

7-1-2015

## Phylogenetics, divergence times and diversification from three genomic partitions in monocots

Kate L. Hertweck  
*National Evolutionary Synthesis Center*

Michael S. Kinney  
*University of La Verne*

Stephanie A. Stuart  
*Macquarie University*

Olivier Maurin  
*University of Johannesburg*

Sarah Mathews  
*Harvard University*

*See next page for additional authors*

Follow this and additional works at: [https://digitalcommons.lsu.edu/biosci\\_pubs](https://digitalcommons.lsu.edu/biosci_pubs)

---

### Recommended Citation

Hertweck, K., Kinney, M., Stuart, S., Maurin, O., Mathews, S., Chase, M., Gandolfo, M., & Pires, J. (2015). Phylogenetics, divergence times and diversification from three genomic partitions in monocots. *Botanical Journal of the Linnean Society*, 178 (3), 375-393. <https://doi.org/10.1111/boj.12260>

This Article is brought to you for free and open access by the Department of Biological Sciences at LSU Digital Commons. It has been accepted for inclusion in Faculty Publications by an authorized administrator of LSU Digital Commons. For more information, please contact [ir@lsu.edu](mailto:ir@lsu.edu).

---

**Authors**

Kate L. Hertweck, Michael S. Kinney, Stephanie A. Stuart, Olivier Maurin, Sarah Mathews, Mark W. Chase, Maria A. Gandolfo, and J. Chris Pires



## Phylogenetics, divergence times and diversification from three genomic partitions in monocots

KATE L. HERTWECK<sup>1,2\*</sup>, MICHAEL S. KINNEY<sup>3</sup>, STEPHANIE A. STUART<sup>4</sup>, OLIVIER MAURIN<sup>5</sup>, SARAH MATHEWS<sup>6</sup>, MARK W. CHASE<sup>7</sup>, MARIA A. GANDOLFO<sup>8</sup> and J. CHRIS PIRES<sup>3</sup>

<sup>1</sup>National Evolutionary Synthesis Center (NESCent), 2024 West Main Street Suite A200, Durham, NC 27705, USA

<sup>2</sup>Division of Biological Sciences, 311 Bond Life Sciences Center, University of Missouri, 1201 Rollins Street, Columbia, MO 65211, USA

<sup>3</sup>Biology Department, University of La Verne, 1950 Third Street, La Verne, CA 91750, USA

<sup>4</sup>Department of Biological Sciences, Faculty of Science, Macquarie University, Sydney, NSW 2109, Australia

<sup>5</sup>Molecular Systematics Laboratory, African Centre for DNA barcoding (ACDB), Department of Botany and Plant Biotechnology, University of Johannesburg, APK Campus, PO Box 524, Auckland Park, 2006, South Africa

<sup>6</sup>Harvard University Herbaria, 22 Divinity Avenue, Cambridge, MA 02138, USA

<sup>7</sup>Jodrell Laboratory, Royal Botanic Gardens, Kew, Richmond, Surrey TW9 3DS, UK

<sup>8</sup>LH Bailey Hortorium, Plant Biology, School of Integrative Plant Science, Cornell University, 410 Mann Library Building, Ithaca, NY 14853, USA

Received 14 February 2014; revised 6 June 2014; accepted for publication 6 January 2015

Resolution of evolutionary relationships among some monocot orders remains problematic despite the application of various taxon and molecular locus sampling strategies. In this study we sequenced and analysed a fragment of the low-copy, nuclear phytochrome C (*PHYC*) gene and combined these data with a previous multigene data set (four plastid, one mitochondrial, two nuclear ribosomal loci) to determine if adding this marker improved resolution and support of relationships among major lineages of monocots. Our results indicate the addition of *PHYC* to the multigene dataset increases support along the backbone of the monocot tree, although relationships among orders of commelinids remain elusive. We also estimated divergence times in monocots by applying newly evaluated fossil calibrations to our resolved phylogenetic tree. Inclusion of early-diverging angiosperm lineages confirmed the origin of extant monocots c. 131 Mya and strengthened the hypothesis of recent divergence times for some lineages, although current divergence time estimation methods may inadequately model rate heterogeneity in monocots. We note significant shifts in diversification in at least two monocot orders, Poales and Asparagales. We describe patterns of diversification in the context of radiation of other relevant plant and animal lineages. © 2015 The Linnean Society of London, *Botanical Journal of the Linnean Society*, 2015, **178**, 375–393.

**ADDITIONAL KEYWORDS:** divergence time estimation – fossil calibration – molecular phylogenetics – monocotyledoneous plants.

\*Corresponding author. Current address: University of Texas at Tyler, 3900 University Blvd, Tyler, TX 75799, USA. E-mail: khertweck@uttyler.edu



**Table 1.** Divergence times for the SL and CG of major monocot lineages

Lineage	Janssen & Bremer, 2004	Anderson & Janssen, 2009	Magallón & Castillo, 2009	Bell <i>et al.</i> , 2010 exponential, lognormal	This study
Data utilized	One gene (plastid DNA) NPRS (r8s)	One gene (plastid DNA) PL (r8s)	Five genes (plastid and nrDNA) Relaxed clock (BEAST)	Three genes (plastid and nrDNA) Relaxed clock (BEAST)	Eight genes (plastid, mitochondrial, nrDNA, <i>PHYC</i> ) PL (r8s) root 160 (140–180)
Dating method (program)					Relaxed clock (multidivtime) mean (95% CI)
Monocots	CG 134*	CG 134*	127	130,146	138 (132–143)
Commelinids	SL N/A	SL N/A	N/A	136,156	147 (142–152)
Acorales	CG 120	CG 120	115	96,103	<i>110 (104–115)</i>
Alismatales	SL 122	SL 122	119	101,107	<i>115 (110–121)</i>
Arecales	CG N/A	CG N/A	N/A	N/A	30 (15–50) <sup>^</sup>
Asparagales	SL 134*	SL 134*	127	130,146	138 (132–143)
Commelinales	CG 128	CG 128	127	107,122	125 (119–131)
Dasygongonaceae	SL 131	SL 131	126	118,136	132 (127–138)
Dioscoreales	CG 110	CG N/A	N/A	31, 33	81 (72–90) <sup>^</sup>
Liliales	SL 120	SL 120	115	96,103	109 (103–114)
Pandanales	CG 119	CG 118	113	92,103	112 (106–118)
Petrosaviales	SL 123	SL 122	119	105,118	<i>115 (110–121)</i>
Poales	CG 124	CG 116	114	70, 76	95 (88–101)
Zingiberales	SL 117	SL 116	111	86, 88	101 (94–107)
	CG 88	CG 78	80	38, 42	68 (56–78) <sup>^</sup>
	SL 114	SL 114	100	96,103	109 (103–114)
	CG 100	CG N/A	N/A	83, 92	110 (104–117)
	SL 119	SL 123	115	105,118	116 (110–122)
	CG 123	CG 124	120	86,114	<i>105 (98–112)<sup>^</sup></i>
	SL 126	SL 116	114	105,118	117 (111–123)
	CG 113	CG 112	109,124	72, 82	<i>103 (95–111)</i>
	SL 117	SL 116	111	105,118	116 (110–122)
	CG 88	CG 78	80	89, 99	<i>108 (99–115)<sup>^</sup></i>
	SL 114	SL 114	100	60 (57–64)	122 (116–127)
	CG 114	CG 114	100	85, 93	99 (94–105) <sup>^</sup>
	SL 114	SL 116	111	96,103	<i>109 (103–114)</i>
	CG 88	CG 78	80	84, 84	76 (65–85) <sup>^</sup>
	SL 114	SL 114	100	86, 88	101 (94–107)

SL, stem lineage; CG, crown group. An asterisk (\*) indicates placement at the root and a fixed date for that node. N/A indicates the date for that node was not reported or was not estimated because of taxonomic sampling. Italics in multidivtime indicates mean dates that fall outside the confidence intervals for the r8s analysis; carets (^) indicate the r8s and multidivtime confidence intervals do not overlap.

mitochondrial) and high-copy nuclear ribosomal (nrDNA) loci (Janssen & Bremer, 2004; Anderson & Janssen, 2009; Magallón & Castillo, 2009). Low-copy nuclear genes provide unlinked loci with which to evaluate separately phylogenetic hypotheses derived primarily from uniparentally inherited and linked plastid markers, which may have alternative evolutionary histories. Moreover, the combination of low-copy nuclear loci with plastid, mitochondrial and high-copy nuclear loci provides a robust dataset with which both to evaluate phylogenetic relationships and to estimate divergence times (Parfrey *et al.*, 2011).

Low-copy nuclear phytochrome genes have been effective in resolving phylogenetic relationships across angiosperms (e.g. Mathews & Sharrock, 1996; Mathews & Donoghue, 1999, 2000; Bennett & Mathews, 2006; McNeal *et al.*, 2013). Nuclear phytochrome genes, a family of red and far/red light-sensing proteins, are well characterized in several angiosperm species and comprise a small number of genes that are evolving independently in angiosperms (Mathews, Lavin & Sharrock, 1995). Phytochrome C is a member of the gene family which diverged prior to the diversification of angiosperms and appears to occur in single copy in monocots (Mathews & Donoghue, 2000).

We sequenced and analysed a small fragment from exon I of the nuclear-encoded *PHYC* gene for most monocot and several outgroup families. *PHYC* data were combined with the multigene data set of Chase *et al.* (2006) to determine if adding this marker improved resolution and support of relationships among the major lineages of monocots, particularly at previously unresolved or weakly supported nodes. We also estimated divergence times by applying new, robust fossil calibrations to a resolved phylogenetic tree calculated from the multi-locus dataset representing all three plant genomes, including *PHYC*. Here, we present refined estimates for the age of monocots and major lineages within them. We utilize three methods to evaluate monocot diversification and interpret resulting patterns in the context of the radiation of other relevant plant and animal lineages.

## MATERIAL AND METHODS

### TAXON SAMPLING

Taxon sampling followed the multilocus data sets of Chase *et al.* (1995, 2000, 2006). These data included 124 species representing all 11 orders of monocots and Dasypogonaceae (Givnish *et al.*, 2006); all extant monocot families circumscribed by APG III (2009) were represented except for three families in Alismatales (Ruppiaceae, Posidoniaceae and Scheuchzeriaceae). Outgroup taxa included the 17 taxa representing

early-diverging angiosperm lineages (Mathews & Donoghue, 1999; Qiu *et al.*, 1999; APG III, 2009) from the Chase datasets (e.g. Chase *et al.*, 2006) and ten additional eudicot taxa (to improve placement of fossil calibrations). Taxon names (and substitutions), voucher information and accession numbers are provided in Supporting Information File S1. Tip labels in all trees correspond to the taxon name from Chase *et al.* (2006).

### DNA EXTRACTION, PCR, CLONING AND SEQUENCING

In most cases the DNA used for amplification was the same as was used in previous molecular phylogenetic studies of the monocots (File S1; Chase *et al.*, 1995, 2000, 2006). Other taxa represented the same genus or family when DNA accessions were unavailable and/or did not amplify; estimations of familial relationships using similar procedures have shown that such substitutions have not had adverse effects on phylogenetic studies at higher taxonomic levels as these families are monophyletic (Qiu *et al.*, 1999; Soltis *et al.*, 2000). Genomic DNA was extracted from fresh or silica-dried leaf material of replacement samples using a modified CTAB procedure (Doyle & Doyle, 1987) using 3–6× CTAB and 2 M NaCl (Smith *et al.*, 1991). For most specimens an approximately 1.2-kb region in exon 1 of the nuclear-encoded *PHYC* gene was amplified using primers c230f and c623r (Mathews & Donoghue, 1999, 2000), and some PCR products produced clean sequences without cloning.

For taxa that did not amplify using this protocol, additional primers were designed manually based on the original primers but made less degenerate for specific monocot orders (File S2). Amplification with the newly designed primers used the Qiagen *Taq* DNA polymerase system in the following 50- $\mu$ L reaction mixture: template DNA ~100 ng, 2  $\mu$ L each primer at 10  $\mu$ M, 5  $\mu$ L 10× Qiagen PCR Buffer (with 15 mM MgCl<sub>2</sub>), an additional 2  $\mu$ L 25 mM MgCl<sub>2</sub>, 4  $\mu$ L 2.5 mM each dNTP and 0.4  $\mu$ L Qiagen *Taq* (5 U  $\mu$ L<sup>-1</sup>). PCR reactions utilized the following conditions: an initial denaturing step of 94 °C for 5 min, 40 cycles at 94 °C for 1 min, 55 °C for 1 min and 72 °C for 90 s, and a final extension step of 72 °C for 20 min. All PCR products were visualized on a 1.5% agarose gel, and 1.2-kb bands were excised and purified, ligated into plasmid and cloned using the TOPO TA Cloning Kit (Invitrogen). We screened at least ten positive (white) colonies using PCR and M13F and M13R primers. The resulting products were purified prior to sequencing, and we obtained at least six complete cloned sequences per taxon.

### PHYLOGENETIC ANALYSIS OF *PHYC*

Trace files for all sequences were assembled into contigs using SeqMan Pro version 7.1.0 (DNASTAR).

Vector ends were identified and trimmed manually. The identity of edited *PHYC* sequences was verified by the presence of easily recognized amino acid sequence hallmarks and in phylogenetic analyses with data from the other *PHY* loci. We performed preliminary analyses of all *PHYC* clones from taxa in each monocot order. Nucleotide sequence alignments within orders were unambiguous and did not contain large insertions/deletions (gaps), and *PHYC* clones from the same taxon clustered together (data not shown).

The final dataset included one randomly chosen *PHYC* clone per taxon. Sequences were translated to amino acids using Mesquite (Maddison & Maddison, 2011) and aligned in MUSCLE (Edgar, 2004a, b); the resulting alignment was used to manually place gaps in the nucleotide matrix. The beginning and end of the matrix were trimmed to minimize missing data; uninformative or ambiguous gaps were excluded from subsequent analysis. Trees that maximized the likelihood of the data (ML trees) with *Amborella trichopoda* Baill. as the outgroup were inferred under the GTR+I+G model of molecular evolution using RAxML v.7.2.8 (Stamatakis, Hoover & Rougemont, 2008), which includes ten slow ML tree searches. See section below on 'Phylogenetic analysis of concatenated data' for best-fit model testing. Support for nodes was evaluated using 1000 rapid bootstrap replicates using the 'f a' option in RAxML. Bayesian phylogenetic trees (BI trees) were inferred in MrBayes v.3.2 (Ronquist *et al.*, 2012) using the GTR+I+G model of evolution; model selection is described below. We implemented default chain and heating parameters (two runs of four chains each); 30 million generations were sampled every 1000 generations with 20% discarded as burn-in. We used two diagnostic methods to evaluate convergence within and between runs, explained here only briefly. First, convergence of molecular evolutionary parameters was examined using Tracer v1.5 (Rambaut & Drummond, 2009); we required ESS (effective sample size) of > 200 for the two combined runs and examined overlapping marginal densities for the two runs. Second, convergence of topologies and posterior probabilities was examined using the program AWTY (Nylander *et al.*, 2008) by comparing patterns of splits between posterior probabilities for both runs. The final tree was visualized with FigTree v1.4 (Rambaut, 2012) and a custom R script implementing R packages ape v3-11 (Paradis, Claude & Strimmer, 2004) and phytools v0.3-72 (Revell, 2011).

#### PHYLOGENETIC ANALYSIS OF CONCATENATED DATA

The *PHYC* data set described above was combined for analysis with the previous seven-gene data set of Chase *et al.* (2006), which includes data from four

plastid loci (*atpB*, *matK*, *ndhF*, *rbcL*), one mitochondrial locus (*atpA*) and two nuclear ribosomal loci (18S and 26S). We added sequences made available on GenBank since the formation of this seven-gene alignment to fill as much missing data as possible (Table 1) compared with previous analyses (Chase *et al.*, 2006). We excluded all characters that were excluded from the final analyses in Chase *et al.* (2006). We used PartitionFinder (Lanfear *et al.*, 2012) to partition our data. Given an alignment of pre-defined data blocks (in this case, the eight-gene dataset), this program selects both the best-fitting partition scheme and models of molecular evolution for each partition. The combined eight-gene dataset was separated by PartitionFinder into three partitions (1, *atpA* and 18S/26S; 2, all four plastid DNA genes; 3, *PHYC*), thus reducing model complexity compared with partitioning by gene.

ML analysis of the combined dataset was conducted as described for the *PHYC* dataset except GTR+I+G was applied to each of the three partitions described above. In addition to this unconstrained analysis, we performed the same ML reconstruction with outgroup topology constrained to the current best estimate of relationships (Moore *et al.*, 2007; Soltis & Soltis, 2013) to improve placement of fossil taxa for divergence time analyses (see below). BI analysis for the combined eight-gene dataset also followed the same parameters described for *PHYC*, except partitioned as determined by PartitionFinder (1, SYM+I+G; 2, GTR+I+G; 3, GTR+I+G). The combined BI analysis continued for 20 million generations with 20% discarded as burn-in.

#### DIVERGENCE TIMES AND DIVERSIFICATION

Fossils were selected from within monocots and the outgroup taxa to constrain divergence time estimates, following the recommendations of Gandolfo, Nixon & Crepet (2008) and Parham *et al.* (2011). The fossil record in monocots (excluding commelinids) was thoroughly discussed by Smith (2013). Crown group (CG) refers to the clade originating at the most recent common ancestor (MRCA) of all extant members of a group, whereas stem lineage (SL) includes both extant and extinct members of a group. Stratigraphic positions of fossils for constraints were transformed to minimum ages using the younger bound of the interval based on the current stratigraphic timescale (Gradstein, Ogg & Schmitz, 2012). Justification for fossils selected for inclusion is provided below and in File S3.

#### Calibration 1

Two fossils (*Monetianthus mirus* E.M.Friis *et al.*, *Carpestella lacunata*) constrained the MRCA of Nym-

phaeales and Illiciales; the placement of this calibration reflects characteristics shared by both lineages. *Monetianthus mirus*, a probably bisexual flower (Friis, Pedersen & Crane, 2001; Friis *et al.*, 2009), was originally published as possessing affinities to Nymphaeales (Friis *et al.*, 2001). Later coding of characters of this fossil for phylogenetic analysis confirmed this placement (Friis *et al.*, 2009). Crepet *et al.* (2004) raised doubts about the inclusion of *M. mirus* in Nymphaeales due to incomplete preservation of the fossil and absence of definitive synapomorphies of Nymphaeaceae. Supporting this fact are the results of a phylogenetic analysis in which *M. mirus* was analysed in combination with another fossil assigned to Nymphaeales, *Microvictoria svitkoana*, and all extant taxa of Nymphaeaceae (Gandolfo, Nixon & Crepet, 2004). This analysis suggested *M. mirus* could alternatively belong to Illiciales. The other fossil, *C. lacunata* (von Balthazar *et al.*, 2008), also possessed an apparently bisexual flower and was analysed using a previously published matrix (Saarela *et al.*, 2007). The lack of resolution in the resulting strict consensus tree left the position of the fossil unresolved, although it shared characters with both Nymphaeaceae and Illiciaceae.

#### Calibration 2

'Unequivocal' similarity to the extant genus *Illicium* L. placed *Illiciospermum pusillum* Frumin & E.M.Friis (Frumin & Friis, 1999) at the SL of Schisandraceae.

#### Calibration 3

Dispersed pollen of *Clavatipollenites minutus* (Brenner & Bickoff, 1992) possesses similarities in morphology, sculpture, and ultrastructure of the wall to pollen produced by modern *Ascarina* Forst. (Chloranthaceae, Walker & Walker, 1984). Several pollen grains show wall structure and other morphologies which make relationships with extant species of the genus uncertain. Given the difficulty in circumscription, uncertain systematic placement and taxonomic sampling of the family, *C. minutus* represents the SL of Chloranthaceae.

#### Calibration 4

*Endressinia brasiliana* (Mohr & Bernardes-de-Oliveira, 2004) was placed at the SL of Magnoliales based on morphological similarities to extant Magnoliales. No affinities for *E. brasiliana* at the family level were established, and it is possible this taxon represents an extinct family.

#### Calibration 5

*Lactoripollenites africanus* (Zavada & Benson, 1987) is placed at the SL of Lactoridaceae; pollen charac-

teristics supporting this relationship include shape, size and ultrastructure characters of the pollen wall.

#### Calibration 6

Phylogenetic analysis of pollen characters (Doyle & Endress, 2010) placed *Spanomera mauldinensis* and *Spanomera marylandensis* (Drinnan *et al.*, 1991) with Buxaceae. We place this constraint at the CG of eudicots because of limited sampling in this portion of the outgroup.

#### Calibration 7

Two fossils constrain the SL of Araceae: *Cobbania corrugata* (Lesquereux, 1876; Stockey, Rothwell & Johnson, 2007) because of overall morphology of the fossil plants, and *Cobbanicarpites amurensis* (Krassilov & Kodrul, 2009) because of fruit morphology and features of the seed coat.

#### Calibration 8

*Mabelia* and *Nuhliantha* (Gandolfo, Nixon & Crepet, 2002) are placed to constrain the CG of Pandanales because of general morphological and anatomical features of preserved flowers as well as pollen *in situ*. Although there is some debate as to the placement of these fossils (Friis, Pedersen & Crane, 2006), phylogenetic analysis nested the taxa in tribe Triurideae (Gandolfo *et al.*, 2002).

#### Calibration 9

*Sabalites magothiensis* and *Palmoxylon cliffwoodensis* (Berry, 1905, 1911, 1916; Daghlian, 1981) constrained the SL of Areaceae (Arecales) because of the general venation pattern in preserved leaves (calibration 9). Although we do include a species from the sister group of the rest of the palms, Calamoideae (Asmussen *et al.*, 2006), sparse taxonomic sampling in this diverse order resulted in our conservative placement of the calibrations at the SL of Arecales.

#### Calibration 10

*Tricostatocarpon silvapinedae* and *Striatornata san-antoniensis* (Rodriguez-de la Rosa & Cevallos-Ferriz, 1994) constrained the SL of Zingiberales based on general external morphology, sculpture and anatomical details. These fossils are associated with an absolute age used for dating the constraint, although *Spiromatospermum chandlerae* (Friis, 1988) occurred in the same time period and was used for previous fossil-based divergence time analysis (Bremer, 2000).

#### Calibration 11

The morphology of phytoliths of Poaceae (Prasad *et al.*, 2005) constrains the SL of Poaceae.

We inferred divergence times using a semiparametric method implemented in r8s v1.70 (Sanderson,

2003) using PL (Sanderson, 2002), the TN algorithm with bound constraints, three initial starts, fossil-based cross validation (Near & Sanderson, 2004) and the combined eight-gene constrained ML tree. The root node representing the CG of angiosperms was fixed at 160 Mya. A test for the application of a molecular clock failed, validating the use of relaxed molecular clock approaches. An optimal smoothing parameter of 3200 was selected by testing values from  $\log \lambda_{10} = 0-3.5$  at intervals of 0.5. We obtained confidence intervals for the PL analysis by performing the same calculations with the early (140 Ma) and late (180 Ma) bounds of the current angiosperm age estimates. See Bell *et al.* (2010) and Magallón (2009) for a complete discussion of current dating of CG angiosperms.

We also estimated divergence times using the 9/25/03 distribution of multidivtime (Thorne & Kishino, 2002), following recommendations of Rutschmann (2005). multidivtime is a Bayesian relaxed clock dating method which uses the paml package (Yang, 2007) to apply substitution models to multiple data partitions and infer a posterior distribution of evolutionary times and rates. We estimated molecular evolution parameters under F84+G, which is the most complex model implemented in multidivtree. We applied the same fossil calibrations described above as hard bounds, including the root node constrained to 160 Mya or younger. Time units for setting priors were in hundreds of millions of years; for example, the mean prior distribution of the root node (rttm) was set to 1.6. The prior distributions for the rate of molecular evolution at the ingroup root node (rtrate) and autocorrelation parameter (brownmean) were estimated from the branch length estimates as recommended by multidivtime documentation and set at 0.05 and 0.7, respectively. Standard deviations for each parameter (rttmsd, rtratesd, brownsd) were set equal to the mean, which is standard practice. We first ran the program under the prior to check that distributions for nodes of interest were sufficiently wide, as recommended by the program documentation. We executed two independent runs with the identical parameters, sampling every 100 cycles, discarding the first 100 000 as burn-in and collecting the following 10 000 samples.

For diversification analyses, we gathered species diversity data from across monocots and tested modifications to the tree to accommodate incomplete sampling. Our analysis samples exemplars from major monocot lineages; each sampled taxon represents from few to many thousands of species. To accommodate this sparse sampling, diversification analyses excluded outgroups (eudicots and early-diverging angiosperm lineages), and monocot taxa were pruned such that only a single tip per monocot family

remained. Using the specified time-calibrated tree, we excluded all outgroups and trimmed all taxa except a single exemplar per monocot family *sensu* APG III (2009). Species counts were obtained from the Angiosperm Phylogeny Website (APW; Stevens, 2001 onwards); if a range of total species was given, we applied the upper bound. Species counts from unsampled families in Alismatales (Ruppiaceae, Posidoniaceae and Scheuchzeriaceae) were added to the closest sampled family, Cymodoceaceae. When APG III lumped families described in APW, totals from these families were added (e.g. Thismiaceae and Burmanniaceae). Species counts by family, including the taxa removed for the family analysis, can be found in File S4.

We used multiple methods to evaluate diversification in monocots. First, we constructed lineage-through-time (LTT; Nee, Mooers & Harvey, 1992) plots in R using ape v3-11 (Paradis *et al.*, 2004) to visualize the rate of diversification across the tree by comparing the shape of plots inferred from different divergence time estimates and both pruned and complete trees. Second, we used apTreeshape v1.4-5 (Bortolussi *et al.*, 2005) as implemented in R to test for shifts in diversification in the family-level r8s tree. apTreeshape is a topology-based test that evaluates tree shape in the context of diversification using the  $\Delta_1$  statistic (Moore, Chan & Donoghue, 2004). Finally, because our tree does not completely sample all monocots, we implemented MEDUSA from the R package geiger v1.99-3.1 (Harmon *et al.*, 2008). This method incorporates both divergence times and species richness by family to fit diversification models in a likelihood framework.

## RESULTS

All trees and the combined data matrix are available on TreeBase (accession number 15722); GenBank accession numbers for all included sequences can be found in File S1.

### PHYC ANALYSIS

The final version of the *PHYC* data set used in this study included 1248 bp (371 aligned amino acids) of exon 1 from 132 taxa; 135 bp representing ambiguous or uninformative gaps were excluded from phylogenetic analysis. In the resulting matrix 81.5% of the positions were variable; 2.86% of the aligned positions included gaps. There were no well-supported differences in topology of the monocots between the ML and BI trees. Although most orders were supported, there was little support for relationships among major clades (Fig. 2). The earliest diverging clades in



Dioscoreales (Nartheciaceae) and Asparagales (Orchidaceae) are not included in their assigned orders, and other relationships for these families were not strongly supported.

#### COMBINED EIGHT-GENE DATA SET AND ANALYSES

The data set that includes the seven loci from Chase *et al.* (2006) and *PHYC* data included 151 taxa, with an aligned length of 11 459 bp, of which 61.1% were variable; 2.9% of positions were missing. The constrained ML, unconstrained ML and BI trees differed primarily in the placement of non-monocot lineages, but were otherwise similar topologically (see below regarding Liliales, Fig. 3). Following Chase *et al.* (2006), areas of conflict for circumscription of crown groups and relationships between major lineages (11 orders and Dasygogonaceae) are highlighted. ML bootstrap scores (BS) > 70 and BI posterior probability (PP) > 70 are reported, with BS > 90 and PP > 95 defined as 'strongly supported'.

Analysis of the combined data set (Fig. 3) resulted in monophyly of the monocots with Acorales as sister to the remaining monocots (BS = 100, PP = 100); monophyly of this monogeneric order is also strongly supported (BS = 100, PP = 100). Placement of Alismatales as the next branching lineage above Acorales is strongly supported (BS = 100, PP = 100) as is the monophyly of this order (BS = 100, PP = 100). Monophyly of Petrosaviales and their position as the next branching lineage after Alismatales are strongly supported (BS = 100, PP = 100 for both). Monophyly of Dioscoreales is strongly supported (BS = 95, PP = 99), and includes Nartheciaceae (unlike the analyses of *PHYC* alone). Monophyly of Pandanales is also weakly to strongly supported (BS = 97, PP = 83). Although taxonomically assigned to Liliales, *Arachnitis* Phil. is grouped with Pandanales in the BI analysis, albeit on a long branch. The ML tree supports the sister relationship of Dioscoreales + Pandanales (BS = 91), the position of Liliales as the next branching lineage above Dioscoreales + Pandanales, (BS = 93), and monophyly of Liliales (BS = 96). Support for the placement of Asparagales as the next branching lineage above Liliales and sister to the commelinids is only moderate (BS = 85, PP = 96); Asparagales, including Orchidaceae, are monophyletic (BS = 99, PP = 99). The commelinids are strongly supported as monophyletic (BS = 98, PP = 97), as are the five lineages that comprise the group (Arecales: BS = 100, PP = 100; Commelinales: BS = 100, PP = 100; Dasygogonaceae: BS = 100, PP = 100; Poales: BS = 100, PP = 100; Zingiberales: BS = 99, PP = 100). The sister relationship of Commelinales + Zingiberales is strongly supported (BS = 99, PP = 99), but the relative placement of the remaining commelinid clades is less clear.

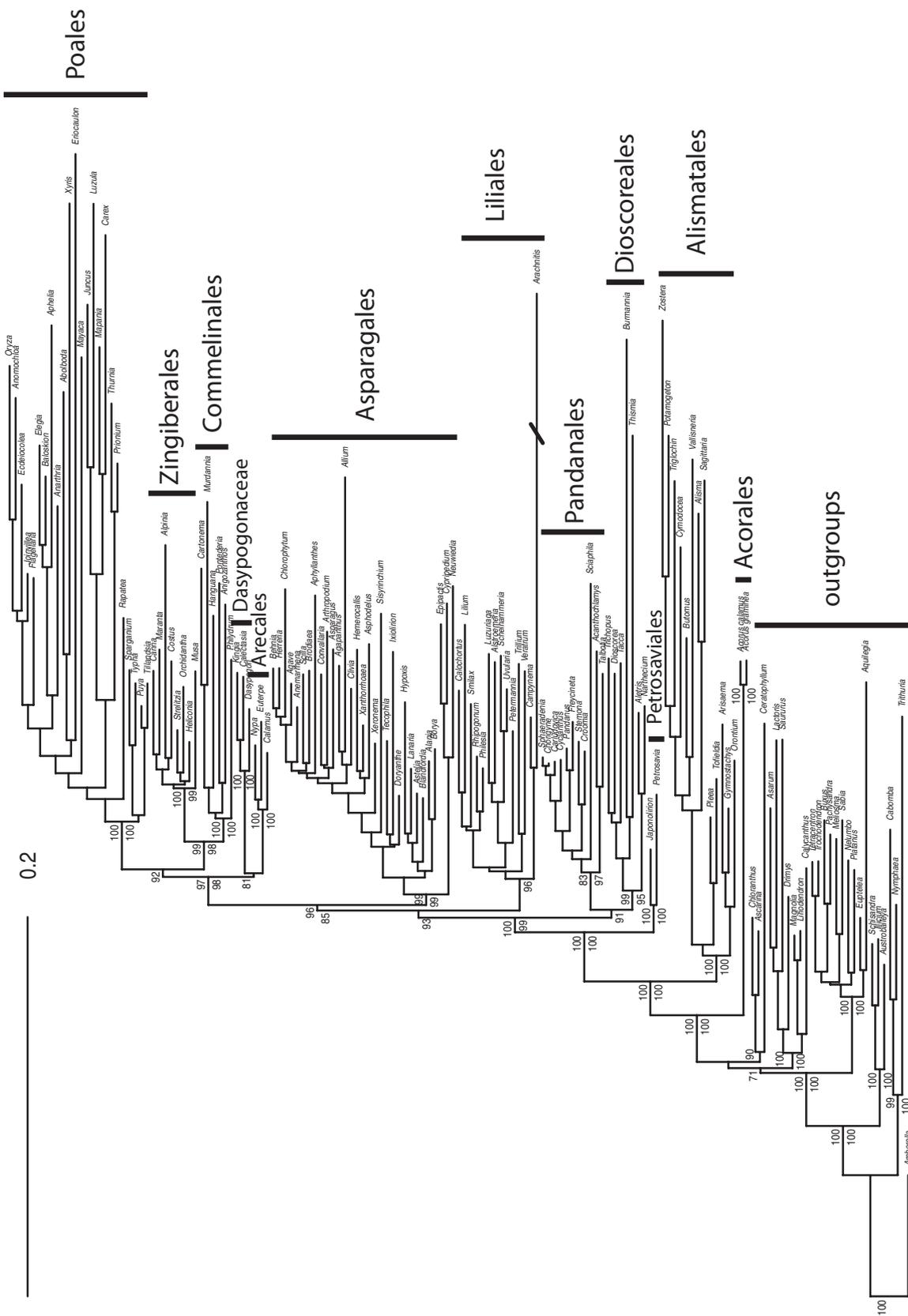
#### DIVERGENCE TIMES AND DIVERSIFICATION

Divergence times for SLs and CGs of monocot clades are shown in Figure 4. Inclusion of earlier diverging taxa allowed us to date the divergence of monocots at 131 Mya (CG 136 Mya, Table 1). Our analyses indicate younger divergence times for several CGs, particularly Alismatales, Arecales, Zingiberales and Commelinales.

Results from the two multidivtime runs did not differ from each other by > 1 Myr per node, suggesting the analyses had reached convergence. Dates for SL divergence did not differ substantially between r8s and multidivtime analysis. multidivtime reported divergence times for seven monocot crown groups which differed significantly from the r8s analysis, in that the confidence intervals for the two algorithms do not overlap and the mean divergence times are > 10 Myr apart. Liliales and Poales are younger in multidivtime compared with r8s, whereas Acorales, Arecales, Dasygogonaceae, Petrosaviales and Zingiberales are older.

We examined the effects of these different age estimates on overall patterns of diversification in monocots using LTT plots. To accommodate the effects of sparse taxon sampling, we emphasized only early portions of evolution in these lineages (Blankers *et al.*, 2013) and assume families may serve as a proxy of important morphological, biogeographical and life history diversity. LTT plots represent diversification by visualizing the estimated time before present (*x* axis) against log of the number of lineages (*y* axis, Fig. 5). The resulting line is a species accumulation curve, which visualizes tree-wide net diversification rates (rate of speciation minus rate of extinction). Overall, the monocot curves (rate of lineage accumulation) increased rapidly before slowing down and then levelling off. The intensity of the slope of the curve increased when only family-level diversity was included from the r8s analysis (Fig. 5A), and the pattern is robust to the age of the root node in r8s (Fig. 5B). The intensity of the slope of the curve also increased when only family-level diversity was included from the multidivtime analysis (Fig. 5C). Finally, tree-wide patterns of lineage accumulation were similar for both r8s and multidivtime, although the latter analysis resulted in a more gradual increase in diversity and earlier asymptote (Fig. 5D).

apTreeshape identified one branch with a significant shift in diversification rate, at the root of the commelinids ( $P = 0.029$ , Fig. 6). Two additional branches possessed marginally significant shifts in diversification: the root of Asparagales ( $P = 0.0979$ ) and the common ancestor of Poales + Zingiberales + Commelinales ( $P = 0.0558$ ).



**Figure 3.** Phylogenetic relationships of monocots inferred from the combined eight-gene dataset. Tree shown is ML inferred using RAxML. Numbers at nodes correspond to ML bootstrap percentages (1000 replicates, only values > 70 are shown) and BI posterior probabilities (only values > 90 are shown). Support percentages are only shown for the tree backbone and monocot crown groups. The hash mark on the branch leading to *Arachnitis* indicates the branch length has been shortened by half for the purposes of visualization. See TreeBase accession 15722 for complete support values.

Stepwise AIC in MEDUSA for the r8s chronogram identified ten branches that exhibited shifts in diversification relative to the background rate (Fig. 6). These shifts were associated with clades possessing both increased and decreased diversity, and were placed in seven monocot orders. The same analysis conducted on the multidivtime chronogram resulted in eight branches with shifts in diversification, all of which were associated with clades possessing decreased diversity (Fig. 6).

## DISCUSSION

Our analyses that combined *PHYC* with the previously assessed plastid, mitochondrial and nuclear ribosomal data set increased support for some previously uncertain relationships, including the placement of Liliales and Asparagales. Dioscoreales (including Nartheciaceae) are supported as sister to Pandanales (Fig. 3). Relationships among some orders of commelinids remain ambiguous. Barrett *et al.* (2013) analysed all 83 plastome genes and found robust support for the same commelinid relationships reported here. However, even their deep plastome sampling was unable to reject alternative topologies, suggesting that multigene data sets including nuclear data are needed to further improve our understanding of monocot phylogeny. Overall, the improved support along the backbone provided a useful context to estimate divergence times and evaluate patterns of monocot diversification. We contend that variation in life history among monocot lineages is associated with patterns of rate heterogeneity which make divergence times particularly challenging to estimate, and resulting dating estimates should be interpreted with these limitations in mind. Regardless, significant shifts in diversification occur across the monocot tree, and may be related to close associations with other plant and animal lineages diversifying at the same time.

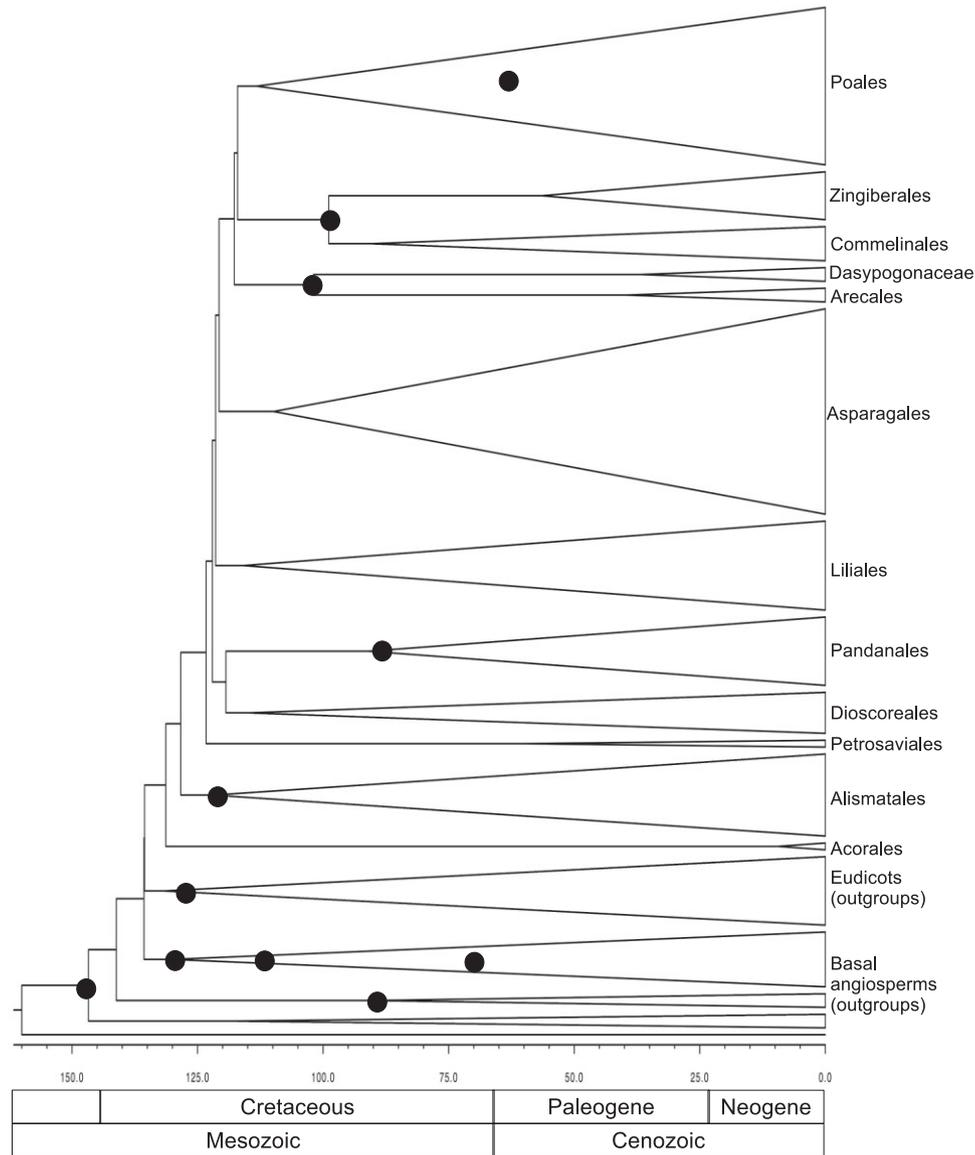
### MONOCOT DIVERSITY AND LIFE HISTORY VARIATION COMPLICATE DIVERGENCE DATING

The *c.* 60 000 species classified as monocots represent substantial diversity in life history traits. Monocots have a high proportion of mycoheterotrophic (MHT) taxa relative to other angiosperm clades of similar size. MHT taxa implement a parasitic nutrition strategy in which carbon is obtained from nearby photosynthetic plants through a shared mycorrhizal network (Leake & Cameron, 2010). Monocots include at least 43 independently evolved MHT lineages placed in five orders: Asparagales, Liliales, Dioscoreales, Pandanales and Petrosaviales (Merckx & Freudenstein, 2010). MHT taxa possess unique morphological and life history traits related to their sym-

biosis with fungi and are divergent at the molecular level. We included representatives of four of these orders in our analysis: *Arachnitis* (Liliales), *Burmanningia* L. and *Thismia* Griff. (Dioscoreales), *Sciaphila* Blume (Pandanales) and *Petrosavia* Becc. (Petrosaviales). These taxa possessed moderate amounts of missing molecular data in our analysis, in part because of plastid gene loss. Moreover, they appear on long terminal branches relative to their neighbours in our analysis (Fig. 3), a phenomenon noted by Merckx *et al.* (2009). Despite these complications, their inclusion was essential for thorough sampling and accurate placement of the fossil constraint for the CG of Pandanales (File S2).

MHT taxa represent a case in which life history variation increases rates of nucleotide evolution among independent monocot lineages. Despite comprising mainly herbaceous species, growth habit also affects evolutionary rates in monocots. Smith & Donoghue (2008) found a shift to lower rates of molecular evolution in palms compared with the rest of monocots, probably due to a longer generation time resulting from a woody habit, which may explain short branches in Arecales for such a species-rich clade (Fig. 3). These examples of life history traits affect the ability of evolutionary models to accurately estimate divergence times. For example, evolutionary rates are autocorrelated across the entire angiosperm tree, but monocots exhibit exceptionally high rate heterogeneity (Bell *et al.*, 2010). In the case of commelinids, Barrett *et al.* (2013) noted that Poales and Arecales possess 'strikingly different internal branch lengths, representing extremes on a continuum of plastome evolutionary rates'. Moreover, Christin *et al.* (2013) found that divergence times estimated using correlated methods were incompatible with fossil evidence in grasses. For our study, methods which assume independent rates consistently failed to converge (data not shown), and variation between the r8s and multidivtime results highlight the level of uncertainty in each analysis (Table 1). Given difficulty in modeling independent and autocorrelated rates simultaneously, as recommended by Bell *et al.* (2010), resulting divergence times must be interpreted in the context of these model violations.

Our estimates of divergence times for monocot clades (Fig. 4, Table 1) indicate most orders diverged in the Cretaceous. Results from multidivtime suggest all monocot clades except Acorales diverged in the Cretaceous. Results from r8s suggest that divergence of four clades occurred later than the Cretaceous. This discordance may result from a complicating effect of life history variation (described above) on estimation of divergence times. First, r8s analysis indicated a Palaeogene divergence for Arecales, supporting a previous divergence estimate of *c.* 37 Mya



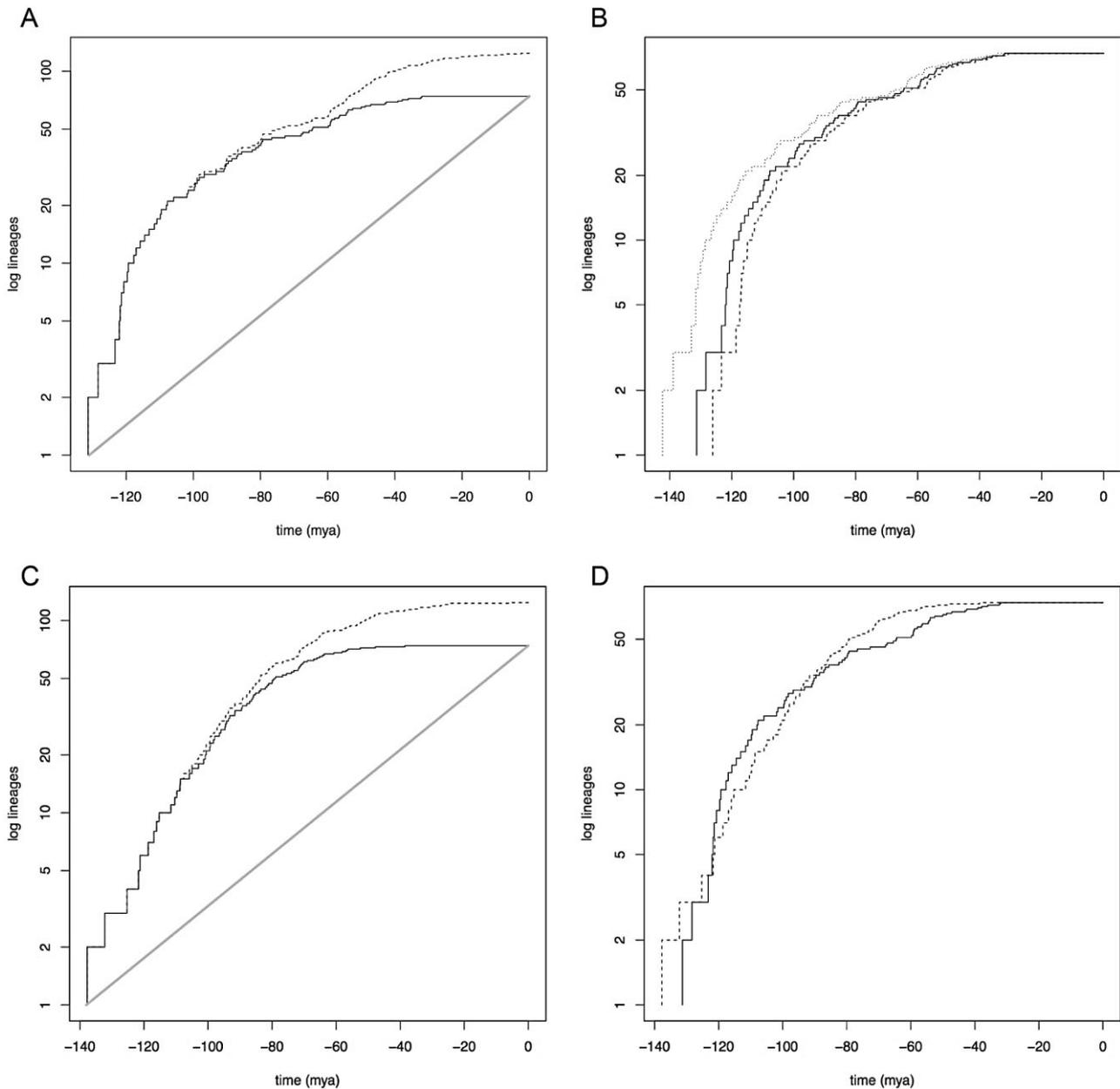
**Figure 4.** Monocot chronogram from r8s analysis. Axis represents age (Mya). Major lineages are collapsed to triangles, the size of which represents the relative number of sampled taxa. Filled circles represent the approximate placement of fossil calibrations. Specific divergence time estimates from both the r8s and multidivtime analyses are given in Table 1.

(Bell *et al.*, 2010). However, abundant Arecales fossils in the Late Cretaceous (e.g. Futey *et al.*, 2012) suggest an older divergence for the order, consistent with multidivtime. Second, the r8s analysis indicated Petrosaviales and Dasygongonaceae diverged later in the Palaeogene. Petrosaviales, however, include MHT taxa and Dasygongonaceae include woody growth habits. The final exception to a Cretaceous divergence for a monocot CG in the r8s analysis is Zingiberales (c. 56 Mya). The date reported by multidivtime agrees with an earlier estimate of divergence (Kress & Specht, 2006), although the reasons behind this discrepancy are not as clear cut as the other exceptions

(but see below). It is possible the complicating effects of life history variation contribute to rate heterogeneity, and the model assumptions of both divergence time estimation methods have been violated at one or more points in the tree.

#### SIGNIFICANT SHIFTS IN DIVERSIFICATION OCCURRED ACROSS THE MONOCOT TREE

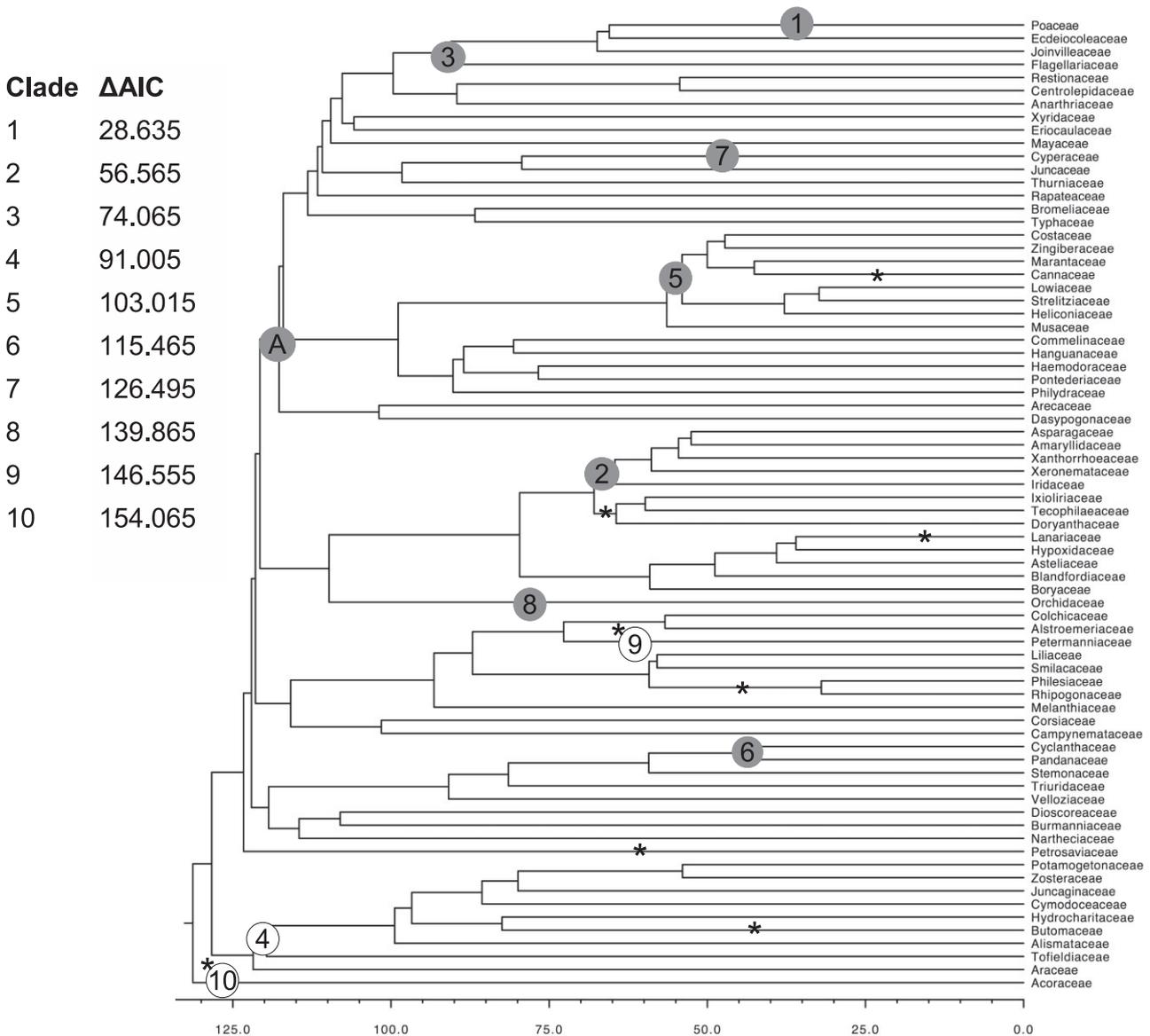
Interpreting general diversification patterns in deep nodes of the phylogenetic trees is difficult, as most current methods possess a bias leading to continual increase in species richness (Ricklefs, 2007). The



**Figure 5.** Lineage-through-time (LTT) plots of monocots. A, total species from the r8s analysis (dashed line) compared with family-level r8s analysis (solid black line). The solid grey line is included for reference, and represents a constant diversification rate over time for the family-level analysis. B, family-level r8s analysis (solid line) with dating confidence intervals (root at 140 and 180 Mya shown as dashed and dotted lines, respectively). C, total species from multidivtime analysis (dashed line) compared with family-level multidivtime analysis (solid black line). The solid grey line is included for reference, and represents a constant diversification rate over time for the family-level analysis. D, family-level r8s analysis (solid line) compared with family-level multidivtime analysis (dashed line).

pattern exhibited in monocots of a sharp initial increase of diversity followed by levelling off of lineage accumulation (Fig. 5) could be attributed to two phenomena: (1) monocots may have been historically diverse, experienced high extinction rates, leaving only a few remnant clades; or (2) monocots

experienced high speciation rates throughout their evolutionary history. Evolutionary modelling suggests that such patterns can only emerge from the latter explanation (Rabosky & Lovette, 2008). The sparse monocot fossil record from the Early Cretaceous also indicates low diversity of ancestral lineages, although



**Figure 6.** Diversification among monocot lineages. Chronogram is the r8s analysis trimmed to one taxon per family. Circles represent shifts in diversification; grey and white circles indicate shifts associated with species-rich and species-poor lineages, respectively. The circle marked A signifies the significant shift at the CG commelinids ( $P = 0.029$ ) selected by a topological diversification test in ApTreeshape (Bortolussi *et al.*, 2005). Numbers in circles indicate the order in which clades were added by stepwise AIC selection in MEDUSA (Harmon *et al.*, 2008) from the r8s chronogram. Asterisks (\*) indicate shifts in diversification from MEDUSA analysis of the multidivtime chronogram, all of which were associated with species-poor lineages.

this could be related to low abundance of early monocots. However, the appearance of relatively high levels of fossil diversity around 65 Mya (e.g. Crane, Friis & Pedersen, 1995) supports a hypothesis of radiation at that time.

This general pattern of diversification probably masks more subtle effects of rate variation among lineages, however. MEDUSA analysis of the r8s chronogram (Fig. 6) confirmed changing rates in com-

melinids by identifying three clades in Poales and one in Zingiberales exhibiting shifting diversification rates. Magallón & Castillo (2009) reported commelinids were included among lineages with some of the highest diversification rates in angiosperms and were also the highest among monocots. However, they also noted relatively low rates of diversification in Asparagales, despite the inclusion of the largest family of angiosperms, Orchidaceae, which they attributed to

the age of the order. Nevertheless, we recover significant shifts in diversification associated with core Asparagales and Orchidaceae in the r8s chronogram. We also identified an additional shift in rate leading to a species-rich clade in Pandanales (Cyclanthaceae + Pandanaceae), compared with two relatively species-poor families (Stemonaceae and Triuridaceae). Shifts in diversification reside in three additional monocot orders and are associated with species-poor lineages, including monogeneric Petermanniaceae (Liliales) and Acoraceae (Acorales). The decrease in diversification in Acoraceae is probably related to the proximity to the species-rich Araceae clade; another shift in diversification also occurs after its divergence from the rest of Alismatales. MEDUSA analysis of the multidivtime chronogram identified fewer nodes associated with shifts in diversification, two of which were identical to the r8s analysis. We prefer the interpretation of diversification as revealed by the MEDUSA analysis of r8s dates for two reasons. First, it is highly unlikely that for such a diverse clade and such a long timeframe that shifts in diversification only occurred in conjunction with species-poor clades. Second, the apTreeshape analysis found only one significant shift in rate, associated with the commelinids. The MEDUSA analysis of multidivtime failed to identify any shifts in the commelinids, whereas the corresponding analysis of r8s dates identified a shift in Poales.

#### MONOCOTS DIVERSIFIED IN THE CONTEXT OF OTHER ORGANISMAL RADIATIONS

Most monocot stem lineages diverged in the mid-Cretaceous, *c.* 100 Mya. Our evidence also suggests monocot diversification and radiation accelerated after the diversification of other major lineages of plants and some animals. Angiosperm-dominated forests composed primarily of rosids (Wang *et al.*, 2009) arose in the Late Cretaceous and created an understory suitable for diversification of the heterosporous ferns and the polypodiaceous families (Schneider *et al.*, 2004; Taylor, Taylor & Krings, 2009). Fern diversification has been attributed to the radiation of angiosperm-dominated forests and subsequent creation of 'new ecospace into which certain lineages could diversify' (Schneider *et al.*, 2004). The overall timing of diversification in monocots closely parallels that of ferns, and may result in part from possession of a similarly herbaceous habit.

Ecospace were appearing as the composition of forests changed but, more importantly, newly emerged diversity in animal lineages important for plant pollination and dispersal were now available. Animal lineages experiencing rapid diversification at this time include placental mammals (Bininda-Emonds *et al.*,

2007), amphibians (Roelants *et al.*, 2007), weevils (McKenna *et al.*, 2009) and ants (Moreau *et al.*, 2006). In fact, specialized pollination modes (including Hymenoptera) are found in 75% of early-diverging monocot families without wind pollination, and specialized pollination increased during the Late Cretaceous to early Palaeogene (Hu *et al.*, 2008). Even more important than the presence of specialized pollinators in the Late Cretaceous was the availability of new seed dispersal mechanisms providing for local adaptation and selection (Crane *et al.*, 1995). A comparison between 77 angiosperm ant-dispersed/non-ant-dispersed sister pairs, including 12 monocot pairs, found that ant-dispersed taxa have diversified more than their sister clades (Lengyel *et al.*, 2009). The presence of fleshy fruits in eight monocot orders (Givnish *et al.*, 2005) suggests shifts in diversification or relatively young divergence times may have been related to concomitant radiation in animal lineages, perhaps in conjunction with dispersal syndromes.

#### CONCLUSION

The fossil record, molecular phylogenetics, extant species diversity and divergence times inferred from evolutionary rates provide a framework to explain historical and contemporary patterns of diversity in monocots. We note several instances of shifting diversification rates in monocots and place these events in the context of the diversification of other organisms. In particular, radiation of ants and other animal lineages relevant to plant pollination and dispersal allowed for rapid diversification in a few key orders. The extraordinary diversity of monocot lineages makes estimation of evolutionary origins particularly challenging. Multiple evolutionary forces appear to have acted on different monocot lineages throughout their history. These findings provide tantalizing hypotheses for future exploration of the causes and consequences of specific episodes during plant evolution.

#### ACKNOWLEDGEMENTS

We thank Jill LeRoy for assistance in data collection; Susana Magallón and Ruth Stockey for advice on fossil calibrations; Mark Beilstein, Nathalie Nagalingum and Mark Clements for assistance with divergence time estimation; and an associate editor and two anonymous reviewers for comments which greatly improved the manuscript. This work was supported by National Science Foundation grants to the Monocot Assembling the Tree of Life group (DEB 0829849), K.L.H. as a scholar at the National Evolutionary Synthesis Center (NESCent, EF-0905606), to

M.A.G. (DEB 0830020, DEB 0918932), to S.M. (DEB 0196150) and by the Arnold Arboretum of Harvard University.

## REFERENCES

- Anderson CL, Janssen T. 2009.** Monocots. In: Hedges SB, Kumar S, eds. *Timetree of life*. New York: Oxford University Press, 203–212.
- APG III. 2009.** An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG III. *Botanical Journal of the Linnean Society* **161**: 105–121.
- Asmussen CB, Dransfield J, Deickmann V, Barfod AS, Pintaud JC, Baker WJ. 2006.** A new subfamily classification of the palm family (Arecaceae): evidence from plastid DNA phylogeny. *Botanical Journal of the Linnean Society* **151**: 15–38.
- von Balthazar M, Pedersen KR, Crane PR, Friis EM. 2008.** *Carpelestella lacunata* gen. et sp. nov., a new basal angiosperm flower from the Early Cretaceous (Early to Middle Albian) of eastern North America. *International Journal of Plant Sciences* **169**: 890–898.
- Barrett CF, Davis JI, Leebens-Mack J, Conran JG, Stevenson DW. 2013.** Plastid genomes and deep relationships among the commelinid monocot angiosperms. *Cladistics* **29**: 65–87.
- Bell CD, Soltis DE, Soltis PS. 2010.** The age and diversification of the angiosperms re-visited. *American Journal of Botany* **97**: 1296–1303.
- Bennett JR, Mathews S. 2006.** Phylogeny of the parasitic plant family Orobanchaceae inferred from phytochrome A. *American Journal of Botany* **93**: 1039–1051.
- Berry EW. 1905.** A palm from the Mid-Cretaceous. *Torreya* **5**: 30–33.
- Berry EW. 1911.** Contributions to the Mesozoic flora of the Atlantic coastal plain-VII. *Bulletin of the Torrey Botanical Club* **38**: 399–424.
- Berry EW. 1916.** Systematic paleontology, Upper Cretaceous: fossil plants. In: Clark WB, ed. *Upper Cretaceous*. Baltimore: Maryland Geological Survey, 757–901.
- Bininda-Emonds ORP, Cardillo M, Jones KE, MacPhee RDE, Beck RMD, Grenyer R, Price SA, Vos RA, Gittleman JL, Purvis A. 2007.** The delayed rise of present-day mammals. *Nature* **446**: 507–512.
- Blankers T, Townsend TM, Pepe K, Reeder TW, Wiens JJ. 2013.** Contrasting global-scale evolutionary radiations: phylogeny, diversification, and morphological evolution in the major clades of iguanian lizards. *Biological Journal of the Linnean Society* **108**: 127–143.
- Bortolussi N, Durand E, Blum M, Francois O. 2005.** apTreeshape: statistical analysis of phylogenetic tree shape. *Bioinformatics* **22**: 363–364.
- Bremer K. 2000.** Early Cretaceous lineages of monocot flowering plants. *Proceedings of the National Academy of Sciences of the United States of America* **97**: 4707–4711.
- Brenner GJ, Bickoff IS. 1992.** Palynology and age of the Lower Cretaceous Basal Kurnub Group from the coastal plain to the northern Negev of Israel. *Palynology* **16**: 137–185.
- Brown JW, Rest JS, Garcìa-Moreno J, Sorenson MD, Mindell DP. 2008.** Strong mitochondrial DNA support for a Cretaceous origin of modern avian lineages. *BMC Biology* **6**: 6.
- Chase MW, Fay MF, Devey DS, Maurin O, Ronsted N, Davies TJ, Pillon Y, Pertersen G, Seberg O, Tamura M, Asmussen CB, Hilu K, Borsch T, Davis JI, Stevenson DW, Pires JC, Givnish TJ, Sytsma KJ, McPherson MA, Graham SW, Rai HS. 2006.** Multigene analyses of monocot relationships: a summary. *Aliso* **22**: 63–75.
- Chase MW, Soltis DE, Olmstead RG, Morgan D, Les DH, Mishler BD, Duvall MR, Price R, Hills HG, Qui Y-L, Kron KA, Rettig JH, Conti E, Palmer JD, Manhart JR, Sytsma KJ, Michaels HJ, Kress WJ, Karol KG, Clark WD, Hedren M, Gaut BS, Jansen RK, Kim K-J, Wimpee CF, Smith JF, Furnier GR, Strauss SH, Xiang Q-Y, Plunkett GM, Soltis PS, Williams SE, Gadek PA, Quinn CJ, Eguiarte LE, Golenberg E, Learn GH, Graham S, Barrett SCH, Dayanandan S, Albert VA. 1993.** Phylogenetics of seed plants: an analysis of nucleotide sequences from the plastid gene *rbcL*. *Annals of the Missouri Botanical Garden* **80**: 528–580.
- Chase MW, Soltis DE, Soltis PS, Rudall PJ, Fay MF, Hahn WH, Sullivan S, Joseph J, Molvray M, Kores PJ, Givnish TJ, Sytsma KJ, Pires JC. 2000.** Higher-level systematics of the monocotyledons: an assessment of current knowledge and a new classification. In: Wilson KA, Morrison DA, eds. *Monocots: systematics and evolution*. Collingwood: CSIRO Publishing, 3–16.
- Chase MW, Stevenson DW, Wilkin P, Rudall PJ. 1995.** Monocot systematics: a combined analysis. In: Rudall PJ, Cribb PJ, Cutler DF, Humphries CJ, eds. *Monocotyledons: systematics and evolution, volume II*. Kew: Royal Botanic Gardens, 685–730.
- Chaw SM, Chang CC, Chen HL, Li WH. 2004.** Dating the monocot–dicot divergence and the origin of core eudicots using whole chloroplast genomes. *Journal of Molecular Evolution* **58**: 424–441.
- Christin PA, Spriggs E, Osborne CP, Stromberg CAE, Salamin N, Edwards EJ. 2013.** Molecular dating, evolutionary rates, and the age of the grasses. *Systematic Biology* **63**: 153–165.
- Crane PR, Friis EM, Pedersen KR. 1995.** The origin and early diversification of angiosperms. *Nature* **374**: 27–33.
- Crepet WL, Gandolfo MA. 2008.** Paleobotany in the post-genomics era: introduction. *Annals of the Missouri Botanical Garden* **95**: 1–2.
- Crepet WL, Nixon KC, Gandolfo MA. 2004.** Fossil evidence and phylogeny: the age of major angiosperm clades based on mesofossil and macrofossil evidence from Cretaceous deposits. *American Journal of Botany* **91**: 1666–1682.
- Daghlian CP. 1981.** A review of the fossil record of monocotyledons. *The Botanical Review* **47**: 517–555.

- Davis JI, Stevenson DW, Petersen G, Seberg O, Campbell LM, Freudenstein JV, Goldman DH, Hardy CR, Michelangeli FA, Simmons MP, Specht CD, 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.
- Doyle JA, Endress PK. 2010.** Integrating Early Cretaceous fossils into the phylogeny of living angiosperms: Magnoliidae and eudicots. *Journal of Systematics and Evolution* **48**: 1–35.
- Doyle JJ, Doyle JL. 1987.** A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochemical Bulletin, Botanical Society of America* **19**: 11–15.
- Drinnan AN, Crane PR, Friis EM, Pedersen KR. 1991.** Angiosperm flowers and tricolpate pollen of buxaceous affinity from the Potomac Group (Mid-Cretaceous) of eastern North America. *American Journal of Botany* **78**: 153–176.
- Edgar RC. 2004a.** MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics* **5**: 113.
- Edgar RC. 2004b.** MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research* **32**: 1792–1797.
- Friis EM. 1988.** *Spirematospermum chandlerae* sp. nov., an extinct species of Zingiberaceae from the North American Cretaceous. *Tertiary Research* **9**: 7–12.
- Friis EM, Crane PR, Pedersen KR. 2011.** *Early flowers and angiosperm evolution*. Cambridge: Cambridge University Press.
- Friis EM, Pedersen KR, von Balthazar M, Grimm GW, Crane PR. 2009.** *Monetianthus mirus* gen. et sp. nov., a Nymphaealean flower from the Early Cretaceous of Portugal. *International Journal of Plant Sciences* **170**: 1086–1101.
- Friis EM, Pedersen KR, Crane PR. 2001.** Fossil evidence of water lilies (Nymphaeales) in the Early Cretaceous. *Nature* **410**: 357–360.
- Friis EM, Pedersen KR, Crane PR. 2006.** Cretaceous angiosperm flowers: innovation and evolution in plant reproduction. *Palaeogeography, Palaeoclimatology, Palaeoecology* **232**: 251–293.
- Frumin S, Friis EM. 1999.** Magnoliid reproductive organs from the Cenomanian–Turonian of north-western Kazakhstan: Magnoliaceae and Illiciaceae. *Plant Systematics and Evolution* **216**: 265–288.
- Futey MK, Gandolfo MA, Zamaloa MC, Cúneo R, Cladera G. 2012.** Arecaceae fossil fruits from the Paleocene of Patagonia, Argentina. *The Botanical Review* **78**: 205–234.
- Gandolfo MA, Nixon KC, Crepet WL. 2002.** Triuridaceae fossil flowers from the Upper Cretaceous of New Jersey. *American Journal of Botany* **89**: 1940–1957.
- Gandolfo MA, Nixon KC, Crepet WL. 2004.** Cretaceous flowers of Nymphaeaceae and implications for complex insect entrapment pollination mechanisms in early angiosperms. *Proceedings of the National Academy of Sciences of the United States of America* **101**: 8056–8060.
- Gandolfo MA, Nixon KC, Crepet WL. 2008.** Selection of fossils for calibration of molecular dating models. *Annals of the Missouri Botanical Garden* **95**: 34–42.
- Givnish TJ, Evans TM, Pires JC, Sytsma KJ. 1999.** Polyphyly and convergent morphological evolution in Comelinales and Commelinidae: evidence from *rbcl* sequence data. *Molecular Phylogenetics and Evolution* **12**: 360–385.
- Givnish TJ, Pires JC, Graham SW, McPherson MA, Prince LM, Patterson TB, Rai HS, Roalson EH, Evans TM, Hahn WH, Millam KC, Meerow AW, Molvray M, Kores PJ, O'Brien HE, Hall JC, Kress WJ, Sytsma KJ. 2006.** Phylogenetic relationships of monocots based on the highly informative plastid gene *ndhF*: evidence for widespread concerted convergence. *Aliso* **22**: 28–51.
- Givnish TJ, Pires JC, Graham SW, McPherson MA, Prince LM, Patterson TB, Rai HS, Roalson EH, Evans TM, Hahn WJ, Millam KC, Meerow AW, Molvray M, Kores PJ, O'Brien HE, Hall JC, Kress WJ, Sytsma KJ. 2005.** Repeated evolution of net venation and fleshy fruits among monocots in shaded habitats confirms a priori predictions: evidence from an *ndhF* phylogeny. *Proceedings of the Royal Society B-Biological Sciences* **272**: 1481–1490.
- Gradstein FM, Ogg JG, Schmitz M. 2012.** *The geologic time scale 2012*. Amsterdam: Elsevier.
- Graham SW, Zgurski JM, McPherson MA, Cherniawsky DM, Saarela JM, Horne ESC, Smith SY, Wong WA, O'Brien HE, Biron VL, Pires JC, Olmstead RG, Chase MW, Rai HS. 2006.** Robust inference of monocot deep phylogeny using an expanded multigene plastid data set. *Aliso* **22**: 3–20.
- Harmon LJ, Weir JT, Brock CD, Glor RE, Challenger W. 2008.** GEIGER: investigating evolutionary radiations. *Bioinformatics* **24**: 129–131.
- Hu S, Dilcher DL, Jarzen DM, Taylor DW. 2008.** Early steps of angiosperm–pollinator coevolution. *Proceedings of the National Academy of Sciences of the United States of America* **105**: 240–245.
- Janssen T, Bremer K. 2004.** The age of major monocot groups inferred from 800+ *rbcl* sequences. *Botanical Journal of the Linnean Society* **146**: 385–398.
- Krassilov V, Kodrul T. 2009.** Reproductive structures associated with *Cobbania*, a floating monocot from the Late Cretaceous of the Amur Region, Russian Far East. *Acta Palaeobotanica* **49**: 233–251.
- Kress WJ, Specht CD. 2006.** The evolutionary and biogeographic origin and diversification of the tropical monocot order Zingiberales. *Aliso* **22**: 619–630.
- Lanfear R, Calcott B, Ho SYW, Guindon S. 2012.** PartitionFinder: combined selection of partitioning schemes and substitution models for phylogenetic analyses. *Molecular Biology and Evolution* **29**: 1695–1701.
- Leake JR, Cameron DD. 2010.** Physiological ecology of mycoheterotrophy. *New Phytologist* **185**: 601–605.
- Leebens-Mack J, Raubeson LA, Cui L, Kuehl JV, Fourcade MH, Chumley TW, Boore JL, Jansen RK, dePamphilis CW. 2005.** Identifying the basal angiosperm node in chloroplast genome phylogenies: sampling one's way

- out of the Felsenstein zone. *Molecular Biology and Evolution* **22**: 1948–1963.
- Lengyel S, Gove AD, Latimer AM, Majer JD, Dunn RR. 2009.** Ants sow the seeds of global diversification in flowering plants. *PLoS ONE* **4**: e5480.
- Lesquereux L. 1876.** On the Tertiary flora of the North American Lignitic, considered as evidence of the age of the formation. In: Hayden FV, ed. *U.S. geological and geographical survey of the Territories, embracing Colorado and parts of adjacent territories; being a report of the exploration for the year 1874*. Washington, DC: Annual Report of U.S. Geological and Geographical Surveys, 275–315.
- Maddison W, Maddison DR. 2011.** *Mesquite*, version 2.75. Available at: <http://mesquiteproject.org>
- Magallón S. 2009.** Flowering plants (Magnoliophyta). In: Hedges SB, Kumar S, eds. *The timetree of life*. Oxford: Oxford University Press, 161–165.
- Magallón S, Castillo A. 2009.** Angiosperm diversification through time. *American Journal of Botany* **96**: 349–365.
- Mathews S, Donoghue MJ. 1999.** The root of angiosperm phylogeny inferred from duplicate phytochrome genes. *Science* **286**: 947–950.
- Mathews S, Donoghue MJ. 2000.** Basal angiosperm phylogeny inferred from duplicate phytochromes A and C. *International Journal of Plant Sciences* **161**: S41–S55.
- Mathews S, Lavin M, Sharrock RA. 1995.** Evolution of the phytochrome gene family and its utility for phylogenetic analyses of angiosperms. *Annals of the Missouri Botanical Garden* **82**: 296–321.
- Mathews S, Sharrock RA. 1996.** The phytochrome gene family in grasses (Poaceae): a phylogeny and evidence that grasses have a subset of the loci found in dicot angiosperms. *Molecular Biology and Evolution* **13**: 1141–1150.
- McKenna DD, Sequeira AS, Marvaldi AE, Farrell BD. 2009.** Temporal lags and overlap in the diversification of weevils and flowering plants. *Proceedings of the National Academy of Sciences of the United States of America* **106**: 7083–7088.
- McNeal JR, Bennett JR, Wolfe AD, Mathews S. 2013.** Phylogeny and origins of holoparasitism in Orobanchaceae. *American Journal of Botany* **100**: 971–983.
- Merckx V, Bakker FT, Huysmans S, Smets E. 2009.** Bias and conflict in phylogenetic inference of myco-heterotrophic plants: a case study in Thismiaceae. *Cladistics* **25**: 64–77.
- Merckx V, Freudenstein JV. 2010.** Evolution of mycoheterotrophy in plants: a phylogenetic perspective. *New Phytologist* **185**: 605–609.
- Mohr B, Bernardes-de-Oliveira M. 2004.** *Endressinia brasiliensis*, a Magnoliacean angiosperm from the Lower Cretaceous Crato Formation (Brazil). *International Journal of Plant Sciences* **165**: 1121–1133.
- Moore BR, Chan KMA, Donoghue MJ. 2004.** Detecting diversification rate variation in supertrees. *Computational Biology* **4**: 487–533.
- Moore MJ, Bell CD, Soltis PS, Soltis DE. 2007.** Using plastid genome-scale data to resolve enigmatic relationships among basal angiosperms. *Proceedings of the National Academy of Sciences of the United States of America* **104**: 19363–19368.
- Moreau CS, Bell CD, Vila R, Archibald SB, Pierce NE. 2006.** Phylogeny of the ants: diversification in the age of angiosperms. *Science* **312**: 101–104.
- Near TJ, Sanderson MJ. 2004.** Assessing the quality of molecular divergence time estimates by fossil calibrations and fossil-based model selection. *Philosophical Transactions of the Royal Society B: Biological Sciences* **359**: 1477–1483.
- Neer S, Mooers AO, Harvey PH. 1992.** Tempo and mode of evolution revealed from molecular phylogenies. *Proceedings of the National Academy of Sciences of the United States of America* **89**: 8322–8326.
- Nylander JAA, Wilgenbusch JC, Warren DL, Swofford DL. 2008.** AWTY (are we there yet?): a system for graphical exploration of MCMC convergence in Bayesian phylogenetics. *Bioinformatics* **24**: 581–583.
- Paradis E, Claude J, Strimmer K. 2004.** APE: analyses of phylogenetics and evolution in R language. *Bioinformatics* **20**: 289–290.
- Parfrey LW, Lahr DJ, Knoll AH, Katz LA. 2011.** Estimating the timing of early eukaryotic diversification with multigene molecular clocks. *Proceedings of the National Academy of Sciences of the United States of America* **108**: 13624–13629.
- Parham JF, Donoghue PCJ, Bell CJ, Calway TD, Head JJ, Holroyd PA, Inoue JG, Irmis RB, Joyce WG, Ksepka DT, Patane JSL, Smith ND, Tarver JE, van Tuinen M, Yang Z, Angielczyk KD, Greenwood JM, Hipsley CA, Jacobs L, Makovicky PJ, Muller J, Smith KT, Theodor JM, Warnock RCM, Benton MJ. 2011.** Best practices for justifying fossil calibrations. *Systematic Biology* **61**: 346–359.
- Prasad V, Stromberg CAE, Alimohammadian H, Sahni A. 2005.** Dinosaur coprolites and the early evolution of grasses and grazers. *Science* **310**: 1177–1180.
- Qiu Y-L, Bernasconi-Quadroni F, Soltis DE, Soltis PS, Zanis MJ, Zimmer EA, Chen Z, Savolainen V, Chase MW. 1999.** The earliest angiosperms: evidence from mitochondrial, plastid and nuclear genomes. *Nature* **402**: 404–407.
- Rabosky DL, Lovette IJ. 2008.** Explosive evolutionary radiations: decreasing speciation or increasing extinction through time? *Evolution* **62**: 1866–1875.
- Rambaut A. 2012.** *FigTree v1.4*. Available at: <http://tree.bio.ed.ac.uk/software/figtree>.
- Rambaut A, Drummond AJ. 2009.** *Tracer v1.5*. Available at: <http://tree.bio.ed.ac.uk/software/tracer/>
- Revell LJ. 2011.** phytools: an R package for phylogenetic comparative biology (and other things). *Methods in Ecology and Evolution* **3**: 217–223.
- Ricklefs RE. 2007.** Estimating diversification rates from phylogenetic information. *Trends in Ecology and Evolution* **22**: 601–610.
- Rodriguez-de la Rosa RA, Cevallos-Ferriz SRS. 1994.** Upper Cretaceous zingiberalean fruits with *in situ* seeds from southeastern Coahuila, Mexico. *International Journal of Plant Sciences* **155**: 786–805.

- Roelants K, Gower DJ, Wilkinson M, Loader SP, Biju SD, Guillaume K, Moriau L, Bossuyt F. 2007. Global patterns of diversification in the history of modern amphibians. *Proceedings of the National Academy of Sciences of the United States of America* **104**: 887–892.
- Ronquist F, Teslenko M, van der Mark P, Ayres DL, Darling A, Höhna S, Larget B, Liu L, Suchard MA, Huelsenbeck JP. 2012. MrBayes 3.2: efficient bayesian phylogenetic inference and model choice across a large model space. *Systematic Biology* **61**: 539–542.
- Rutschmann F. 2005. *Bayesian molecular dating using PAML/multidivtime: a step-by-step manual. v1.5*. Zurich: University of Zurich.
- Saarela JM, Rai HS, Doyle JA, Endress PK, Mathews S, Marchant A, Briggs BG, Graham SW. 2007. Hydatellaceae identified as a new branch near the base of the angiosperm phylogenetic tree. *Nature* **446**: 312–315.
- Sanderson MJ. 2002. Estimating absolute rates of molecular evolution and divergence times: a penalized likelihood approach. *Molecular Biology and Evolution* **19**: 101–109.
- Sanderson MJ. 2003. r8s: inferring absolute rates of molecular evolution and divergence times in the absence of a molecular clock. *Bioinformatics* **19**: 301–302.
- Sanderson MJ, Doyle JA. 2001. Sources of error and confidence intervals in estimating the age of angiosperms from *rbcL* and 18S rDNA data. *American Journal of Botany* **88**: 1499–1516.
- Schneider H, Schuettpelz E, Pryer KM, Cranfill R, Magallón S, Lupia R. 2004. Ferns diversified in the shadow of angiosperms. *Nature* **428**: 553–557.
- Smith JF, Sytsma KJ, Shoemaker JS, Smith RL. 1991. A qualitative comparison of total cellular DNA extraction protocols. *Phytochemical Bulletin, Botanical Society of America* **23**: 2–9.
- Smith SA, Donoghue MJ. 2008. Rates of molecular evolution are linked to life history in flowering plants. *Science* **322**: 86–89.
- Smith SY. 2013. The fossil record of noncommelinid monocots. In: Wilkin P, Mayo SJ, eds. *Early events in monocot evolution*. Cambridge: Cambridge University Press, 39–59.
- Soltis DE, Soltis PS, Chase MW, Mort ME, Albach DC, Zanis M, Savolainen V, Hahn WH, Hoot SB, Fay MF, Axtell M, Swensen SM, Prince LM, Kress WJ, Nixon KC, Farris JS. 2000. Angiosperm phylogeny inferred from 18S rDNA, *rbcL*, and *atpB* sequences. *Botanical Journal of the Linnean Society* **133**: 381–461.
- Soltis P, Soltis D. 2013. Angiosperm phylogeny: a framework for studies of genome evolution. In: Greilhuber J, Dolezel J, Wendel JF, eds. *Plant genome diversity Volume 2*. Vienna: Springer, 1–11.
- Stamatakis A, Hoover P, Rougemont J. 2008. A rapid bootstrap algorithm for the RAxML web servers. *Systematic Biology* **57**: 758–771.
- Stevens PF. 2001 onwards. *Angiosperm phylogeny website* Available at: <http://www.mobot.org/MOBOT/research/APweb/>. Accessed 2013.
- Stockey RA, Rothwell GW, Johnson KR. 2007. *Cobbania corrugata* gen. et comb. nov. (Araceae): a floating aquatic monocot from the upper Cretaceous of western North America. *American Journal of Botany* **94**: 609–624.
- Taylor EL, Taylor TN, Krings M. 2009. *Paleobotany: the biology and evolution of fossil plants*. Burlington: Academic Press.
- Thorne JL, Kishino H. 2002. Divergence time and evolutionary rate estimation with multilocus data. *Systematic Biology* **51**: 689–702.
- Walker JW, Walker AG. 1984. Ultrastructure of Lower Cretaceous angiosperm pollen and the origin and early evolution of flowering plants. *Annals of the Missouri Botanical Garden* **71**: 464–521.
- Wang H, Moore MJ, Soltis PS, Bell CD, Brockington SF, Alexandre R, Davis CC, Latvis M, Manchester SR, Soltis DE. 2009. Rosid radiation and the rapid rise of angiosperm-dominated forests. *Proceedings of the National Academy of Sciences of the United States of America* **106**: 3853–3858.
- Yang Z. 2007. PAML 4: phylogenetic analysis by maximum likelihood. *Molecular Biology and Evolution* **24**: 1586–1591.
- Zavada MS, Benson JM. 1987. First fossil evidence for the primitive angiosperm family Lactoridaceae. *American Journal of Botany* **74**: 1590–1594.

## SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article at the publisher's web-site:

**File S1.** Taxon and voucher information. Family assignments follow APGIII. An asterisk (\*) indicates a mycoheterotrophic taxon. A carat (^) indicates a GenBank accession submitted for this study. Voucher information for non-PHYC taxa is available through the GenBank accession. Names in parentheses represent taxon substitutions for that locus. KewDNA refers to Kew DNA Bank accession number (<http://apps.kew.org/dnabank>). Non-monocot lineages include order in the first column.

**File S2.** PHY C primers designed to amplify monocot orders.

**File S3.** Fossils used for calibration of divergence times. Node label refers to text in the Methods. Constrained nodes indicate the lineage for which each minimum date is assigned (SL, stem lineage; CG, crown group). Published ages of fossils were transformed to absolute ages when necessary using the younger bound of the interval based on the current stratigraphic timescale (Gradstein *et al.*, 2012).

**File S4.** Monocot taxonomy and species counts used for diversification analyses.