *rbc*L sequence (see below). This lack of support for

Figure 1 Four species of Corsiaceae included in this study: (a) *Corsia* cf. *brassii* P.Royen (photo: S.P. Lyon); (b) *Corsia* cf. *huonensis* P.Royen (photo: S.P. Lyon), (c) *Arachnitis uniflora* Phil. (photo: M. Renny); and (d) *Corsia* cf. *boridiensis* P.Royen (photo: S.P. Lyon).

the phylogenetic position of Corsiaceae has been explicitly noted by many authors (Fay *et al.*, 2006; Kim *et al.*, 2013; Petersen *et al.*, 2013). In summary, phylogenetic analyses based on DNA sequence data are not entirely in accordance with earlier classifications, as these phylogenetic analyses mostly suggest a close relationship between Corsiaceae and Liliales, rather than a close relationship of Corsiaceae with other mycoheterotrophic taxa. The precise phylogenetic position and monophyly of Corsiaceae also remain ambiguous, which has prevented attempts to address questions about the intriguing disjunct distribution pattern of the family.

Disjunct distribution patterns across the southern Pacific Ocean are observed in many angiosperm families. For a long time, it was thought that these distributions were explained by a vicariant scenario resulting from the breakup of Gondwana (Raven & Axelrod, 1972), but recent insights suggest that vicariance cannot be the only explanation. Particularly in plants, Sanmartín & Ronquist (2004) found incongruence between Southern Hemisphere area cladograms and the generally accepted sequential breakup of Gondwana. This incongruence indicates that ancient vicariance patterns may have been obscured by more recent dispersal and/or extinction events. Subsequent biogeographical inferences based on dated phylogenies of many groups, such as *Aristotelia* (Elaeocarpaceae), *Gaultheria* (Ericaceae) and *Muehlenbeckia* (Polygonaceae), indicated that long-distance dispersal between Australasia and South America is the most likely explanation of the distribution patterns of these taxa (Crayn *et al.*, 2006; Bush *et al.*, 2009; Schuster *et al.*, 2013). Conversely, vicariance has been suggested to explain the extant distributions of a few other taxa (e.g. tribe Embothriinae, Proteaceae, Barker *et al.*, 2007; *Raukaua*, Araliaceae, Mitchell *et al.*, 2012). A notorious case is the family Nothofagaceae, for which both long-distance dispersal (e.g. Cook & Crisp, 2005) and vicariance (Heads, 2006) have been proposed independently as the most likely mechanism explaining its distribution.

The mycoheterotrophic family Corsiaceae is a prime candidate for testing ancient Gondwanan vicariance, although this is contingent on resolving its monophyly and phylogenetic affinities. We therefore aimed to (1) test the monophyly of Corsiaceae, (2) define its exact phylogenetic position, and (3) estimate divergence times, in order to investigate whether its disjunct distribution is consistent with ancient vicariance events or more recent long-distance dispersal. Because the phylogenetic placement of Corsiaceae has proved to be a difficult problem (e.g. Kim *et al.*, 2013; Petersen *et al.*, 2013), probably due to the rate variation that is typical of mycoheterotrophic plants (e.g. Merckx *et al.*, 2006, 2009), we primarily used nuclear and mitochondrial data, which were demonstrated to be informative both for other mycoheterotrophic taxa and for Liliales as a whole (e.g. Mennes *et al.*, 2013; Petersen *et al.*, 2013). The use of plastid data has generally lagged in mycoheterotrophic plant groups, because non-photosynthetic plants tend to lack the photosynthetic genes that are commonly used for plant phylogenetic inference (e.g. Olmstead & Palmer, 1994). Recent advances in sequencing technology have, however, facilitated the recovery of whole plastid genomes from photosynthetic and non-photosynthetic monocots alike (e.g. Givnish *et al.*, 2010; Barrett & Davis, 2012; Logacheva *et al.*, 2014). We used this approach to recover plastid genes sets from *Arachnitis* and *Corsia*, allowing us to corroborate the monophyly and phylogenetic position of Corsiaceae inferred from mitochondrial and nuclear data. We then estimated divergence times using a subset of the nuclear and mitochondrial data, based on an extensive set of monocot taxa and multiple fossil calibration points.

MATERIALS AND METHODS

Taxon sampling

Nuclear and mitochondrial data

Four different nuclear and mitochondrial datasets were compiled to study different evolutionary questions. (1) To maximize the taxon sampling for testing the monophyly of Corsiaceae as well as its position in the monocots, a three-locus dataset (nuclear 18S rDNA and mitochondrial *atp*A and *mat*R) was assembled for 33 species of Liliales (representing all nine recognized families) and sequence data from most monocot orders. (2) To maximize the analysed sequence data for Liliales for inferring the position of Corsiaceae, a five-locus dataset (nuclear 18S rDNA and mitochondrial *atp*A, *mat*R, *cob* and *nad*5) was assembled for 33 taxa of Liliales, with Pandanales, Dioscoreales and Alismatales included as outgroups. (3) To estimate general divergence times in monocots, a two-locus dataset (nuclear 18S rDNA and mitochondrial *atp*A) similar to the one used by Mennes *et al.* (2013) was assembled that contained sequence data from all monocot orders. (4) To obtain divergence-time estimates in Liliales, based on more data (i.e. more reliably, possibly with a smaller 95% confidence interval), a three-locus dataset (nuclear 18S rDNA and mitochondrial *atp*A and *mat*R) was assembled that contained only sequence data from Liliales. These latter two datasets were analysed using a fossil-calibrated uncorrelated lognormal relaxed-clock approach (see Table S1 in Appendix S1 of the Supporting Information for details of accessions and vouchers). These four analyses, based on different subsets of the same nuclear and mitochondrial regions for monocots, collectively provide insights into the evolutionary biogeographical history of Corsiaceae based on two of the three plant genomes. (See Appendix S2 for details of DNA extraction, PCR and sequencing.)

Plastid genome data

As independent confirmation of the position of Corsiaceae, we also retrieved plastid gene sets for 82 genes from 63 taxa, including newly sequenced species from Liliales and Pandanales (see Table S1 in Appendix S1). These data represent 78 protein-coding genes and four rDNA genes (Appendix S2). The taxon sampling includes all monocot orders and the monocot family Dasypogonaceae, and includes 11 additional angiosperms (eudicots, magnoliids and ANITA-grade taxa). (See Appendix S2 for details of DNA extraction, PCR and sequencing.)

Phylogenetic analyses

Nuclear and mitochondrial data

Assembly of DNA sequences, manual editing and alignment were performed using Geneious Pro 6.1.7 (Biomatters, Auckland, New Zealand; available at: http://www.geneious.com/). For phylogenetic inference, we used maximum likelihood (ML) as optimality criterion, as well as Bayesian inference (BI) (see Appendix S2 for further details). We repeated the analysis of the three-locus dataset with *Arachnitis* and *Corsia* excluded in turn, to evaluate the individual and independent placement of each genus.

Plastid data

We inferred a plastid-based phylogeny using maximum likelihood. We partitioned the data by gene and codon position for a total of 241 initial data partitions. We used PartitionFinder 1.1.1 (Lanfear *et al.*, 2012) to group together data partitions with similar models of DNA substitution. We evaluated branch support for the ML analysis using 500 bootstrap replicates (Felsenstein, 1985) (see Appendix S2 for further details).

Estimation of divergence times

Divergence times of taxa in Corsiaceae were estimated by performing two unlinked uncorrelated lognormal relaxed-clock analyses using beast 2.1.3 (Bouckaert *et al.*, 2014) (see Appendix S2 for the complete workflow). The Yule process of speciation (Gernhard, 2008) was set as a tree prior in both analyses. In the first analysis, based on the two-locus dataset described above, calibration-point priors were modelled as lognormal distributions, with the ages of six dated angiosperm fossils used as offset values, following a recent review (Iles *et al.*, 2015). (1) Based on fossils of *Spirematospermum chandlerae*, the stem node of Zingiberales was constrained to a minimum age of 72.1 Ma (Friis, 1988; Sohl & Owens, 1991). (2) Based on fossils of *Sabalites carolinensis*, the crown node of Arecaceae was constrained to a minimum age of 83.6 Ma (Berry, 1914; Gohn *et al.*, 1992). (3) Based on fossils of *Changii indicum*, the crown node of Poaceae was constrained to a minimum age of 66 Ma (Courtillot & Renne, 2003; Prasad *et al.*, 2011). (4) Based on fossils of *Liliacidites*, the stem node of Asparagales was constrained to a minimum age of 93 Ma (Ramirez *et al.*, 2007; Gustafsson *et al.*, 2010). (5) Based on fossils of *Ripogonum tasmanicum*, the stem node of Ripogonaceae was constrained to a minimum age of 51 Ma (Carpenter *et al.*, 2007; Conran *et al.*, 2009). (6) Based on fossils of *Luzuriaga*, the stem node of *Luzuriaga* was constrained to a minimum age of 23.2 Ma (Lindqvist & Lee, 2009; Conran *et al.*, 2014). This node corresponds to the crown node of Alstroemeriaceae under our taxon sampling. The root age of the tree was modelled as a normal distribution (mean, 0.0; sigma, 1.0), constrained to an offset age of 134 Ma, based on the crown age of the monocots inferred by Bremer (2000). The substitution rate was set to 1.0 and priors for each substitution rate were modelled as gamma distributions with default parameter values. The gamma shape prior of the substitution model and the priors of the mean and standard deviation of the unlinked uncorrelated lognormal clock models were set as exponential distributions. The default settings were kept, except for the mean of the clock models, which was set to 10.0. All other priors were kept to their default settings. Clades in the phylogeny were constrained according to the latest insights in monocot phylogeny (Chase *et al.*, 2006; Merckx *et al.*, 2009; Soltis *et al.*, 2011; Barrett *et al.*, 2013; Mennes *et al.*, 2013), as well as the Liliales phylogeny from this study. Clades were constrained to order level, and clades within Dioscoreales, Pandanales and Liliales were constrained to family or intrafamilial level. The chain length of the Markov chain Monte Carlo (MCMC; Geyer, 1991) algorithm was 200 million generations, sampling one tree per 10,000 generations. In the second analysis, only the fossil calibration points from Liliales were included, using the following parameter settings. The substitution rate was set to 1.0, priors for each substitution were modelled as gamma distributions (default parameter values), the gamma shape prior of the substitution model and the priors of the mean and standard deviation of the clock models were set as exponential distributions. The mean of the clock models was set to 10.0, and the remaining settings were kept at their default settings. Additionally, the posterior probability distribution obtained for the crown age of Liliales resulting from the first divergence-times estimation was set as an additional secondary calibration point. This prior was modelled as a normal distribution (mean, 85 Ma; sigma, 9.7 Myr). All remaining prior settings were identical to the first analysis and kept to their default settings.

RESULTS

Phylogenetic analyses

The BI phylogenetic inferences for the complete three-locus dataset (5741 bp) recovered Corsiaceae as monophyletic with strong support (100% posterior probability, PP; 100% bootstrap support, BS). This clade was sister to Campynemataceae (Liliales), also with strong support (100% PP, 100% BS; Fig. 2). *Arachnitis* and *Corsia* were recovered in this position independently in analyses of the equivalent data set that excluded either *Arachnitis* (see Fig. S1 in Appendix S3; based on 5822 bp) or *Corsia* (see Fig. S2 in Appendix S3; based on 5881 bp). The phylogenies resulting from the BI analyses are

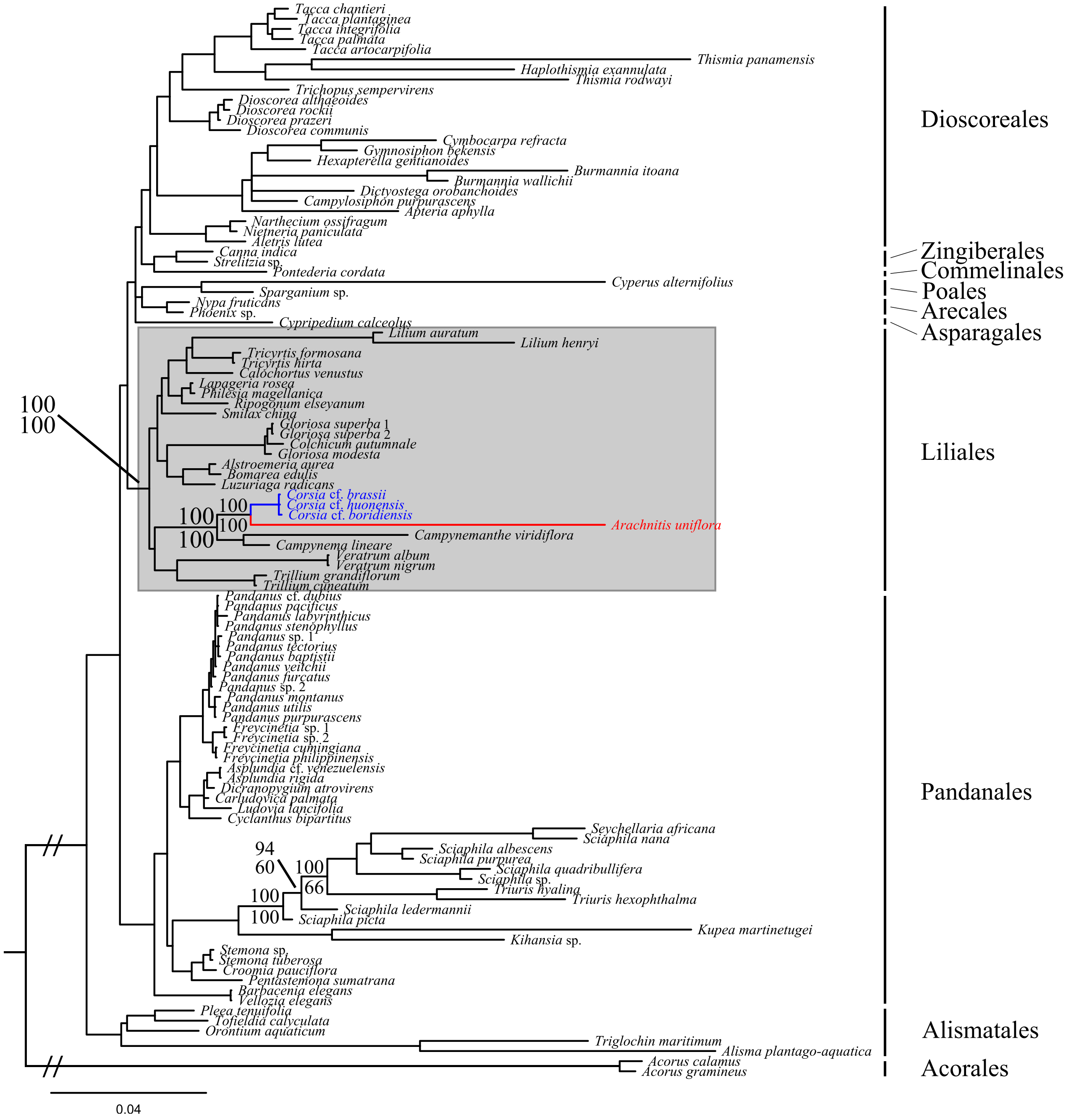

Figure 2 Phylogeny of Liliales and related monocot orders based on a Bayesian-inference (BI) analysis of a three-locus molecular dataset (nuclear 18S rDNA and mitochondrial *atp*A and *mat*R), including all available *Corsia* (blue) and *Arachnitis* (red) species from this study. The grey shaded area shows the Liliales taxa. Values above branches are posterior probabilities (expressed as percentages) resulting from the BI analysis; those below branches are bootstrap support percentages resulting from the maximum-likelihood (ML) analysis. Support values are only given for Corsiaceae and the weakly conflicting Triuridaceae taxa (see text). Dashes ('//') indicate branches that are longer than indicated in the figure. The scale bar indicates the number of substitutions per site. (See Figs S1 & S2 in Appendix S3 for phylogenies based on the same analyses, but excluding *Arachnitis* and *Corsia*, respectively.)

not in conflict with those resulting from the ML analyses (not shown), although removal of either *Arachnitis* or *Corsia* resulted in a slight topological change within Triuridaceae for both BI and ML analyses (concerning the relative position of *Triuris* spp. and *Sciaphila ledermannii*); this was not a strong conflict, in that only the BI analysis depicted a relationship with strong support (100% PP; Fig. 2, Figs S1 & S2). In the analysis that excluded *Arachnitis*, the standard deviation of split frequencies was 0.015 (i.e. higher than 0.01) after 1,000,000 additional generations.

The phylogeny resulting from the BI analysis of the five-locus dataset (Fig. 3; 6965 bp) is not in conflict with that

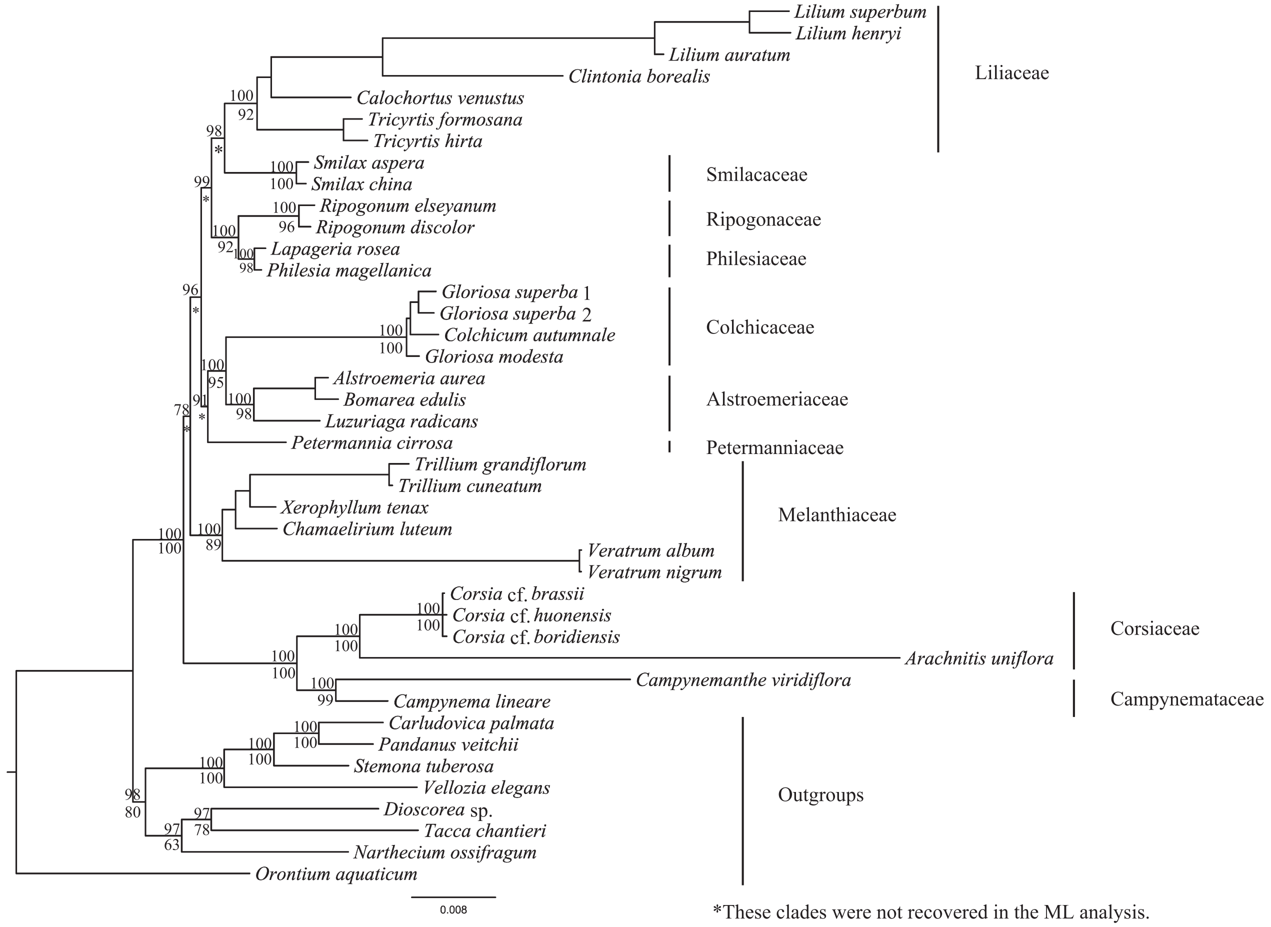

Figure 3 Phylogeny of Liliales based on a Bayesian-inference (BI) analysis of a five-locus molecular dataset (nuclear 18S rDNA and mitochondrial *atp*A, *mat*R, *cob* and *nad*5). Values above branches are posterior probabilities (expressed as percentages) resulting from the BI analysis; those below branches are bootstrap support percentages resulting from the maximum-likelihood (ML) analysis. The scale bar indicates the number of substitutions per site.

resulting from the ML analysis (not shown). The trees resulting from separate analyses of individual genes (see Figs S3–S7 in Appendix S3) do not show highly supported conflict. In the separate *nad*5 analysis, the standard deviation of split frequencies was 0.023 after 1,000,000 additional generations. Additionally, in one of the runs, a single effective sample size (ESS) did not reach a sufficient value (i.e. above 100), namely the ESS of tree length (sum of all branch lengths, TL)=69. With the current taxon sampling, the monophyly of Liliales is strongly supported (100% PP; 100% BS), and all the families in Liliales were recovered as monophyletic. Within Corsiaceae, *Corsia* was represented by three species that formed a strongly supported monophyletic group (100% PP; 100% BS). A group comprising Corsiaceae and Campynemataceae was also recovered as monophyletic (100% PP; 100% BS). Limited support (78% PP; not recovered in ML analysis) was found for a clade consisting of all Liliales families excluding Corsiaceae and Campynemataceae, sister to the latter two families.

The plastid dataset contained full gene data (protein-coding genes and plastid rDNA genes) for representatives of all monocot orders, supplemented by additional angiosperms. A total of 83 plastid genes were retrieved for most species (*ycf*1 was retrieved but not included in the analysis due to alignment difficulties). Corsiaceae was represented by *Arachnitis uniflora* and *Corsia* cf. *boridiensis*. For *A. uniflora*, 20 non-photosynthetic genes were retrieved, and for *C.* cf. *boridiensis*, 27 non-photosynthetic genes were retrieved (see Table S2 in Appendix S1). No photosynthetic genes were retrieved as open reading frames in the plastid genomes for Corsiaceae, although several pseudogenes were recovered. A number of tRNA sequences were also retrieved, but were not included in the analysis. With the current taxon sampling (which lacks any representatives of Colchicaceae, Ripogonaceae, Petermanniaceae or Philesiaceae), the relationships in the order Liliales agree with the results from the five-locus analysis (Fig. 4). Corsiaceae and Campynemataceae are sister to all other Liliales and, notably, the clade containing all other sampled members of Liliales is strongly supported (100% BS). The relationships among the rest of the monocots also agree with those in recent studies (e.g. Soltis *et al.*, 2011; Barrett *et al.*, 2013).

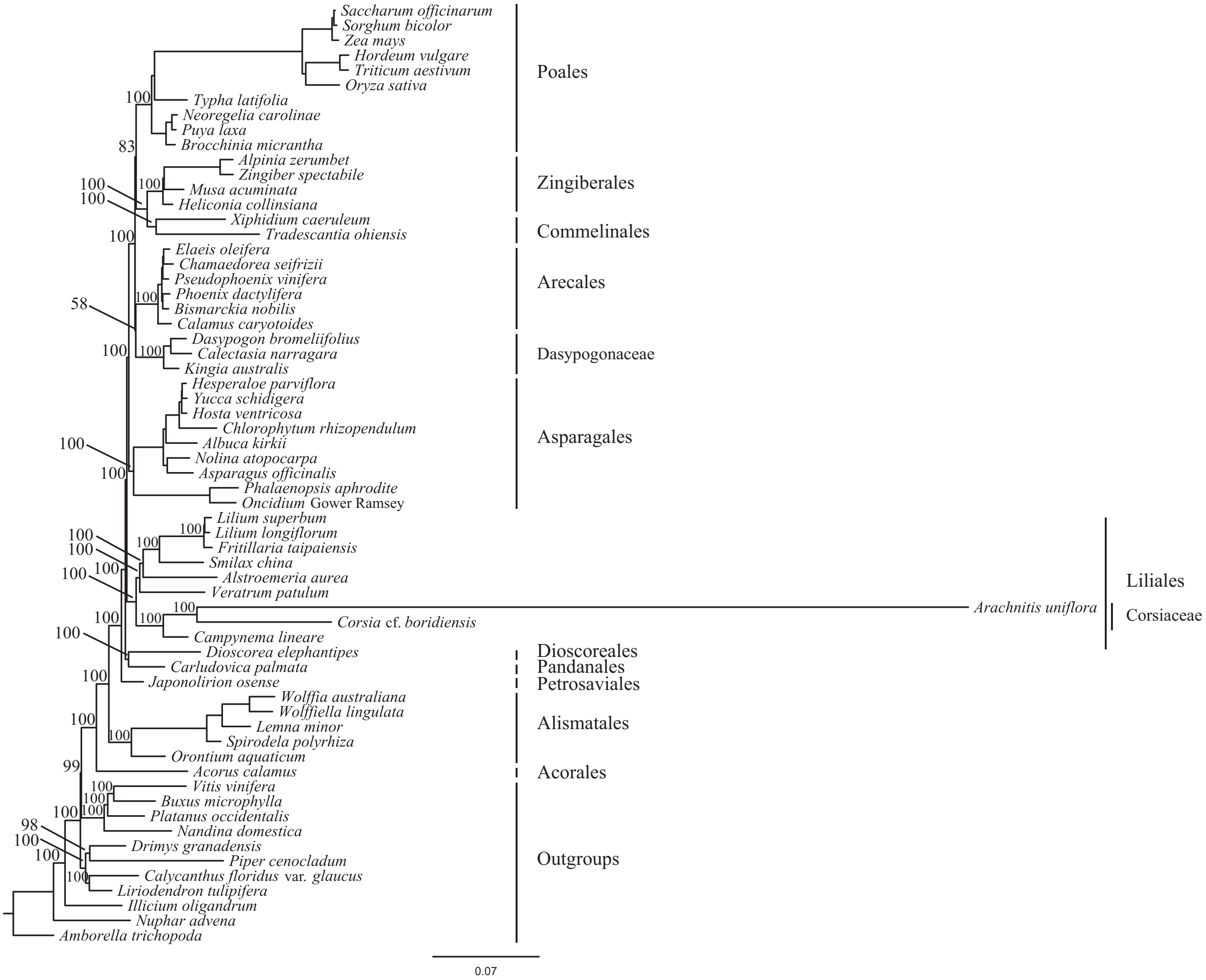

Figure 4 Phylogeny of monocots based on a maximum-likelihood (ML) analysis of an 82-gene molecular plastid genome dataset. Values above branches are bootstrap support percentages. The scale bar indicates the number of substitutions per site.

Estimation of divergence times

The estimated ages for Liliales clades are roughly the same in both analyses, except for Corsiaceae, which had notably older age estimates in the three-locus analysis (Table 1). For Corsiaceae, the three-locus analysis resulted in a mean stem age of 70 Ma (95% confidence interval, CI, 47–93 Ma; Fig. 5). The mean crown age of Corsiaceae was estimated at 53 Ma (95% CI, 30–76 Ma; Fig. 5). The two-locus analysis resulted in a mean stem age of 54 Ma (95% CI, 30–78 Ma; see Fig. S8 in Appendix S3), and a mean crown age of 36 Ma (95% CI, 14–58 Ma; Fig. S8). The estimates obtained in the two-locus monocot-wide analysis for the crown and stem-node ages of monocot orders are generally in accordance with the findings of earlier studies (Table 2). Within Liliales, the results of Chacón *et al.* (2012) resemble our results in showing a crown age of Campynemataceae of 36.5 Ma (95% CI, 13–61 Ma), compared to our estimate of 41 Ma (95% CI, 16–70 Ma). Differences between the two studies are probably the result of differences in the use of a fossil calibration point in *Luzuriaga*, reflecting our taxon sampling – we conservatively constrained the crown node of Alstroemeriaceae using this fossil, whereas Chacón *et al.* (2012) were able to constrain a subclade in this family using the same fossil, because of their denser taxon sampling (see Iles *et al.*, 2015, for justification).

DISCUSSION

Phylogenetic affinities and monophyly of Corsiaceae

The monophyletic relationship revealed between *Corsia* and *Arachnitis* (Figs 2–4) supports the current circumscription of the family (van Royen, 1972; Cronquist, 1981; Dahlgren *et al.*, 1985; Neinhuis & Ibisch, 1998; Zhang *et al.*, 1999; Jones & Gray, 2008). Although the members of Corsiaceae form a well-supported clade, the genus *Arachnitis* is always placed on a rather long branch (Figs 2–4). Thismiaceae taxa

Table 1 Estimated divergence times of Liliales.

Family	Node	Age estimation (mean and 95% confidence interval in Ma)	
		Two-locus monocot-wide analysis	Three-locus Liliales analysis
Alstroemeriaceae	Stem	44 (28–61)	45 (27–63)
	Crown	26 (23.2–34)	28 (23.2–39)
Campynemataceae	Stem	54 (30–78)	70 (47–93)
	Crown	38 (12–62)	41 (16–70)
Colchicaceae	Stem	44 (28–61)	45 (27–63)
	Crown	10 (2–21)	9 (3–17)
Corsiaceae	Stem	54 (30–78)	70 (47–93)
	Crown	36 (14–58)	53 (30–76)
Liliaceae	Stem	59 (44–73)	56 (42–69)
	Crown	53 (38–68)	46 (31–60)
Liliales	Stem	110 (97–122)	n/a
	Crown	85 (67–104)	90 (75–106)
Melanthiaceae	Stem	74 (59–90)	77 (63–93)
	Crown	54 (32–77)	59 (36–80)
Petermanniaceae	Stem	56 (37–74)	n/a
	Crown	n/a	n/a
Philesiaceae	Stem	53 (51–58)	53 (51–56)
	Crown	6 (0.3–16)	6 (0.4–14)
Ripogonaceae	Stem	53 (51–58)	53 (51–56)
	Crown	n/a	n/a
Smilacaceae	Stem	59 (44–73)	56 (42–69)
	Crown	n/a	n/a

are also placed on long branches, suggesting that the weakly recovered affinities between *Arachnitis* and *Thismia* in an earlier analysis (Neyland & Hennigan, 2003) were the result of long-branch attraction (Felsenstein, 1978), possibly exacerbated by the relatively limited taxon sampling in that study. In our analyses, *Arachnitis* is sister to Campynemataceae rather than to *Thismia* when *Corsia* is excluded (and *Corsia* emerges as sister to Campynemataceae after the removal of *Arachnitis*) (Figs S1 & S2). Moreover, both the analyses resulting from the nuclear and mitochondrial datasets (Figs 2 & 3) and the plastid dataset (Fig. 4) independently showed strong support for a sister-group relationship between *Arachnitis* and *Corsia*. These results collectively indicate that the placement of *Arachnitis* inferred in this study is not an artefact of long-branch attraction; the family is monophyletic under the current taxon sampling. The inclusion of more species of Corsiaceae, particularly *Corsiopsis*, would further test the monophyly of the family.

The sister-group relationship of Corsiaceae and Campynemataceae inferred from all data sources investigated here (Figs 2–4) identifies Campynemataceae as the closest chlorophyllous relatives of the fully mycoheterotrophic Corsiaceae. Corsiaceae and Campynemataceae then form a clade that is sister to the rest of Liliales, well supported by the plastid genome data. These findings moreover are in accordance with the previous molecular analyses of Davis *et al.* (2004), Fay *et al.* (2006) and Petersen *et al.* (2013), based on analyses that included a single taxon of each family, although we recovered higher support values for the clade containing Corsiaceae and Campynemataceae. Erdtman (1952) had earlier suggested the affinity of Corsiaceae with Campynemataceae, based on similarities in pollen surface patterning. Based on plastid *rbc*L sequence data, Kim *et al.* (2013) had found a possible sister-group relationship for Corsiaceae with a clade comprising commelinids, Asparagales and Liliales. However, BLAST results of this sequence (obtained from the authors) showed high sequence similarity (95.8%) with *Aglaia* (Meliaceae, core eudicots), suggesting that the sequence does not belong to *Arachnitis*. Furthermore, *rbc*L is not present in the draft circular plastid genome of *A. uniflora*.

All the studied datasets indicate that Liliales forms a well-supported monophyletic order (Figs 2–4) containing the same monophyletic families as in other recent phylogenetic studies of Liliales (Kim *et al.*, 2013; Petersen *et al.*, 2013). The clade containing all sampled members of Liliales except Corsiaceae and Campynemataceae is more strongly supported by the plastid data than by the mitochondrial and nuclear data. The order is supported by some strong morphological synapomorphies, especially the presence of perigonal (tepal-borne) nectaries, compared with septal (ovary-borne) nectaries in many other monocots (Rudall *et al.*, 2000; Smets *et al.*, 2000). The absence of septal nectaries is often linked with congenital carpel fusion (Rudall, 2002; Remizowa *et al.*, 2010). The relationships between other families within Liliales are roughly in accordance with previous molecular analyses (Kim *et al.*, 2013; Petersen *et al.*, 2013) (Fig. 3). The relationships between most other monocot lineages (Fig. 4) are also in accordance with earlier studies (e.g. Soltis *et al.*, 2011; Barrett *et al.*, 2013).

There are few obvious morphological synapomorphies shared by the families Campynemataceae and Corsiaceae, partly because Corsiaceae have reduced vegetative characters.

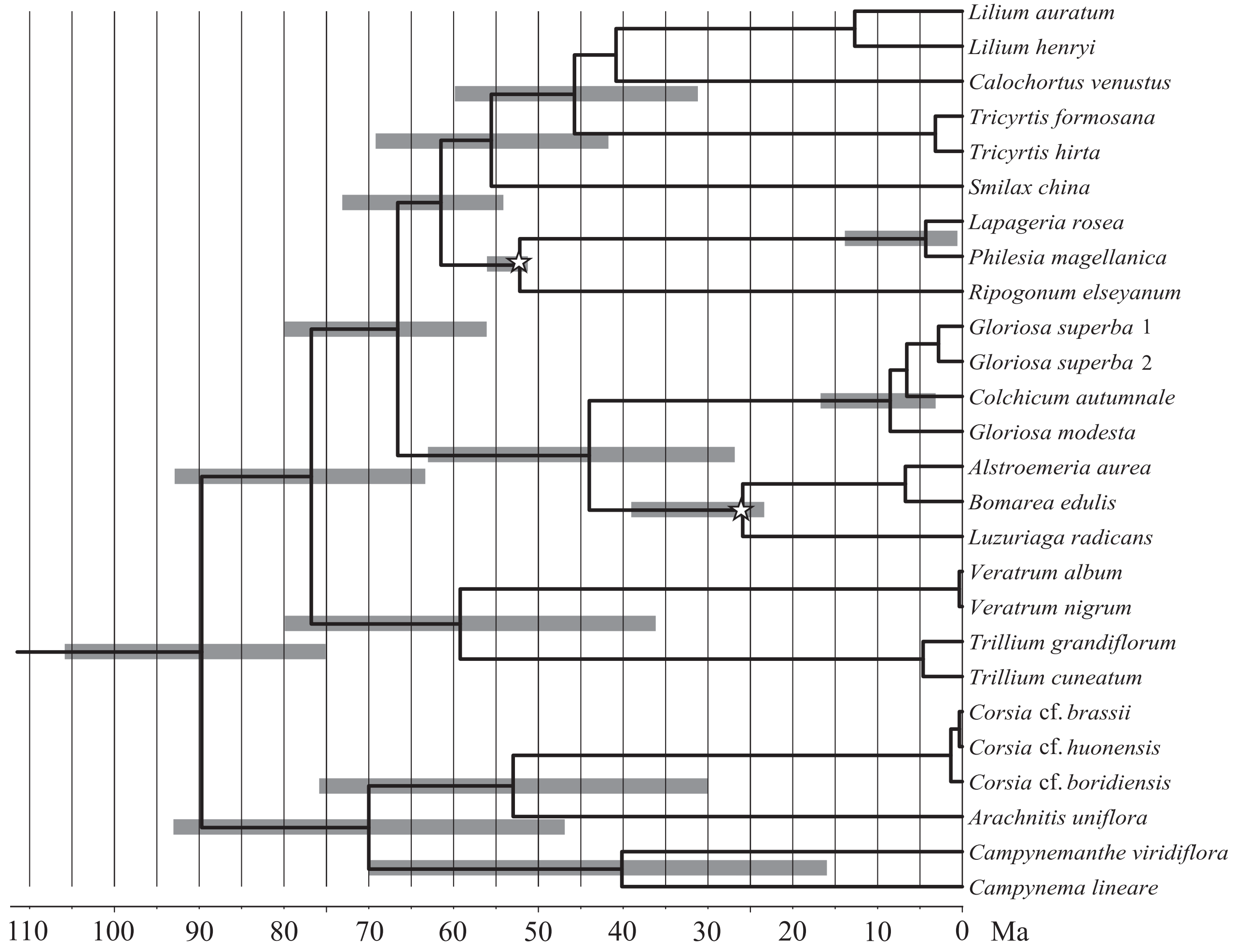

Figure 5 Maximum clade credibility tree resulting from the divergence time estimation of Liliales inferred using the uncorrelated lognormal relaxed clock method, based on a three-locus molecular dataset (nuclear 18S rDNA and mitochondrial *atp*A and *mat*R), two fossil calibration points, one secondary calibration point and assuming a Yule process of speciation (see text). The scale bar and the numbers in the figure both represent time (million years ago; Ma), and bars around nodes represent 95% confidence intervals. Nodes with 95% confidence interval bars were constrained. Stars indicate fossil calibration points (see text).

Erdtman (1952) suggested similarities in pollen sexine sculpturing between Corsiaceae and Campynemataceae (as 'Campynematoideae'). However, more extensive comparative studies of pollen in Liliales (C.A. Furness and P.J. Rudall, RBG Kew, personal communication) show that both families have reticulate or microreticulate pollen sculpturing of a type that is relatively common in Liliales and could represent a plesiomorphic condition for the order as a whole. Pollen of Corsiaceae and *Campynema* is monosulcate (porate in *Corsia*), but inaperturate in *Campynemanthe*. Perhaps the strongest similarity between the two families is the shared presence of an inferior ovary, a feature that is otherwise rare in Liliales, although also present in Alstroemeriaceae (Rudall, 2002; Rudall & Eastman, 2002). The inferior ovary is unilocular in *Arachnitis* and probably also in *Corsiopsis*, but trilocular in *Corsia* and Campynemataceae. The style is almost entirely fused in *Corsia* (albeit trilobed at the tip), but has separate styluli in *Arachnitis* and Campynemataceae, whereas the stigmas are sessile and fused in *Corsiopsis* (Zhang *et al.*, 1999). The primary synapomorphy of Corsiaceae is the presence of a labellum that is (uniquely among monocots) formed from the outer median tepal (Rudall & Bateman, 2002). The labellum is reflexed forwards in *Arachnitis* and *Corsiopsis*, but more upright in some species of *Corsia,* in which it has a prominent callus that could be nectariferous, like the perigonal nectaries of other Liliales, or (perhaps more likely) could represent osmophores. No labellum is present in Campynemataceae, although the tepals also possess thickened regions.

Biogeographical history of Corsiaceae

The estimated divergence times are roughly in accordance with previous studies, as shown by overlapping 95% confidence intervals (Table 2; Vinnersten & Bremer, 2001; Janssen & Bremer, 2004; Merckx *et al.*, 2008; Chacón *et al.*, 2012; Mennes *et al.*, 2013). Differences from earlier studies are likely to be the result of different fossil calibration points or

Table 2 Estimated divergence times of monocots.

		95% confidence interval (Ma)			
Order	Node	This study	Mennes *et al.* (2013)	Merckx *et al.* (2008)	Janssen & Bremer (2004) (Ma)
Acorales	Stem	132–136	134–138	134	n/a
	Crown	4–46	4–52	7–44	n/a
Alismatales	Stem	116–135	116–137	123–133	131
	Crown	89–132	90–134	97–133	128
Petrosaviales	Stem	108–132	108–133	121–132	126
	Crown	9–91	8–96	87–102	123
Dioscoreales	Stem	96–123	96–123	119–130	124
	Crown	85–116	85–116	113–126	123
Pandanales	Stem	96–123	96–123	119–130	124
	Crown	70–110	69–110	116–130	114
Liliales	Stem	97–122	97–121	109–131	124
	Crown	67–104	43–105	78–131	117
Asparagales	Stem	94–118	94–116	98–126	122
	Crown	63–113	93–96	101–127	119
Arecales	Stem	87–109	90–93	94–122	120
	Crown	84–90	14–71	15–98	110
Commelinales	Stem	72–82	83–85	83–114	114
	Crown	47–82	47–84	50–104	110
Zingiberales	Stem	72–82	83–85	91–116	114
	Crown	26–66	30–75	52–96	88
Poales	Stem	88–111	84–105	89–120	117
	Crown	82–104	78–98	88–116	113

a different tree topology used, coupled with the use of less comprehensive gene sampling [i.e. the use of a single *rbc*L sequence by Vinnersten & Bremer (2001) and Janssen & Bremer (2004)]. The two divergence-time estimates in this study indicate similar ages for families in Liliales (Table 1). Notable differences are found in the estimated stem and crown ages of Corsiaceae, however. This source of uncertainty may be caused by the branch leading to *Arachnitis* being much longer than that leading to *Corsia*, indicating rather strong heterogeneity of substitution rates, which we suspect might have had an effect on the estimated divergence times. Moreover, the low support values for the sister clade of Corsiaceae and Campynemataceae (the rest of Liliales) resulting from analyses of nuclear and mitochondrial data, might further complicate the analysis. Inspection of ESS values indicates that the two-locus monocot dataset only reached moderately high values (between 200 and 300) for Corsiaceae, Campynemataceae and the clade consisting of both families, after 200 million generations. The three-locus Liliales dataset reached ESS values above 3000 for the same clades, and we therefore conclude that this dataset may provide more reliable divergence-time estimates for these nodes.

The inferred sister-group relationship between the genera *Arachnitis* and *Corsia* establishes Corsiaceae as a genuine example of a disjunct South American–Australasian plant lineage. Based on the three-locus Liliales dataset, the estimated age for the crown node of Corsiaceae (53 Ma; 95% CI, 30–76 Ma; Fig. 5) overlaps tightly with age estimates of the tectonic splits between Australasia and South America/Antarctica (35.5–52 Ma) and South America and Antarctica (36 Ma) (Scotese *et al.*, 1988; Veevers *et al.*, 1991; Woodburne & Case, 1996). This implies that the current disjunct distribution of Corsiaceae is best explained by Gondwanan vicariance, although the lower boundary of the 95% CI (30 Ma) leaves the possibility that the distribution of the family is the result of a more recent dispersal event. Moreover, recent dispersal events combined with extinction (possibly in Australia) cannot be completely ruled out.

Nevertheless, our results are consistent with a scenario in which Corsiaceae provides a rare example of vicariance between South America and Australasia in Southern Hemisphere biogeography. The mycoheterotrophic lifestyle of the species of Corsiaceae might have prevented subsequent dispersal events, as these species can only grow in symbiosis with mycorrhizal fungi. *Arachnitis uniflora* is known to be highly specialized on a clade of arbuscular mycorrhizal fungi (Bidartondo *et al.*, 2002), which may have limited its dispersal capabilities in the past. Both Corsiaceae and the Tasmanian/New Caledonian Campynemataceae are distributed in the southern Pacific region. The inferred mean stem age of Corsiaceae, which is also the date of its split from Campynemataceae (70 Ma; 95% CI, 47–93 Ma; Fig. 5), is on average older than the age estimates of the tectonic splits between Australasia and South America/Antarctica (35.5–52 Ma), and between South America and Antarctica (36 Ma) (Scotese *et al.*, 1988; Veevers *et al.*, 1991; Woodburne & Case, 1996). This suggests that the two families diverged in Gondwana, although the lower boundary of the 95% CI (47 Ma) leaves the possibility open that the split between the families is more recent than the mentioned tectonic splits. The divergence between *Campynema* (Tasmania) and *Campynemanthe* (New Caledonia), however, is estimated to be more recent (16–70 Ma; Fig. 5) than the plate-tectonic split between Australia and New Caledonia/New Zealand (80 Ma) (Raven & Axelrod, 1972; Scotese *et al.*, 1988). This implies that the current distribution of Campynemataceae is explained by a long-distance dispersal event, as found for many New Caledonian taxa (e.g. Sapotaceae; Swenson *et al.*, 2014). The age estimates based on the two-locus monocot-wide dataset are younger for these families, particularly for Corsiaceae (Table 1), which makes long-distance dispersal events slightly more likely. The mean value for the estimated crown age of Corsiaceae (36 Ma), however, is still in accordance with the described scenarios. We assume that including the Chinese *Corsiopsis* would not alter the findings of this study, although it might involve an additional scenario (e.g. a long-distance dispersal event) within Corsiaceae. Moreover, the placement of *Corsiopsis* might alter the monophyly of Corsiaceae as currently circumscribed. There is, however, only a single collection for *Corsiopsis* (Zhang *et al.*, 1999), and it is not clear whether the genus is extant.

Plastome evolution in Corsiaceae

Our plastome dataset indicates gene loss and putative pseudogenization of genes involved in photosynthesis [i.e. genes involved in ATP synthase, NADH dehydrogenase, electron transport (cytochrome), and photosystems I and II; Wicke *et al.*, 2011] compared with chlorophyllous Liliales species (Tables S1 & S2). We can assume that pseudogenization and the initial plastid gene-loss events related to the loss of photosynthesis took place between the inferred crown and stem ages of Corsiaceae, which implies an apomorphy age for mycoheterotrophy in Corsiaceae of between 30 and 93 Ma. *Arachnitis* seems to have a more degraded plastid genome, as several further genes (the group II intron maturase *mat*K and open reading frames encoding the plastid genes of uncertain function *ycf*1 and *ycf*2) appear to be lost from it but are retained in *Corsia*. Because the split between these genera was estimated as 53 Ma (95% CI, 30–76 Ma), these gene losses in *Arachnitis* are likely to have occurred in the last 76 Myr. The *mat*K locus is hypothesized to be one of the last genes lost after the origin of mycoheterotrophy (Barrett & Davis, 2012). The largest class of genes to be retained include genetic apparatus genes (ribosomal protein genes, ribosomal RNA regions and tRNA genes), similar to the findings in the mycoheterotrophic genus *Petrosavia* (Petrosaviaceae), although some genes directly involved in photosynthesis were retained in that genus (Logacheva *et al.*, 2014) but are not present in Corsiaceae. The retained non-photosynthetic plastid genes show relatively longer branches in the plastid phylogeny (Fig. 4) than those based on the nuclear and mitochondrial data (Fig. 3), and not all genes in this functional class have been retained in the plastid genome, showing that losses are not restricted to photosynthesis-related genes. These results will be addressed in a follow-up study.

ACKNOWLEDGEMENTS

The authors thank Elke Bellefroid, Martin Bidartondo, Barbara Briggs, Tom Givnish, Ray Neyland, Gerda van Uffelen, Ed de Vogel and Rogier van Vugt for their help and assistance in obtaining plant material. Furthermore, we thank William Iles for assisting with the selection of fossil calibration points, Marcel Eurlings for technical support and Mauricio Renny for providing photographs. V.M. is financially supported by NWO (Veni fellowship 863.11.018); support from an NSERC Discovery grant to S.W.G., and U.B.C. and NSERC postgraduate fellowships to V.L. are also gratefully acknowledged.

REFERENCES

Barker, N.P., Weston, P.H., Rutschmann, F. & Sauquet, H. (2007) Molecular dating of the ‘Gondwanan’ plant family Proteaceae is only partially congruent with the timing of the break-up of Gondwana. *Journal of Biogeography*, **34**, 2012–2027.

Barrett, C.F. & Davis, J.I. (2012) The plastid genome of the mycoheterotrophic *Corallorhiza striata* (Orchidaceae) is in the relatively early stages of degradation. *American Journal of Botany*, **99**, 1513–1523.

Barrett, C.F., Davis, J.I., Leebens-Mack, J., Conran, J.G. & Stevenson, D.W. (2013) Plastid genomes and deep relationships among the commelinid monocot angiosperms. *Cladistics*, **29**, 65–87.

Bentham, G. & Hooker, J.D. (1883) Ordo CLXVIII Burmanniaceae. *Genera plantarum*, Vol. 3 (ed. by G. Bentham and J.D. Hooker), pp. 455–460. Reeve & Co., London.

Berry, E.W. (1914) *The Upper Cretaceous and Eocene floras of South Carolina and Georgia*. United States Geological Survey, Washington, DC.

Bidartondo, M.I., Redecker, D., Hijri, I., Wiemken, A., Bruns, T.D., Domínguez, L., Sérsic, A., Leake, J.R. & Read, D.J. (2002) Epiparasitic plants specialized on arbuscular mycorrhizal fungi. *Nature*, **419**, 389–392.

Bouckaert, R., Heled, J., Kühnert, D., Vaughan, T., Wu, C.-H., Xie, D., Suchard, M.A., Rambaut, A. & Drummond, A.J. (2014) BEAST2: a software platform for Bayesian evolutionary analysis. *PLoS Computational Biology*, **10**, e1003537.

Bremer, K. (2000) Early Cretaceous lineages of monocot flowering plants. *Proceedings of the National Academy of Sciences USA*, **97**, 4707–4711.

Bush, C.M., Wagstaff, S.J., Fritsch, P.W. & Kron, K.A. (2009) The phylogeny, biogeography and morphological evolution of *Gaultheria* (Ericaceae) from Australia and New Zealand. *Australian Systematic Botany*, **22**, 229–242.

Carpenter, R.J., Jordan, G.J. & Hill, R.S. (2007) A toothed Lauraceae leaf from the Early Eocene of Tasmania. *International Journal of Plant Sciences*, **168**, 1191–1198.

Chacón, J., Camargo de Assis, M., Meerow, A.W. & Renner, S.S. (2012) From East Gondwana to Central America:

historical biogeography of the Alstroemeriaceae. *Journal of Biogeography*, **39**, 1806–1818.

Chase, M.W., Fay, M.F., Devey, D.S. *et al.* (2006) Multigene analyses of monocot relationships: a summary. *Aliso*, **22**, 63–75.

Conran, J.G., Carpenter, R.J. & Jordan, G.J. (2009) Early Eocene *Ripogonum* (Liliales: Ripogonaceae) leaf macrofossils from southern Australia. *Australian Systematic Botany*, **22**, 219–228.

Conran, J.G., Bannister, J.M., Mildenhall, D.C., Lee, D.E., Chacón, J. & Renner, S.S. (2014) Leaf fossils of *Luzuriaga* and a monocot flower with in situ pollen of *Liliacidites contortus* Mildenh. & Bannister sp. nov. (Alstroemeriaceae) from the Early Miocene. *American Journal of Botany*, **101**, 141–155.

Cook, L.G. & Crisp, M.D. (2005) Not so ancient: the extant crown group of *Nothofagus* represents a post-Gondwanan radiation. *Proceedings of the Royal Society B: Biological Sciences*, **272**, 2535–2544.

Courtillot, V.E. & Renne, P.R. (2003) On the ages of flood basalt events. *Comptes Rendus Geoscience*, **335**, 113–140.

Crayn, D.M., Rossetto, M. & Maynard, D.J. (2006) Molecular phylogeny and dating reveals an Oligo-Miocene radiation of dry-adapted shrubs (former Tremandraceae) from rainforest tree progenitors (Elaeocarpaceae) in Australia. *American Journal of Botany*, **93**, 1328–1342.

Cronquist, A. (1981) *An integrated system of classification of flowering plants*. Columbia University Press, New York.

Dahlgren, R.M.T., Clifford, H.T. & Yeo, P.F. (1985) *The families of the monocotyledons: structure, evolution and taxonomy*. Springer, Berlin.

Davis, J.I., Stevenson, D.W., Petersen, G., Seberg, O., Campbell, L.M., Freudenstein, J.V., Goldman, D.H., Hardy, C.R., Michelangeli, F.A., Simmons, M.P., Specht, C.D., Vergara-Silva, F. & Gandolfo, M. (2004) A phylogeny of the monocots, as inferred from *rbcL* and *atpA* sequence variation, and a comparison of methods for calculating jackknife and bootstrap values. *Systematic Botany*, **29**, 467–510.

Engler, A. (1889) Burmanniaceae. *Die Natürlichen Pflanzenfamiliën, II. Teil* (ed. by A. Engler and K. Prantl), pp. 44–51. Wilhelm Engelmann, Leipzig.

Erdtman, G. (1952) *Pollen morphology and plant taxonomy – angiosperms: an introduction to palynology*. Almqvist & Wiksell, Stockholm, Sweden.

Fay, M.F., Chase, M.W., Rønsted, N., Devey, D.S., Pillon, Y., Pires, J.C., Petersen, G., Seberg, O. & Davis, J.I. (2006) Phylogenetics of Liliales: summarized evidence from combined analyses of five plastid and one mitochondrial loci. *Aliso*, **22**, 559–565.

Felsenstein, J. (1978) Cases in which parsimony or compatibility methods will be positively misleading. *Systematic Zoology*, **27**, 401–410.

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

Friis, E.M. (1988) *Spirematospermum chandlerae* sp. nov., an extinct species of Zingiberaceae from the North American Cretaceous. *Tertiary Research*, **9**, 149–157.

Gernhard, T. (2008) The conditioned reconstructed process. *Journal of Theoretical Biology*, **253**, 769–778.

Geyer, C.J. (1991) Markov chain Monte Carlo maximum likelihood. *Computing science and statistics: Proceedings of the 23rd Symposium on the Interface* (ed. by E.M. Keramidas), pp. 156–163. Interface Foundation of North America, Fairfax Station, VA.

Givnish, T.J., Ames, M., McNeal, J.R., McKain, M.R., Steele, P.R., dePamphilis, C.W., Graham, S.W., Pires, J.C., Stevenson, D.W., Zomlefer, W.B., Briggs, B.G., Duvall, M.R., Moore, M.J., Heaney, J.M., Soltis, D.E., Soltis, P.S., Thiele, K. & Leebens-Mack, J.H. (2010) Assembling the tree of the monocotyledons: plastome sequence phylogeny and evolution of Poales. *Annals of the Missouri Botanical Garden*, **97**, 584–616.

Gohn, G.S., Dowsett, H.J. & Sohl, N.F. (1992) *Biostratigraphy of the Middendorf Formation (Upper Cretaceous) in a corehole at Myrtle Beach, South Carolina*. USGS Bulletin 2030. United States Geological Survey, Denver, CO.

Gustafsson, A.L., Verola, C.F. & Antonelli, A. (2010) Reassessing the temporal evolution of orchids with new fossils and a Bayesian relaxed clock, with implications for the diversification of the rare South American genus *Hoffmannseggella* (Orchidaceae: Epidendroideae). *BMC Evolutionary Biology*, **10**, 177.

Heads, M. (2006) Panbiogeography of *Nothofagus* (Nothofagaceae): analysis of the main species massings. *Journal of Biogeography*, **33**, 1066–1075.

Ibisch, P.L., Neinhuis, C. & Rojas, N.P. (1996) On the biology, biogeography, and taxonomy of *Arachnitis* Phil. nom. cons. (*Corsiaceae*) in respect to a new record from Bolivia. *Willdenowia*, **26**, 321–332.

Iles, W.J.D., Smith, S.Y., Gandolfo, M.A. & Graham, S.W. (2015) A review of monocot fossils suitable for molecular dating analyses. *Botanical Journal of the Linnean Society*, DOI: 10.1111/boj.12233 [Epub ahead of print].

Janssen, T. & Bremer, K. (2004) The major monocot groups inferred from 800+ *rbcL* sequences. *Botanical Journal of the Linnean Society*, **146**, 385–398.

Jones, D.L. & Gray, B. (2008) *Corsia dispar* D.L.Jones & B. Gray (Corsiaceae), a new species from Australia, and a new combination in *Corsia* for a New Guinea taxon. *Austrobaileya*, **7**, 717–722.

Jonker, F.P. (1938) *A monograph of the Burmanniaceae*. Mededeelingen van het Botanisch Museum en Herbarium van de Rijks Universiteit te Utrecht 51. University of Utrecht.

Kim, J.S., Hong, J.-K., Chase, M.W., Fay, M.F. & Kim, J.-H. (2013) Familial relationships of the monocot order Liliales based on a molecular phylogenetic analysis using four plastid loci: *matK*, *rbcL*, *atpB* and *atpF-H*. *Botanical Journal of the Linnean Society*, **172**, 5–21.

Lanfear, R., Calcott, B., Ho, S.Y.W. & Guindon, S. (2012) PartitionFinder: combined selection of partitioning schemes and substitution models for phylogenetic analyses. *Molecular Biology and Evolution*, **29**, 1695–1701.

Leake, J.R. (1994) The biology of myco-heterotrophic ('saprophytic') plants. *New Phytologist*, **127**, 171–216.

Lindqvist, J.K. & Lee, D.E. (2009) High-frequency paleoclimate signals from Foulden Maar, Waipiata Volcanic Field, southern New Zealand: an Early Miocene varved lacustrine diatomite deposit. *Sedimentary Geology*, **222**, 98–110.

Logacheva, M.D., Schelkunov, M.I., Nuraliev, M.S., Samigullin, T.H. & Penin, A.A. (2014) The plastid genome of mycoheterotrophic monocot *Petrosavia stellaris* exhibits both gene losses and multiple rearrangements. *Genome Biology and Evolution*, **6**, 238–246.

Maas-van de Kamer, H. & Weustenfeld, T. (1998) Triuridaceae. *The families and genera of flowering plants, Vol. 3, Lilianae (except Orchidaceae)* (ed. by K. Kubitzki), pp. 452–458. Springer, Berlin.

Mennes, C.B., Smets, E.F., Moses, S.N. & Merckx, V.S.F.T. (2013) New insights in the long-debated evolutionary history of Triuridaceae (Pandanales). *Molecular Phylogenetics and Evolution*, **69**, 994–1004.

Merckx, V.S.F.T., Schols, P., Maas-van de Kamer, H., Maas, P., Huysmans, S. & Smets, E.F. (2006) Phylogeny and evolution of Burmanniaceae (Dioscoreales) based on nuclear and mitochondrial data. *American Journal of Botany*, **93**, 1684–1698.

Merckx, V.S.F.T., Chatrou, L.W., Lemaire, B., Sainge, M.N., Huysmans, S. & Smets, E.F. (2008) Diversification of myco-heterotrophic angiosperms: evidence from Burmanniaceae. *BMC Evolutionary Biology*, **8**, 178.

Merckx, V.S.F.T., Bakker, F.T., Huysmans, S. & Smets, E.F. (2009) Bias and conflict in phylogenetic inference of myco-heterotrophic plants: a case study in Thismiaceae. *Cladistics*, **25**, 64–77.

Merckx, V.S.F.T., Smets, E.F. & Specht, C.D. (2013a) Biogeography and conservation. *Mycoheterotrophy: the biology of plants living on fungi* (ed. by V.S.F.T. Merckx), pp. 103–156. Springer, New York.

Merckx, V.S.F.T., Kissling, J., Hentrich, H., Janssens, S.B., Mennes, C.B., Specht, C.D. & Smets, E.F. (2013b) Phylogenetic relationships of the mycoheterotrophic genus *Voyria* and the implications for the biogeographic history of Gentianaceae. *American Journal of Botany*, **100**, 712–721.

Mitchell, A., Li, R., Brown, J.W., Schönberger, I. & Wen, J. (2012) Ancient divergence and biogeography of *Raukaua* (Araliaceae) and close relatives in the southern hemisphere. *Australian Systematic Botany*, **25**, 432–446.

Neinhuis, C. & Ibisch, P.L. (1998) Corsiaceae. *The families and genera of vascular plants, Vol. 3, Lilianae (except Orchidaceae)* (ed. by K. Kubitzki), pp. 198–201. Springer, Berlin.

Neyland, R. & Hennigan, M. (2003) A phylogenetic analysis of large-subunit (26S) ribosome DNA sequences suggests that the Corsiaceae are polyphyletic. *New Zealand Journal of Botany*, **41**, 1–11.

Olmstead, R.G. & Palmer, J.D. (1994) Chloroplast DNA systematics: a review of methods and data analysis. *American Journal of Botany*, **81**, 1205–1224.

Petersen, G., Seberg, O. & Davis, J.I. (2013) Phylogeny of the Liliales (Monocotyledons) with special emphasis on data partition congruence and RNA editing. *Cladistics*, **29**, 274–295.

Prasad, V., Strömberg, C.A.E., Leaché, A.D., Samant, B., Patnaik, R., Tang, L., Mohabey, D.M., Ge, S. & Sahni, A. (2011) Late Cretaceous origin of the rice tribe provides evidence for early diversification in Poaceae. *Nature Communications*, **2**, 480.

Ramirez, S.R., Gravendeel, B., Singer, R.B., Marshall, C.R. & Pierce, N.E. (2007) Dating the origin of the Orchidaceae from a fossil orchid with its pollinator. *Nature*, **448**, 1042–1045.

Raven, P.H. & Axelrod, D.I. (1972) Plate tectonics and Australasian paleobiogeography. *Science*, **176**, 1379–1386.

Remizowa, M.V., Sokoloff, D.D. & Rudall, P.J. (2010) Evolutionary history of the monocot flower. *Annals of the Missouri Botanical Garden*, **97**, 617–645.

van Royen, P. (1972) Sertulum Papuanum 17: Corsiaceae of New Guinea and surrounding areas. *Webbia*, **27**, 223–255.

Rübsamen, T. (1986) *Morphologische, embryologische und systematische Untersuchungen an Burmanniaceae und Corsiaceae (mit Ausblick auf die Orchidaceae-Apostasioideae)*. Dissertationes Botanicae 92. J. Cramer, Berlin.

Rudall, P.J. (2002) Homologies of inferior ovaries and septal nectaries in monocotyledons. *International Journal of Plant Sciences*, **163**, 261–276.

Rudall, P.J. & Bateman, R.M. (2002) Roles of synorganisation, zygomorphy and heterotopy in floral evolution: the gynostemium and labellum of orchids and other lilioid monocots. *Biological Reviews*, **77**, 403–441.

Rudall, P.J. & Eastman, A. (2002) The questionable affinities of *Corsia* (Corsiaceae): evidence from floral anatomy and pollen morphology. *Botanical Journal of the Linnean Society*, **138**, 315–324.

Rudall, P.J., Stobart, K.L., Hong, W.P., Conran, J.G., Furness, C.A., Kite, G. & Chase, M.W. (2000) Consider the lilies: systematics of Liliales. *Proceedings of the Second International Conference on the Comparative Biology of the Monocots, Sydney, Vol. 1, Monocots: systematics and evolution* (ed. by K.L. Wilson and D.A. Morrison), pp. 347–359. CSIRO, Melbourne.

Sanmartín, I. & Ronquist, F. (2004) Southern Hemisphere biogeography inferred by event-based models: plant versus animal patterns. *Systematic Biology*, **53**, 216–243.

Schuster, T.M., Setaro, S.D. & Kron, K.A. (2013) Age estimates for the buckwheat family Polygonaceae based on sequence data calibrated by fossils and with a focus on the Amphi-Pacific *Muehlenbeckia*. *PLoS ONE*, **8**, e61261.

Scotese, C.R., Gahagan, L.M. & Larson, R.L. (1988) Plate tectonic reconstructions of the Cretaceous and Cenozoic ocean basins. *Tectonophysics*, **155**, 27–48.

Smets, E.F., Ronse Decraene, L.P., Caris, P. & Rudall, P.J. (2000) Floral nectaries in monocotyledons: distribution and evolution. *Proceedings of the Second International Conference on the Comparative Biology of the Monocots, Sydney, Vol. 1, Monocots: systematics and evolution* (ed. by K.L. Wilson and D.A. Morrison), pp. 230–240. CSIRO, Melbourne.

Sohl, N.F. & Owens, J.P. (1991) Cretaceous stratigraphy of the Carolina coastal plain. *The geology of the Carolinas: Carolina Geological Society fiftieth anniversary volume* (ed. by J.W. Horton and V.A. Zullo), pp. 191–220. University of Tennessee Press, Knoxville, TN.

Soltis, D.E., Smith, S.A., Cellinese, N. *et al.* (2011) Angiosperm phylogeny: 17 genes, 640 taxa. *American Journal of Botany*, **98**, 704–730.

Stevenson, D.W. & Loconte, H. (1995) Cladistic analysis of monocot families. *Monocotyledons: systematics and evolution* (ed. by P.J. Rudall, P.J. Cribb, D.F. Cutler and C.J. Humphries), pp. 543–578. Kew, Royal Botanic Gardens.

Swenson, U., Nylinder, S. & Munzinger, J. (2014) Sapotaceae biogeography supports New Caledonia being an old Darwinian island. *Journal of Biogeography*, **41**, 797–809.

Thorne, R.F. (1983) Proposed new realignments in the angiosperms. *Nordic Journal of Botany*, **3**, 85–117.

Veevers, J.J., Powell, C.M. & Roots, S.R. (1991) Review of sea floor spreading around Australia. I. Synthesis of the patterns of spreading. *Australian Journal of Earth Sciences*, **38**, 373–389.

Vinnersten, A. & Bremer, K. (2001) Age and biogeography of major clades in Liliales. *American Journal of Botany*, **88**, 1695–1703.

Wicke, S., Schneeweiss, G.M., dePamphilis, C.W., Müller, K.F. & Quandt, D. (2011) The evolution of the plastid chromosome in land plants: gene content, gene order, gene function. *Plant Molecular Biology*, **76**, 273–297.

Woodburne, M.O. & Case, J.A. (1996) Dispersal, vicariance, and the Late Cretaceous to early Tertiary land mammal biogeography from South America to Australia. *Journal of Mammalian Evolution*, **3**, 121–161.

Zhang, D., Saunders, R.M.K. & Hu, C.-M. (1999) *Corsiopsis chinensis* gen. et sp. nov. (Corsiaceae): first record of the family in Asia. *Systematic Botany*, **24**, 311–314.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article:

Appendix S1 Supplementary tables (Tables S1 & S2).
Appendix S2 Supplementary materials and methods.
Appendix S3 Supplementary figures (Figs S1–S8).

BIOSKETCH

Constantijn Mennes is a PhD student in Systematic Botany at Naturalis Biodiversity Center and Leiden University, Leiden, The Netherlands. His research focuses on the phylogeny, biogeography and evolutionary history of mycoheterotrophic plants, with an emphasis on tropical taxa. Recent projects involve the plant families Triuridaceae, Polygalaceae and Gentianaceae.

Author contributions: C.M., V.M. and E.S. initiated the research; C.M., P.R. and V.M. conceived the ideas; C.M., V.L. and S.L. acquired the data; C.M., V.L. and S.G. analysed the data; C.M. led the writing; C.M., V.L. and P.R. contributed to the writing; all authors discussed the ideas and commented on the manuscript.

Editor: Malte Ebach