

Nested radiations and the pulse of angiosperm diversification: increased diversification rates often follow whole genome duplications

David C. Tank^{1,2}, Jonathan M. Eastman^{1,2}, Matthew W. Pennell^{1,2}, Pamela S. Soltis³, Douglas E. Soltis^{3,4}, Cody E. Hinchliff⁵, Joseph W. Brown⁵, Emily B. Sessa⁴ and Luke J. Harmon^{1,2}

¹Department of Biological Sciences, University of Idaho, Moscow, ID 83844, USA; ²Institute for Bioinformatics and Evolutionary Studies, University of Idaho, Moscow, ID 83844, USA;

³Florida Museum of Natural History, University of Florida, Gainesville, FL 32611, USA; ⁴Department of Biology, University of Florida, Gainesville, FL 32611, USA; ⁵Department of Ecology and Evolutionary Biology, University of Michigan, Ann Arbor, MI 48109, USA

Summary

Author for correspondence:

David C. Tank

Tel: +01 208 885 7033

Email: dtank@uidaho.edu

Received: 10 August 2014

Accepted: 1 May 2015

New Phytologist (2015)

doi: 10.1111/nph.13491

Key words: angiosperm diversification rates, clade age and diversity, modeling evolutionary diversification using stepwise AIC (MEDUSA), nested radiations, polyploidy, whole-genome duplication radiation lag-time model.

- Our growing understanding of the plant tree of life provides a novel opportunity to uncover the major drivers of angiosperm diversity.
- Using a time-calibrated phylogeny, we characterized hot and cold spots of lineage diversification across the angiosperm tree of life by modeling evolutionary diversification using stepwise AIC (MEDUSA). We also tested the whole-genome duplication (WGD) radiation lag-time model, which postulates that increases in diversification tend to lag behind established WGD events.
- Diversification rates have been incredibly heterogeneous throughout the evolutionary history of angiosperms and reveal a pattern of 'nested radiations' – increases in net diversification nested within other radiations. This pattern in turn generates a negative relationship between clade age and diversity across both families and orders. We suggest that stochastically changing diversification rates across the phylogeny explain these patterns. Finally, we demonstrate significant statistical support for the WGD radiation lag-time model.
- Across angiosperms, nested shifts in diversification led to an overall increasing rate of net diversification and declining relative extinction rates through time. These diversification shifts are only rarely perfectly associated with WGD events, but commonly follow them after a lag period.

Introduction

With over 250 000 species (Crane *et al.*, 1995; Judd *et al.*, 2007; and probably many more: e.g. Joppa *et al.*, 2011), angiosperms are among the most impressive radiations of terrestrial organisms. In addition to their incredible species diversity, angiosperms are ecologically, functionally, and morphologically diverse, having risen to dominance in most terrestrial environments in a (geologically) short period of time since their major diversification in the Cretaceous (Lidgard & Crane, 1990; Crane *et al.*, 1995; Dilcher, 2001). Within this massive radiation of species, it remains a mystery why some plant groups are much more diverse than others. For example, the cosmopolitan family Asteraceae (which includes > 23 000 species, including daisies and sunflowers, in a wide variety of habitats) is < 50 million yr old (Bremer & Gustafsson, 1997; Kim, 2005; Bell *et al.*, 2010; Beaulieu *et al.*, 2013), while its sister group, Calyceraceae, is restricted to South America and comprises only 60 species (Stevens, 2001-onwards; Bremer *et al.*, 2004).

The varying tempo of angiosperm diversification has long fascinated biologists (Darwin, 1903). This mystery has only been compounded as many of the recalcitrant branches of the angiosperm tree of life have been resolved and a comprehensive view of relationships among the major lineages of angiosperms has emerged (Soltis *et al.*, 2011). At the same time, methods for identifying shifts in the rate of net diversification (speciation minus extinction) have advanced considerably, allowing us to pinpoint 'hot' and 'cold' spots on phylogenetic trees (e.g. where rates of speciation and extinction have changed; Alfaro *et al.*, 2009). Despite decades of work characterizing angiosperm diversification from both paleontological and phylogenetic perspectives (Lidgard & Crane, 1990; Sanderson & Donoghue, 1994; Crane *et al.*, 1995; Dilcher, 2001; Magallón & Sanderson, 2001; Davies & Barraclough, 2004; Magallón & Castillo, 2009; Smith *et al.*, 2011), we still do not have a clear idea of the drivers of differential diversification across plant species. There have been many hypotheses put forward to explain the dramatic rise to dominance of angiosperms, and explanations are often based on ecological

and physiological innovations and the functional traits associated with these (Berendse & Scheffer, 2009; Brodribb & Feild, 2010; Labandeira, 2010; Feild *et al.*, 2011; reviewed in Augusto *et al.*, 2014). Recently, with the availability of genome-scale data representing an increasingly wide variety of angiosperm lineages, and a stabilizing phylogenetic framework to investigate genome evolution in a comparative framework, there has been increased interest in the role of paleopolyploidy – or ancient whole-genome duplication (WGD) – with respect to increased rates of lineage diversification in angiosperms (Schranz *et al.*, 2012; Vanneste *et al.*, 2014).

Recent genomic investigations indicate that polyploidy is ubiquitous among angiosperms, with evidence for an ancient WGD event preceding the origin of the clade itself (Jiao *et al.*, 2011; *Amborella* Genome Project, 2013). Genomic data also suggest other major ancient WGD events in angiosperms (Van de Peer *et al.*, 2009, 2010), including two WGDs that occurred in close temporal succession early in eudicot evolution (Jiao *et al.*, 2012) with other events close to the origin of monocots, Poales, and Solanales (Soltis *et al.*, 2009). At least 50 independent ancient WGDs are distributed across flowering plant phylogeny (Cui *et al.*, 2006; Soltis *et al.*, 2009; Van de Peer *et al.*, 2009, 2010; M. S. Barker, pers. comm.). With this increased understanding of the ubiquitous nature of polyploidy in angiosperm histories and the phylogenetic location of these paleopolyploid events, it has been recognized that a common pattern in the tree of life – species-poor lineages subtending species-rich crown groups – tends to follow established WGD events in angiosperms (Soltis *et al.*, 2009, 2014b; Schranz *et al.*, 2012). Schranz *et al.* (2012) erected a formal hypothesis – the WGD radiation lag-time model – based on observations of tree imbalance in several angiosperm lineages that have been associated with paleopolyploidy, which postulates that upticks in diversification rates tend to follow WGD events, but only after lag times that may span millions of years (Schranz *et al.*, 2012). However, there has been no formal quantitative test of this hypothesis across the angiosperm tree of life.

Our growing understanding of the structure of the plant tree of life provides a rare opportunity to investigate angiosperm diversification to try to uncover the major drivers that have led to the diversity of plant species on Earth. The convergence of a well-sampled, well-resolved, and well-supported angiosperm phylogeny and the methods for detecting heterogeneity in diversification rates from phylogenetic information – including topology, branch lengths, and clade richnesses – in a single coherent framework sets the stage for addressing the question ‘what drives differential diversification rates and the resulting striking disparities in clade diversities across angiosperms?’

Here we use phylogenetic and taxonomic data to investigate angiosperm diversification dynamics. First, we characterize hot and cold spots of diversification across the angiosperm tree of life. We identify a heterogeneous pattern of rapid radiations nested within other rapid radiations, which we refer to as ‘nested radiations’ – similar to the ‘repeated radiations’ observation of Soltis *et al.* (2004), but here with formal diversification rate analyses – and suggest that stochastically changing

diversification rates across lineages generate this pattern (Stadler *et al.*, 2014). Second, we show that increases in diversification tend to lag behind major paleopolyploid events. These delayed bursts may suggest that genome duplications promote, but are not sufficient to cause, increased diversification. We suggest how such duplications might interact with other biotic and abiotic influences to lead to the nested pattern of radiations that we see across plant species.

Materials and Methods

Methodological overview

We present fully detailed explanations of our methods in the following sections. As a general overview, we first extended a time-calibrated analysis of the Soltis *et al.* (2011) angiosperm phylogeny (Zanne *et al.*, 2014) to incorporate phylogenetic uncertainty. We then used the modeling evolutionary diversification using stepwise AIC (MEDUSA) approach developed by Alfaro *et al.* (2009; extended by Pennell *et al.*, 2014) to identify a set of clade-specific shifts in diversification rate across angiosperms. Finally, we placed nine well-characterized angiosperm WGDs on the phylogeny to test the WGD radiation lag-time model (Schranz *et al.*, 2012) of angiosperm diversification.

Time-calibrated phylogeny

We obtained a phylogenetic tree with branch lengths proportional to time from a previous publication (Zanne *et al.*, 2014). Full details of tree construction are available in the original paper. Briefly, we compiled sequences for eight plastid genes (*atpB*, *matK*, *ndhF*, *psbBTN*, *rbcl*, *rpoC2*, *rps16* and *rps4*) from Soltis *et al.* (2011), and used a by-gene partitioned maximum likelihood analysis in RAxML v. 7.4.1 (Stamatakis, 2006; Ott *et al.*, 2007) to obtain a 639-taxon tree. Tree topology was constrained according to Soltis *et al.* (2011) to ensure concordance with well-supported deep relationships among taxa despite (relatively) limited genetic data. This ML tree was time-scaled using 39 fossil calibrations and a penalized likelihood approach (see Zanne *et al.*, 2014 for full details). The fossil calibrations used by Zanne *et al.* (2014) represent a reliable set of fossils that could be confidently identified and placed on the phylogeny, and, in addition to Zanne *et al.* (2014), these fossils have all been used in other comprehensive large-scale dating analyses in plants (Bell *et al.*, 2010; Smith *et al.*, 2010; Beaulieu *et al.*, 2013). Rate smoothing was conducted by penalized likelihood (TREEPL; Smith & O’Meara, 2012) using a smoothing parameter of 0.1 that was optimized on the maximum likelihood tree. For each calibration, both minimum and maximum age constraints were applied, where minimum age constraints corresponded to the age of the fossil used in previous analyses (Bell *et al.*, 2010; Smith *et al.*, 2010; Beaulieu *et al.*, 2013) and maximum age constraints were calculated from the upper 97.5% of the lognormal distribution with means and standard deviations following the lognormal priors used for the Bayesian

divergence time estimates of Bell *et al.* (2010), Smith *et al.* (2010), and Beaulieu *et al.* (2013; see Zanne *et al.*, 2014 supplementary Table 2 for full details). In addition to the fossil calibrations, Zanne *et al.* (2014) constrained the root node with a minimum age of 301 million yr (Myr) and a maximum of 366 Myr following the results of Smith *et al.* (2010) and recommendations of Clarke *et al.* (2011). To account for uncertainty in divergence time estimation, we repeated this analysis pipeline for a set of 1024 bootstrapped data sets. For each of these replicates, we resampled the genetic sequence data with replacement to obtain new alignments with the same length as the original, and then repeated all of the analyses described above to obtain a chronogram with branch units in millions of years. We applied all the diversification rate analyses described in the following section across this full set of 1024 bootstrapped chronograms.

Identifying diversification shifts

We used MEDUSA (Alfaro *et al.*, 2009; Pennell *et al.*, 2014) to identify shifts in diversification rates along branches in the seed plant phylogeny. After initially fitting a constant rate birth–death model of diversification to the phylogenetic tree, MEDUSA uses a step-wise addition algorithm to infer phylogenetically local shifts in the rates of two diversification parameters: net diversification ($r = \lambda - \mu$) and relative extinction ($\epsilon = \mu/\lambda$), where λ is the speciation rate (birth) and μ is the rate of extinction (death). Rate shifts are retained if including the shift substantially improves the sample-size corrected Akaike Information Criterion (AIC) score (AICc; Burnham & Anderson, 2002). The version of MEDUSA we used for this analysis has been improved from the original implementation in three ways: it (1) considers a mixture of pure-birth and birth–death processes in shifted lineages; (2) uses an AIC threshold that is corrected for tree size; and (3) considers both forward and backward selection in model choice (see Pennell *et al.*, 2014 for more details).

For this MEDUSA analysis, we collapsed our 639-taxon phylogeny into an exemplar tree, with one representative for each of the 325 sampled seed plant families. Because MEDUSA requires that all missing species be assigned to a tip clade in the full tree, we mapped species richnesses to each familial exemplar using a comprehensive systematic resource, The Plant List (2010). We were able to account for a large proportion (0.949) of known species richness in seed plants. Inferences of rate heterogeneity in the diversification process of seed plants are drawn from the compilation of MEDUSA analyses across our full bootstrapped distribution of time-calibrated trees ($n = 1024$; Supporting Information Table S1).

To summarize the time-course of speciation and extinction from our MEDUSA analysis, we divided each tree into 1-Myr time intervals. For each time interval, we calculated the minimum, maximum, and mean of all branch-specific rate estimates of net diversification (r) and relative extinction (ϵ) for all branches occurring in that interval. This results in a time-series of rate estimates for each tree. We then repeated this procedure over all of the bootstrap trees, and summarized the results by

computing the average and standard deviation of the mean rates from each tree, as well as the overall range of the parameter estimates across all the trees. When resolution is lost at the family level (as a result of the incorporation of unsampled diversity), the average parameter values represent only those lineages with resolution, resulting in an increasingly smaller number of lineages contributing to the calculation of mean rates towards the present.

Clade age–diversity relationships

We tested for a relationship between clade age and diversity at the level of plant families and orders. To do this, we correlated the species diversity of each family ($n = 325$) and each order ($n = 66$) with its stem age using linear regression.

Whole-genome duplications and correlation with diversification rates

Finally, we explored the relationship between WGDs and shifts in diversification rates. To do this, we placed nine well-characterized WGDs at the family level and above on the tree and asked if there is a correspondence between polyploidization and increases in net diversification rate. We selected these nine WGDs based on both the strength of evidence in support of a WGD at a particular node and our ability to precisely place a WGD on our phylogeny given the sampling. Because we collapsed clades at the family level to incorporate unsampled diversity in our diversification rate analyses, we cannot place any WGD events within a family. In addition, because the hypothesized phylogenetic locations of many WGDs inferred from genomic data are often based on very limited taxonomic sampling, it is difficult to precisely place these events on a phylogeny.

Table 1 shows the nine WGD events that were used to test for a correlation with increased rates of diversification. In all cases, these events were identified through comparative genomic analyses. The first two represent established WGDs that occurred in the ancestor of all angiosperms (the ϵ event; Jiao *et al.*, 2011) and the subsequent radiation of the core eudicots (the γ event; Jiao *et al.*, 2012). The third established paleopolyploidization represents an ancient genome duplication in monocots. Based on integrated syntenic and phylogenomic analyses, Jiao *et al.* (2014) revealed a WGD shared by all monocots sampled to date (the τ event). It is important to note that this event is separate from the previously characterized ρ and σ events in monocots (Paterson *et al.*, 2012). However, because no noncommelinid monocot genomes have been sampled thus far, the phylogenetic placement of the τ event is more uncertain. Jiao *et al.* (2014) hypothesized that τ may represent a WGD in one of the early diverging monocot lineages, but given that sampling to date has been limited only to the *Commelinidae* (i.e. palms, bananas, and grasses), we considered several alternative placements for this event that ranged from the commelinid clade, which is the most recent place this could have occurred given the available data, to the entire monocot clade, including intermediate alternative placements along the backbone of the clade (Table 1). In their

Table 1 Whole-genome duplications (WGDs) investigated for correspondence with increased diversification rates

	WGD description	Reference	Nearest <i>r</i> increase	n-dist	Myr
1	Angiospermae (ϵ)	Jiao <i>et al.</i> (2011)	<i>Mesangiospermae</i>	3	49.2
2	<i>Gunneriidae</i> (γ)	Jiao <i>et al.</i> (2012)	MRCA of (<i>Superrosidae</i> + <i>Superasteridae</i>)	1	0.6
3	<i>Commelinidae</i> (τ)	Jiao <i>et al.</i> (2014)	MRCA of (<i>Arecaceae</i>) and (<i>Commelinales</i> + <i>Zingiberales</i>)	1	2.3
3a	<i>Monocotyledonae</i> (τ)	Jiao <i>et al.</i> (2014)	MRCA of (<i>Commelinidae</i> + <i>Asparagales</i>)	4	39.4
3b	<i>Nartheciidae</i> (τ)	Jiao <i>et al.</i> (2014)	MRCA of (<i>Commelinidae</i> + <i>Asparagales</i>)	3	26.5
3c	<i>Petrosaviidae</i> + <i>Dioscoreales</i> (τ)	Jiao <i>et al.</i> (2014)	MRCA of (<i>Commelinidae</i> + <i>Asparagales</i>)	2	5.7
3d	<i>Petrosaviidae</i> (τ)	Jiao <i>et al.</i> (2014)	MRCA of (<i>Commelinidae</i> + <i>Asparagales</i>)	1	3.0
3e	<i>Commelinidae</i> + <i>Asparagales</i> (τ)	Jiao <i>et al.</i> (2014)	MRCA of (<i>Commelinidae</i> + <i>Asparagales</i>)	0	0.0
4	<i>Poaceae</i> (ρ , σ)	Paterson <i>et al.</i> (2012); Jiao <i>et al.</i> (2014)	NA	NA	NA
5	Magnoliales + Laurales	Cui <i>et al.</i> (2006)	Lauraceae	4	26.8
6	Ranunculales	Cui <i>et al.</i> (2006); Pabón-Mora <i>et al.</i> (2013)	None	NA	NA
6a	<i>Papaveraceae</i> + <i>Ranunculaceae</i>	Pabón-Mora <i>et al.</i> (2013)	None	NA	NA
7	<i>Asteraceae</i>	Barker <i>et al.</i> (2008)	<i>Asteraceae</i>	0	0.0
8	<i>Brassicaceae</i> (α)	Haudry <i>et al.</i> (2013); Kagale <i>et al.</i> (2014)	NA	NA	NA
9	<i>Brassicaceae</i> + <i>Limnanthaceae</i> (β)	Barker <i>et al.</i> (2009)	MRCA of (<i>Capparaceae</i> + <i>Brassicaceae</i>)	3	35.9
9a	Core Brassicales [<i>Resedaceae</i> + <i>Brassicaceae</i>] (β)	Barker <i>et al.</i> (2009)	MRCA of (<i>Capparaceae</i> + <i>Brassicaceae</i>)	1	19.9
9b	<i>Brassicaceae</i> + <i>Bataceae</i> (β)	Schranz <i>et al.</i> (2011)	MRCA of (<i>Capparaceae</i> + <i>Brassicaceae</i>)	2	28.8

WGD description, location of WGD, with names following Soltis *et al.* (2011); reference, citation for each WGD; nearest *r* increase, location of the nearest increase in net diversification that was identified in > 75% of trees (*r*); n-dist, the number of nodes between WGDs and the nearest tip-ward increase in net diversification; Myr, the lag time (in millions of years) between WGD and shifted clade calculated using divergence time estimates from the maximum likelihood tree. Rows 3(a–e), 6(a), and 9(a,b), represent alternative placements of these WGDs (see text). MRCA, most recent common ancestor; NA, not applicable.

analyses, Jiao *et al.* (2014) also disambiguate the placement of the ρ and σ events, confidently placing both of these WGDs along the lineage leading to grasses (Table 1).

In a pioneering comparative genomic study of paleopolyploidy, Cui *et al.* (2006) identified several WGD events in angiosperms, including a shared WGD between Magnoliales and Laurales, and an additional duplication in Ranunculales. While these WGDs have been widely cited (Soltis *et al.*, 2009; Van de Peer *et al.*, 2009), given the limited sampling in this study (i.e. a single taxon representing each order), generalizing these duplications to the ordinal level is tentative. Using preliminary data from the 1000 plants (1KP) transcriptome project, we were able to confirm both of these events (P. S. Soltis & D. E. Soltis, unpublished data). From transcriptomes for four species of Magnoliales and six species of Laurales, initial plots of the divergence of duplicate gene pairs in terms of substitutions per synonymous site per year (K_s), indicate a shared WGD that is unique to these orders. Likewise, K_s analyses from seven species of Ranunculales revealed a shared WGD as suggested by Cui *et al.* (2006). However, because diploidization can obscure these patterns – especially at deep phylogenetic levels – these data are part of an ongoing, more detailed study that will integrate these comparative genomic analyses with syntenic and phylogenomic techniques (e.g. as in Jiao *et al.*, 2014). Nevertheless, here we include a WGD event shared by Magnoliales and Laurales, as well as a Ranunculales WGD (Table 1). For Ranunculales, additional evidence from phylogenetic analyses of MADS-box genes in Ranunculales supports a duplication early in the diversification of Ranunculales, but it is not clear whether this

duplication occurred before or after the divergence of Eupteleaceae (Pabón-Mora *et al.*, 2013). Therefore, for the ranunculid WGD, we investigated an alternative placement excluding Eupteleaceae, the sister group to the rest of the clade.

The seventh WGD that we considered is the well-studied polyploidization event that characterizes the sunflowers (*Asteraceae*; Barker *et al.*, 2008). Finally, we also considered the α and β WGDs that have been extensively studied in Brassicales (Bowers *et al.*, 2003; Barker *et al.*, 2009; Schranz *et al.*, 2011; Haudry *et al.*, 2013; Kagale *et al.*, 2014). While the phylogenetic location of the α event has been confidently placed on the lineage leading to *Brassicaceae* (Haudry *et al.*, 2013; Kagale *et al.*, 2014), because of limited genomic sampling in Brassicales, the placement of the β WGD is more problematic. Based on the sampling from Barker *et al.* (2009; *Brassicaceae*, *Cleomaceae*, and *Caricaceae*), the deepest placement of this event is following the divergence of *Caricaceae* (*Brassicaceae* + *Limnanthaceae* in Table 1). However, Barker *et al.* (2009) hypothesize that this duplication probably represents a paleopolyploidization of the core Brassicales (*Resedaceae* + *Brassicaceae* in Table 1), and Schranz *et al.* (2011) suggest that the β WGD occurred following the divergence of *Limnanthaceae* (*Brassicaceae* + *Bataceae* in Table 1). Therefore, we chose to include these three alternatives as possible placements for the β WGD (Table 1).

Given these WGDs, we first asked if there was a perfect correspondence between inferred increases in net diversification and polyploidization events. To accomplish this, we first determined how many of our nine identified WGD events corresponded

exactly with one of the 15 inferred increases in net diversification rate from our MEDUSA analyses (Table 2). We considered several alternative placements for three of the WGDs events (Table 1); we tested all 36 ($6 \times 2 \times 3$) combinations of these alternative placements. For each set of WGD placements, we counted the number of corresponding diversification rate shifts, and we compared this number with how many correspondences one would expect under a simple null model where nine WGDs were placed randomly on branches in the tree.

We then tested the WGD radiation lag-time hypothesis, which predicts a delayed elevated net diversification rate following polyploidization (Schranz *et al.*, 2012). In this case, our hypothesis is that polyploidization events will elevate net diversification rates after some delay (measured here as nodal distance). We chose nodal distance (as opposed to absolute time) to evaluate this hypothesis, because there is no expectation of exactly how long the lag time should be. The verbal model of Schranz *et al.* (2012) recognizes that the delayed shift in diversification rate may take many millions of years. In addition, the precise temporal locations of both the identified diversification rate shifts and the WGDs are unclear. We visualize both of these as being at a node, when in reality they happened somewhere along the stem lineage. Neither MEDUSA nor the approaches used to identify the phylogenetic position of WGDs provide meaningful time estimates for where these events happened. Therefore, we chose to use nodal distance with various cut-offs that represent a reasonable expectation for the time-lag hypothesis. Three of the WGD events (Poaceae, Brassicaceae, and Asteraceae; Table 1) are tip lineages. In these cases, we cannot test the WGD radiation lag-time hypothesis, because we have no resolution in the tree within these lineages (i.e. the tree was collapsed to the family level for our diversification rate analyses); we excluded these tip lineages from the lag-time test. For each of the other six WGD events, we determined whether or not any of the 15 identified shifts in elevated diversification rates (Table 2) followed within a set number of nodes on the tree (we used three nodes as our cut-off, but results were robust to cut-offs of one though four, which is the maximum distance of any rate shift from a WGD event). We used the number of WGDs that were followed by a rate shift as a test statistic, and compared this to a null distribution generated by again placing six WGDs randomly on branches in the tree. As before, we repeated this test for all 36 sets of potential WGD placements.

Results

Data availability and analysis pipelines

All data sets and R-scripts necessary to perform the analyses presented here are available at <https://github.com/harmonlab/angio-pulse>.

Time-calibrated phylogeny

Our extension of the divergence time estimates of Zanne *et al.* (2014) resulted in a bootstrap set of 1024 time-calibrated trees. This bootstrap set was used as a distribution of divergence time

estimates to account for phylogenetic (temporal) uncertainty with respect to penalized likelihood analyses. Stem ages for angiosperm families (and associated 95% confidence intervals) retained for our diversification analyses are shown in Table S2.

Variation in diversification rates among clades

Our analyses identified a total of 41 shifts in the maximum likelihood estimate (MLE) tree and 142 total unique shifts in diversification rate across the set of bootstrapped trees (Table S1). However, incorporating uncertainty in divergence times reduces this number to 27 major shifts that occur in at least 75% of the bootstrap set (Table 2; Fig. 1). Shifts were distributed throughout the tree, revealing heterogeneous patterns of diversification across angiosperms, but with most shift points clustered in the denser parts of the tree between 50 and 125 million yr ago (Ma) (Fig. 2a,b). These shifts often show a nested pattern of rate shifts within a clade followed by further shifts within a subclade of that original clade.

Our calculated diversification rates through time show high variability among branches and across the bootstrap distribution of trees. In general, we revealed a trend of increasing net diversification rates (r) and high, but decreasing, turnover (high ϵ) rates through time (Fig. 2; Tables 2, S1). Because clades were collapsed at the family level to incorporate unsampled diversity, there is a marked leveling off of the mean rate estimates for both net diversification and relative extinction using this summary approach (Fig. 2c,d).

Clade age–diversity relationships

Consistent with previous studies of angiosperm diversification (Magallón & Sanderson, 2001; Magallón & Castillo, 2009), we found a significant negative relationship between age and the natural logarithm of species richness considered at the familial level ($R^2 = 0.016$; $P = 0.0227$; Fig. S1A) and at the ordinal level ($R^2 = 0.144$; $P = 0.0017$; Fig. S1B). Stem-group ages for the 325 angiosperm families summarized across our distribution of trees are provided in Table S2.

WGDs and diversification rate

We first tested for a perfect correspondence between diversification-rate shifted lineages and polyploidization events. We found that either one or two of nine investigated polyploidization nodes were perfectly associated with diversification upticks (Fig. 3; Table 1; Asteraceae and *Commelinidae* + Asparagales, when this putative placement was selected to represent of the τ event (3e in Table 1)). We can only reject the null hypotheses in the six scenarios (out of 36) with two exact matches.

We then tested the WGD radiation lag-time hypothesis by testing for a delayed correspondence between diversification-rate shifted lineages and polyploidization (excluding three tip lineages where we would be unable to detect a delayed upturn). Using a cut-off of three nodes, we found that either three or four (out of a possible six) polyploidization nodes show delayed

Table 2 MEDUSA (modeling evolutionary diversification using stepwise AIC) estimates for a set of primary shifts in diversification finding support in at least 75% of bootstrap replicates

Description	Label	Support	Ma	r	ε	Δ_r	Δ_ε
MRCA of (Capparaceae + Brassicaceae)	26	0.94	34.7	0.2187	0.00	0.1614	-0.31
Cactaceae	27	1.00	29.1	0.2626	NA	0.1359	NA
Asteraceae	25	1.00	48.8	0.2101	NA	0.1292	NA
MRCA of (Caryophyllaceae) and (Stegnospermataceae)	23	1.00	62.2	0.1267	0.00	0.0910	-0.03
MRCA of (Gentianales + Solanales) and (Boraginaceae + Lamiales)	19	0.90	76.3	0.1287	0.52	0.0821	0.61
Fabaceae	21	1.00	69.3	0.1452	NA	0.0809	NA
Piperaceae	20	0.98	74.8	0.1012	NA	0.0534	NA
MRCA of (Araliaceae) and (Myodocarpaceae + Apiaceae)	18	0.99	76.5	0.0996	0.00	0.0469	0.00
MRCA of (Vochoysiaceae + Myrtaceae) and (Melastomataceae)	15	0.76	88.2	0.1102	0.00	0.0451	-0.44
Annonaceae	17	0.99	81.2	0.0927	NA	0.0451	NA
Lauraceae	13	0.98	93.0	0.0857	NA	0.0379	NA
<i>Mesangiospermae</i>	1	1.00	218.8	0.0473	0.22	0.0349	-0.66
MRCA of (<i>Superrosidae</i> + <i>Superasteridae</i>)	7	0.98	119.2	0.0645	0.54	0.0174	0.35
MRCA of (<i>Commelinidae</i> + Asparagales)	4	0.92	135.6	0.0626	0.94	0.0154	0.71
MRCA of (Arecaceae) and (Commelinales + Zingiberales)	5	0.92	123.8	0.0629	0.00	0.0003	-0.94
MRCA of (<i>Campanulidae</i> + <i>Lamiidae</i>)	8	1.00	106.2	0.0612	0.01	-0.0030	-0.54
MRCA of (Paeoniaceae) and (Altingiaceae)	9	0.98	103.6	0.0377	0.00	-0.0277	-0.50
MRCA of (Rhabdodendraceae) and (Simmondsiaceae)	16	0.83	87.5	0.0293	0.00	-0.0342	-0.65
Ceratophyllaceae	2	0.95	186.1	0.0075	NA	-0.0401	NA
Huertales	14	0.78	90.5	0.0223	0.00	-0.0409	-0.67
Acoraceae	3	0.97	172.4	0.0040	NA	-0.0435	NA
Blandfordiaceae	6	1.00	119.6	0.0116	NA	-0.0505	NA
MRCA of (Aphloiaceae + Strasburgeriaceae)	12	1.00	93.9	0.0074	0.00	-0.0538	-0.63
Berberidopsidales	10	1.00	97.9	0.0043	0.00	-0.0599	-0.55
Curtisiaceae	11	0.96	96.1	0.0000	NA	-0.0645	NA
MRCA of (Montiniaceae) and (Sphenocleaceae + Hydroleaceae)	22	1.00	63.1	0.0324	0.00	-0.0984	-0.68
Plocospermataceae	24	0.92	61.5	0.0000	NA	-0.1288	NA

Where indicated, values are averaged across bootstrap-replicate trees.

Description, topological position of the shifted lineage, with clade names (in italics) following Cantino *et al.* (2007) and Soltis *et al.* (2011), and family and order names following APGIII (The Angiosperm Phylogeny Group, 2009); label, temporal ordering of each shift, beginning with the oldest (i.e.

Mesangiospermae); support, fraction of trees in which shift received statistical support; Ma, mean estimate of age of shift (millions of years ago); r , maximum likelihood estimate of net diversification; ε , maximum likelihood estimate of relative extinction; Δ_r , mean magnitude of change in net diversification relative to immediate ancestor (see Fig. 1a for further details); Δ_ε , mean magnitude of change in relative extinction relative to immediate ancestor (see Fig. 1a for further details). Because diversification rate analyses were collapsed to the family level, relative extinction (ε and Δ_ε) cannot be calculated and is noted as NA (not applicable) for those clades. MRCA, most recent common ancestor.

diversification upticks (Fig. 3; Table 1; *Angiospermae*, *Gunneridae*, the monocot τ event in all positions except 3b, and all placements of the Brassicales β event). In all 36 cases we can reject the null hypothesis ($P < 0.001$ – 0.011), and using the alternate cut-offs of one, two, or four nodes, we could reject the null hypothesis of a random association in 20, 30, or all 36 cases, respectively, and thus support a delayed association between polyploidy and an uptick in diversification rate that ranges in time from 0.6 Ma (*Gunneridae* WGD and *Superasteridae* + *Superrosidae* shift) to 49.2 Ma (*Angiospermae* WGD and *Mesangiospermae* shift; Fig. 3; Table 1).

Discussion

Heterogeneous diversification

We found a striking amount of heterogeneity in diversification rates across angiosperms (Table 2; Figs 1, 2). Previous work has suggested that the overall increased diversification rates in angiosperms are associated with rate shifts that occurred after the initial diversification of angiosperms (Sanderson & Donoghue, 1994).

Later work used a variety of methods to locate and characterize these shifts (Magallón & Sanderson, 2001; Davies & Barraclough, 2004; Magallón & Castillo, 2009; Smith *et al.*, 2011; Fiz-Palacios *et al.*, 2011). In contrast to these studies, the analyses we performed here utilized all aspects of the phylogenetic data – including the topology, branch lengths, and clade richnesses – in a single coherent framework based on birth–death models.

Heterogeneity in diversification rates may occur either temporally (e.g. speciation rates slowing through time) or among lineages. Our analyses focus on the latter, as MEDUSA is a clade-based approach used to test for shifts in diversification rates between lineages (Alfaro *et al.*, 2009). An assumption of temporal-based methods is that changes in diversification rates occur homogeneously across the tree. However, for a large group such as angiosperms, this assumption is almost certainly violated: previous studies (Magallón & Sanderson, 2001; Davies & Barraclough, 2004; Magallón & Castillo, 2009; Smith *et al.*, 2011) have found substantial differences in diversification rates among angiosperm clades. Furthermore, our level of sampling at the species level is both low (639 exemplar taxa out of >250 000 known angiosperms) and highly

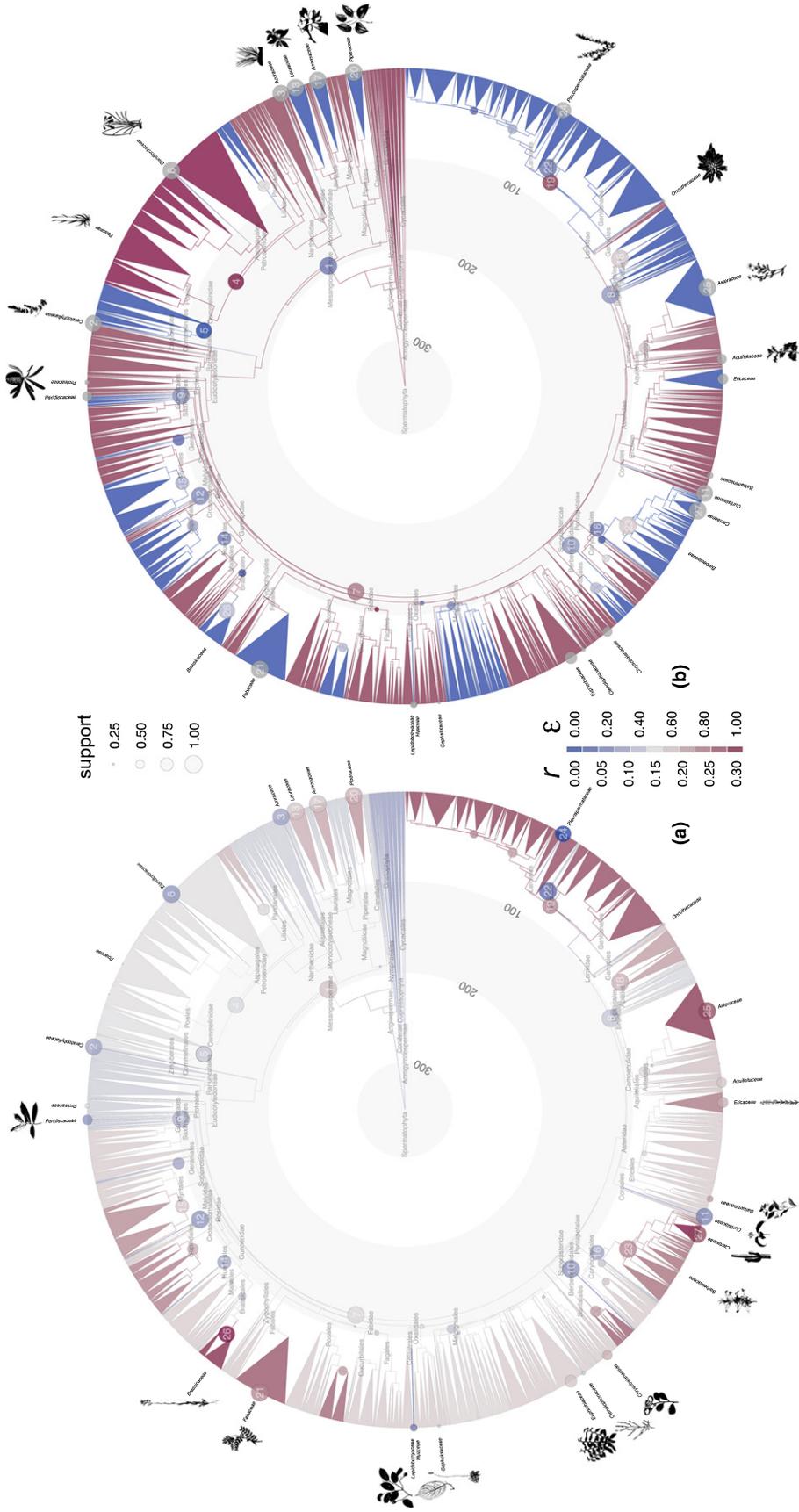
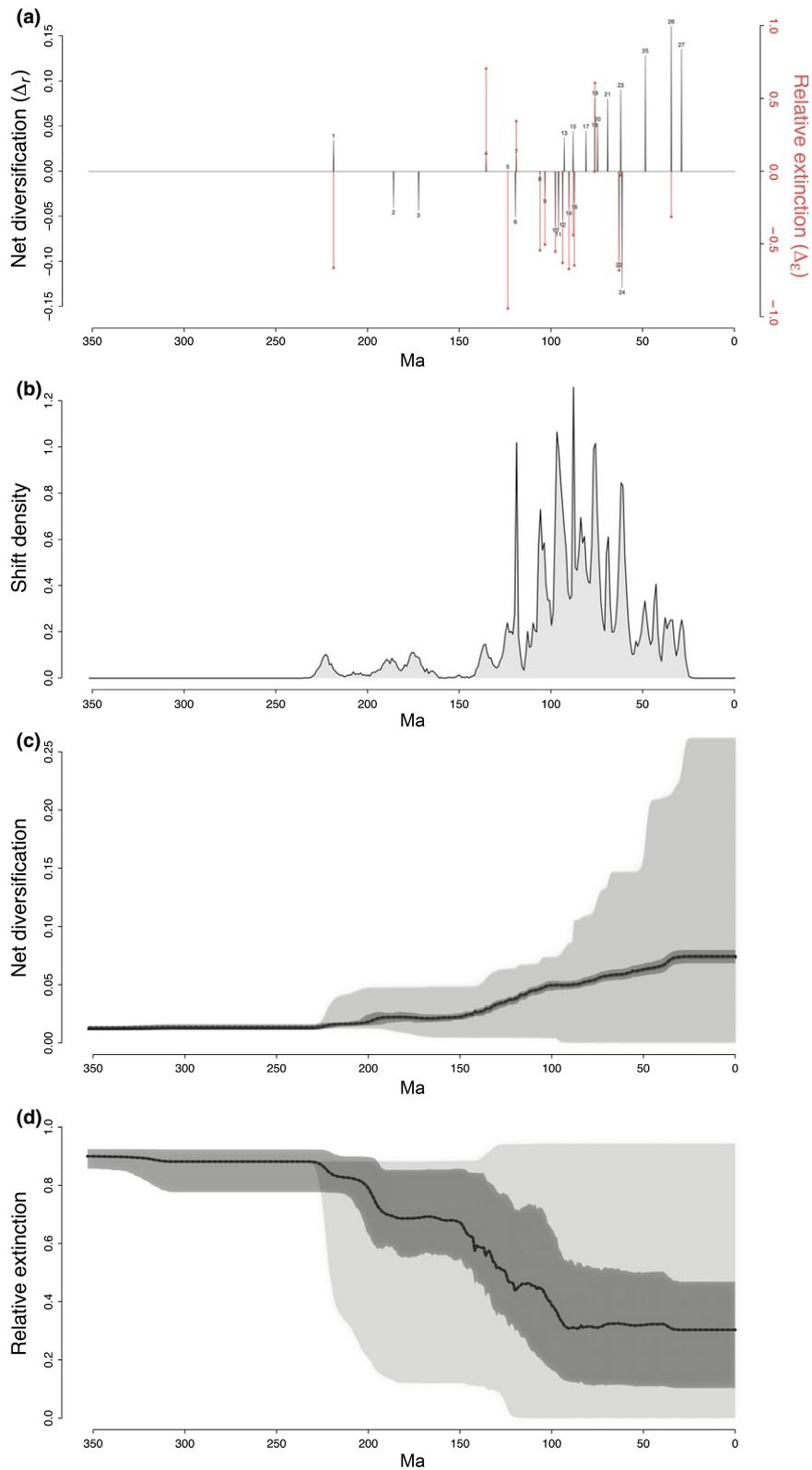


Fig. 1 Summary of MEDUSA (modeling evolutionary diversification using stepwise AIC) parameter estimates and shifts for net diversification (r ; a) and relative extinction (ϵ ; b), plotted on the rate-smoothed maximum likelihood estimate (MLE) of the phylogeny (see text). Collapsed clades are drawn as triangles that are proportional to family-level species richnesses. Colors of branches correspond to parameter estimates (see keys for r and ϵ) for the MLE; colors within circles reflect magnitudes of change in r (a) or ϵ (b) from the immediate phylogenetic background to the shifted lineage, averaged across the distribution of bootstrap replicates. Bolder colors are associated with more extreme values of r and ϵ or with greater magnitudes of shifts (Δ_r and Δ_ϵ). In (b), gray within circles is used where a (unresolved) family is estimated to have independent rates of diversification. Estimating relative extinction is problematic in these circumstances because of the lack of resolution in the subtree (Rabosky *et al.*, 2007). The size of circles indicates shift support as the relative frequency of trees in the distribution for which MEDUSA recovers a given shifted lineage (see 'support' key). For clarity, support for shifts at unresolved clades is indicated at the outermost extents of those clades (e.g. Cactaceae). Major lineages are labeled internally, and families with support for a shift in greater than 5% of analyses across the distribution of trees are enlarged as tip labels. These same families have representative icons, found either in (a) or in (b). All other tip labels are minimized. Circles with numeric labels correspond to those found in Fig. 2 and Table 2.



nonrandom so that inferences based on temporal patterns using available methodologies are likely to be misleading (Cusimano & Renner, 2010).

Still, we attempted to summarize the temporal pattern of rate shifts inferred using the MEDUSA algorithm (Alfaro

et al., 2009) by plotting the timing of rate shifts calculated from 1-Myr time intervals across the trees (Fig. 2). The majority of shifts in diversification rate occur in the interval between *c.* 125 and 50 Ma (Fig. 2a,b), even when we accounted for some of the uncertainty in estimating divergence times

Fig. 2 Summary of results from MEDUSA (modeling evolutionary diversification using stepwise AIC), focusing on 27 shifts that are well supported across a distribution of bootstrap replicates (i.e. found in > 75% of bootstrap replicates). (a) Mean timing and nature of primary shifts: red points depict magnitude of change in relative extinction comparing a shifted lineage (ϵ_s) to its immediate ancestor (ϵ_a) such that $\Delta_\epsilon = \epsilon_s - \epsilon_a$. Similarly, black segments indicate ancestor–descendant changes in net diversification ($\Delta_r = r_s - r_a$). Numeric labels correspond to those in Table 2 and are ordered through time; for example, the first shift (at c. 219 million yr ago (Ma), for *Mesangiospermae*; Table 2) shows a moderate increase in net diversification and a precipitous drop in relative extinction in comparison to the immediate background rates. (b) Uncertainty in timing of primary shifts, compiled from MEDUSA analyses for each tree in the distribution of bootstrap replicates. Temporal bins of 1 million yr (Myr) were used to assess the shift density through time (see text). Broader peaks are those for which there is more uncertainty in timing (e.g. the mesangiosperm shift between c. 230 and 200 Ma). Densities may exceed 1 if multiple shifts overlap in timing. (c, d) Trends through time in estimated means (black) and ranges (gray, where dark gray represents uncertainty in estimated means across the bootstrapped distribution of trees, and light gray represents the absolute ranges across all trees) are shown for net diversification (c) and relative extinction (d). Parameter estimates for all extant lineages in a given 1-Myr bin of time, and across all bootstrap replicates, were used to generate panels (c) and (d). As a result of the inability to estimate relative extinction in clades that lack resolution (Rabosky *et al.* 2007), estimates associated with shifts in unresolved clades were excluded from (d) and (a).

(Fig. 2b). This is perhaps unsurprising as this interval coincides with the rise of many of the major lineages of angiosperms, especially among the eudicots. This period also encompasses the Cretaceous–Paleogene (K–Pg) extinction event, which paleobotanical studies suggest led to the extinction of up to c. 60% of plant species in some regions (Wilf & Johnson, 2004), and which provoked major changes in regional floras (McElwain & Punyasena, 2007). Furthermore, this time

period coincides with a nonrandom association of WGD events (Vanneste *et al.*, 2014).

We observe two general temporal trends in angiosperm diversification. The origin of *Mesangiospermae* marks the beginning of a trend toward gradually increasing rates of net diversification coupled with a trend of decreasing relative extinction (Fig. 2c,d). Following this, we further observe a trend toward increasing among-lineage heterogeneity in both net diversification and

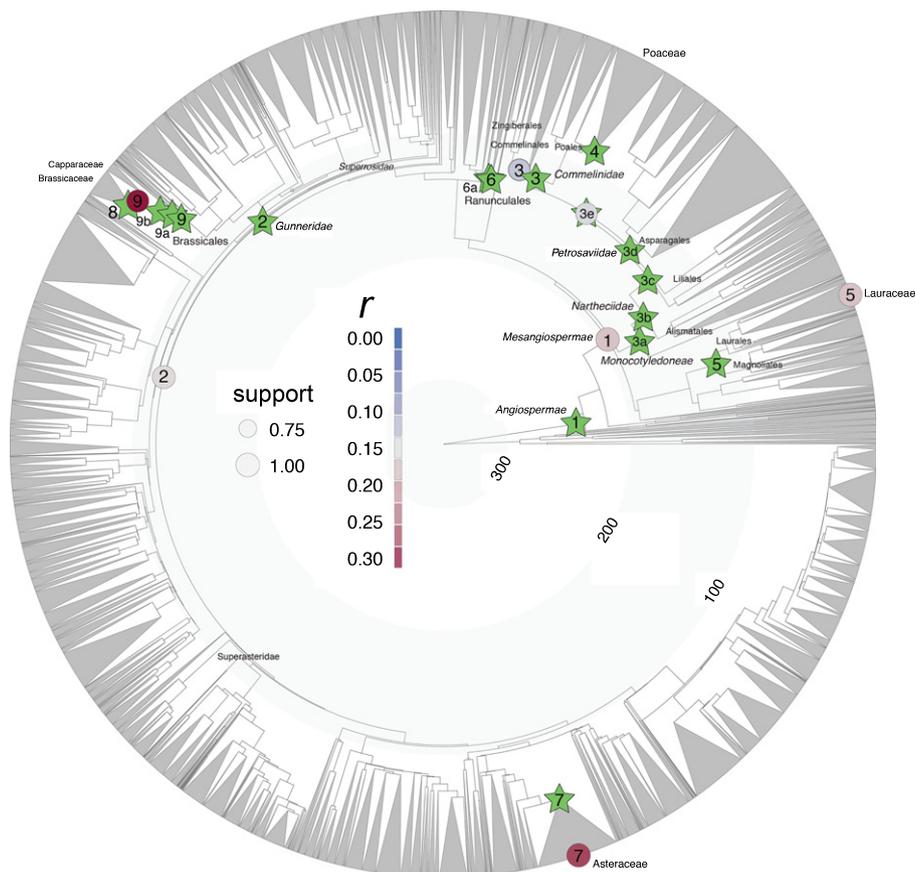


Fig. 3 Phylogenetic location of whole-genome duplication (WGD) events and the corresponding increases in net diversification that support the WGD radiation lag-time model (Schranz *et al.*, 2012). Collapsed clades are drawn as triangles proportional to family-level species richnesses. Green stars indicate WGD events, and colored circles show the location of increases in net diversification. Colors within circles reflect magnitudes of change in r (net diversification) from the immediate phylogenetic background to the shifted lineage, averaged across the distribution of bootstrap replicates; bolder colors are associated with greater magnitudes of shifts (Δ_r). The size of circles indicates shift support as the relative frequency of trees in the distribution for which MEDUSA (modeling evolutionary diversification using stepwise AIC) recovers a given shifted lineage (see ‘support’ key). For clarity, support for shifts at unresolved clades is indicated at the outermost extents of those clades (e.g. Asteraceae). Both WGDs and associated diversification rate shifts are numbered according to Table 1, and major lineages are labeled.

relative extinction. This large variance suggests that methods implicitly assuming rate homogeneity across the tree may be inappropriate for this data set (but see Morlon *et al.*, 2011). Many hypotheses have been put forward to explain the increase in diversification rates among angiosperm lineages (Stebbins, 1970, 1971, 1974, 1981; reviewed in Gorelick, 2001; Davies & Barraclough, 2004; Vamossi & Vamossi, 2011; reviewed in Augusto *et al.*, 2014), and we do not attempt to distinguish between these here, but hope that other researchers will be able to use our results to test these hypotheses and generate new ones.

Modeling approaches that consider diversification rates as a function of time (Pybus & Harvey, 2000; Rabosky, 2006; Rabosky & Lovette, 2008; Morlon *et al.*, 2010, 2011; Stadler, 2011) include some that can model variation both across clades and through time (Morlon *et al.*, 2011; Rabosky, 2014). Such approaches (reviewed in Pennell & Harmon, 2013; Pyron & Burbrink, 2013; Morlon, 2014) represent an interesting next-step analysis for angiosperm radiations, especially as more fully resolved time-trees emerge.

Negative age–diversity relationship across plant clades

Across angiosperm clades, we found a significant negative relationship between clade age and the natural logarithm of species richness (Fig. S1; see also Table S2). This result confirms the findings of previous studies (Magallón & Sanderson, 2001; Magallón & Castillo, 2009), but is notable given our more complete sampling of angiosperm diversity included here and the updated divergence time analyses. A negative or nonexistent correlation between clade age and clade richness has also been reported for a number of other groups (e.g. avian tribes: Ricklefs, 2006; major squamate clades: Ricklefs *et al.*, 2007; snakes: Pyron & Burbrink, 2012), and one recent analysis provides evidence that this may be a common feature across the tree of life (Rabosky *et al.*, 2012).

A number of recent papers (Rabosky, 2009a,b, 2010, 2012; Rabosky & Adams, 2012; Rabosky *et al.*, 2012) have argued that the absence of a positive relationship between clade age and species richness calls into question the validity of using birth–death models (and their variants). Drawing on results from simulation (Rabosky, 2009a, 2010), these authors conclude that heterogeneity in diversification rates cannot alone explain the lack of relationship between age and richness in empirical data (Rabosky, 2009a, 2010; Rabosky & Adams, 2012; Rabosky *et al.*, 2012), and that such a pattern suggests that speciation and extinction rates are influenced by ecological limits to diversity. These authors go on to claim that methods based on birth–death models (e.g. MEDUSA) are inappropriate where a positive age–diversity relationship is not supported (Rabosky & Adams, 2012; Rabosky *et al.*, 2012).

However, in a recent simulation study, Stadler *et al.* (2014) examined this question in further detail, demonstrating that when stem ages are used to estimate the relationship with diversity (as we did in this study), both negative and positive slopes can be produced by a variety of processes, including rate heterogeneous models. Additionally, Stadler *et al.* (2014) found that the results depend strongly on how higher taxa are delimited,

which is in turn determined by an odd mix of biology and psychology. Stadler *et al.*'s (2014) simulations clearly demonstrate that it is possible to generate inverse relationships between clade age and richness under models without ecological limits. Importantly, weak or negative age–diversity relationships are expected under ‘nested radiation’ conditions as revealed by our diversification rate analyses.

Our arguments against an ‘ecological limits’ interpretation of the age–diversity relationship do not mean ecological interactions have been inconsequential to the diversification of angiosperms. Indeed, there is some evidence from the fossil record (Alroy, 2008, 2010), and increasingly from reconstructed molecular phylogenies (Phillimore & Price, 2008; Rabosky & Lovette, 2008; Rabosky & Glor, 2010), that diversity-dependent speciation and extinction have played a prominent role in the generation of biodiversity. However, it is important to note that the presence or absence of ecological limits cannot be reliably judged from the sign of the relationship between lineage age and richness (Stadler *et al.*, 2014).

WGDs as a driver of nested radiations

WGDs have long been recognized as important drivers of speciation in plants (Clausen *et al.*, 1945; Stebbins, 1947, 1950), but studies explicitly linking rates of diversification and polyploidy have been relatively few (Mayrose *et al.*, 2011; Soltis *et al.*, 2014a; see Zhan *et al.*, 2014 for an example in fish), and these have focused primarily on the evolutionary consequences of recent polyploid formation, with contrasting results (see Soltis *et al.*, 2014a for a review of these analyses). With the rise of genome-scale sequencing studies and statistical methods for identifying WGD events across angiosperms, we are gaining confidence in the placement of a number of ancestral genome duplications (e.g. seed plants and angiosperms: Jiao *et al.*, 2011; monocots: Jiao *et al.*, 2014; Brassicales: Barker *et al.*, 2009; Schranz *et al.*, 2011; Brassicaceae: Haudry *et al.*, 2013; Kagale *et al.*, 2014; Asteraceae: Barker *et al.*, 2008), but tests of the effect of ancient WGDs on rates of diversification have not been performed before now. Schranz *et al.* (2012) presented a verbal model – the WGD radiation lag-time model – that hypothesizes that WGDs often result in diversification rate increases, but following a delay (of potentially millions of years), resulting in tree imbalance following the duplication event. This suggests that WGDs promote, but are not sufficient to cause, increased diversification.

We investigated the link between upticks in rates of diversification and nine well-documented ancient WGDs that we were able to place on our tree, incorporating uncertainty in the placement of several of these (Table 1), and demonstrated significant statistical support for a nonrandom association between WGD events and a delayed increase in rates of diversification (Fig. 3). Explanations for the lag in diversification rate increases following WGD remain unclear, and could simply reflect early extinction events – radiations that do not involve WGD are often asymmetrical too (Mooers & Heard, 1997). Nevertheless, Schranz *et al.* (2012) provide a potential scenario to explain the observed lag pattern between WGDs and subsequent upticks in

diversification. Following their model, once a WGD event occurs, it then contributes to the subsequent evolution of a key defining trait(s). Importantly, these traits do not necessarily arise immediately or spur diversification, but are assumed to be the result of increased evolutionary potential as a result of polyploidization. Noting an observed pattern of species-poor lineages geographically restricted to the center of origin of a lineage and widespread, species-rich sister clades, their model implies that initial diversification events following the WGD occur in the region that is the center of origin, and only after millions of years, a dispersal event(s) sparks the crown group radiation. Schranz *et al.* (2012) postulate that this dispersal may not be initially driven by the WGD event, but could be in response to any number of factors, including climatic or geological factors, or plant–animal interactions (e.g. herbivory or pollination). Nonetheless, a key innovation resulting from WGD may ultimately underlie the long-term success of that lineage, but such key innovations may be difficult to discern and even more difficult to associate directly with WGDs. However, the creative roles of hybridization and polyploidization have long been recognized (Levin, 1983; Arnold, 1992; Soltis *et al.*, 2014b,c). The dynamic nature of polyploid genomes – including alterations in gene content, gene number, gene arrangement, gene expression, and transposon activity – may trigger evolutionary novelty that is manifested in biochemical, cellular, morphological, or physiological features. Although most such novel features are likely to be maladaptive, rare variants or combinations of newly generated variation may serve as key innovations that spur diversification. The interplay between features arising from WGDs and both biotic and abiotic factors has probably shaped these patterns of nested radiations, but such linkages are largely unexplored.

That said, there are a few putative examples of the creative force of WGD in key-trait evolution. Schranz *et al.* (2011) demonstrated support for a causative link between two WGDs in Brassicales (the α and β duplications) and diversification of novel glucosinolate defense pathways in this clade, and while they hypothesized a putative link with diversification rate shifts in the clade, this was not explicitly tested. Here we show a clear link between a large-magnitude increase in net diversification and a decrease in the extinction fraction in the Brassicaceae + Cappara-ceae clade following the β duplication in Brassicales (Figs 1, 3; Tables 1, 2). Likewise, in the basal-eudicot lineage Ranunculales, Pabón-Mora *et al.* (2013) demonstrated that selection-mediated asymmetric sequence diversification, the generation of novel motifs, differences in codon substitutions, and putative differences in protein–protein interactions of ranunculiid-specific duplications in floral MADS-box genes explain the functional differences among these gene copies across Ranunculales. Finally, a recent comparative analysis of 41 whole-genome sequences from angiosperms revealed a striking nonrandom association between successful paleopolyploidization events and the Cretaceous–Paleogene (K–Pg) extinction event *c.* 66 Ma (Vanneste *et al.*, 2014). Vanneste *et al.* hypothesized that successful establishment of polyploid lineages may be promoted during times of environmental stress, and that the evolutionary potential

of polyploids in the context of dramatic environmental and ecological perturbations at the time of WGD may help explain the stark contrast in the proposed evolutionary fates of polyploids. While these examples are compelling, more mechanistic links between WGD, shifts in diversification rates, and the evolution of novel key-traits are needed, and therefore, we echo Schranz *et al.*'s (2012) call for 'ecological and genomic comparisons of species-rich crown groups with their species-poor sister-groups that share a common WGD history as a means to provide new insights into the radiation lag-time paradox.'

Summary and conclusions

We show that diversification rates have been incredibly heterogeneous throughout the history of angiosperms. The pervasive pattern of radiations nested within other radiations generates a negative relationship between age and diversity across both families and orders, although we emphasize that such negative age–richness relationships are potentially consistent with other processes. Across angiosperms, stochastically changing diversification rates have led to an overall increasing rate of net diversification and declining relative extinction rates through time. Finally, we show that diversification shifts are only rarely perfectly associated with WGD events but commonly follow them after a lag period, providing statistical support for the WGD radiation lag-time hypothesis linking genome duplications and diversification rates.

There are some important caveats to our analyses. First, as molecular phylogenies include only extant species, most approaches to diversification rate analyses (Magallón & Sander-son, 2001; Rabosky, 2006; Morlon *et al.*, 2010; Stadler, 2011) using these kinds of data restrict the diversification rates to be positive throughout the clade (but see Morlon *et al.*, 2011). This assumption has received criticism from paleontologists, as periods of high extinction rates are frequently inferred from the fossil record (Quental & Marshall, 2010). Additionally, MEDUSA uses a stepwise AIC algorithm that can suffer from statistical shortcomings (Mundry & Nunn, 2009; May & Moore, 2014). We consider our analyses conservative, and these questions are worth revisiting as better methods (Mayrose *et al.*, 2011; Soltis *et al.*, 2014a) and data become available. Second, we considered only the few ancient WGD events that have been well characterized and that we could place on our tree at or above the family level, but many more such events have occurred within angiosperms. Our unresolved tree provides little information about shifts in diversification closer to the present day. Without the proper resolution in the tree, MEDUSA can never detect shifts within any of our families, even though such shifts are known to have occurred in many clades (e.g. Asteraceae and Poaceae; Smith *et al.*, 2011). However, because our null hypothesis involves random associations between shifts and WGDs, the small number of WGDs that we were able to test is unlikely to lead to false support for the lag-time hypothesis. As more fully resolved trees become available, we can expect to learn more about the patterns and processes of angiosperm diversification in relation to WGDs at a shallower phylogenetic level.

As our picture of the whole tree of life comes into focus, statistical analyses like those employed here can bring clarity to additional parts of the tree, allowing us to understand the macroevolutionary forces that have shaped diversities of clades through space and time. In angiosperms, we can explain some of the striking differences in diversity across clades as delayed responses to ancient WGDs. However, much variation still remains to be explained. Future analyses promise a deeper understanding of the shape of the angiosperm tree of life.

Acknowledgements

We would like to thank the organizers of the 'Evolutionary plant radiations: where, when, why & how?' meeting in Zurich (Switzerland) for organizing this special issue, P. Linder and three anonymous reviewers for insightful comments that improved this paper, Simon Uribe-Convers, Eric Roalson, and Stephen Smith for help with divergence time bootstrap analyses, Michael Alfaro for helpful comments on the diversification rate analyses, the NESCent 'Tempo and mode of plant trait evolution' working group for helpful discussions, and the Phylogenetics Reading Group (PuRGe) at the University of Idaho and Washington State University for the always lively discussions that have helped shape many of our ideas. We also thank the University of Idaho Institute for Bioinformatics and Evolutionary Studies Computational Resources core, supported in part by NIH/NCRR grants P20RR16448 and P20RR016454. MWP was supported by an NSERC postgraduate fellowship. This work was also supported by NSF grants DEB-0919499 and DEB-1208912 to L.J.H., and DEB-1253463 to D.C.T.

References

- Alfaro ME, Santini F, Brock C, Alamillo H, Dornburg A, Rabosky DL, Carnevale G, Harmon LJ. 2009. Nine exceptional radiations plus high turnover explain species diversity in jawed vertebrates. *Proceedings of the National Academy of Sciences, USA* 106: 13410–13414.
- Alroy J. 2008. Dynamics of origination and extinction in the marine fossil record. *Proceedings of the National Academy of Sciences, USA* 105: 11536–11542.
- Alroy J. 2010. Geographical, environmental and intrinsic biotic controls on Phanerozoic marine diversification. *Palaeontology* 53: 1211–1235.
- Amborella Genome Project. 2013. The *Amborella* genome and the evolution of flowering plants. *Science* 342: 1241089.
- Arnold ML. 1992. Natural hybridization as an evolutionary process. *Annual Review of Ecology and Systematics* 23: 237–261.
- Augusto L, Davies TJ, Delzon S, De Schrijver A. 2014. The enigma of the rise of angiosperms: can we untie the knot? *Ecology Letters* 17: 1326–1338.
- Barker MS, Kane NC, Matvienko M, Kozik A, Michelmore RW, Knapp SJ, Rieseberg LH. 2008. Multiple paleopolyploidizations during the evolution of the Compositae reveal parallel patterns of duplicate gene retention after millions of years. *Molecular Biology and Evolution* 25: 2445–2455.
- Barker MS, Vogel H, Schranz ME. 2009. Paleopolyploidy in the Brassicales: analyses of the *Cleome* transcriptome elucidate the history of genome duplications in *Arabidopsis* and other Brassicales. *Genome Biology and Evolution* 1: 391–399.
- Beaulieu JM, Tank DC, Donoghue MJ. 2013. A Southern Hemisphere origin for campanulid angiosperms, with traces of the break-up of Gondwana. *BMC Evolutionary Biology* 13: 80.
- Bell CD, Soltis DE, Soltis PS. 2010. The age and diversification of the angiosperms re-visited. *American Journal of Botany* 97: 1296–1303.
- Berendse F, Scheffer M. 2009. The angiosperm radiation revisited, an ecological explanation for Darwin's 'abominable mystery'. *Ecology Letters* 12: 865–872.
- Bowers JE, Chapman BA, Rong J, Paterson AH. 2003. Unravelling angiosperm genome evolution by phylogenetic analysis of chromosomal duplication events. *Nature* 422: 433–438.
- Bremer K, Friis EM, Bremer B. 2004. Molecular phylogenetic dating of asterid flowering plants shows early Cretaceous diversification. *Systematic Biology* 53: 496–505.
- Bremer K, Gustafsson MH. 1997. East Gondwana ancestry of the sunflower alliance of families. *Proceedings of the National Academy of Sciences, USA* 94: 9188–9190.
- Brodribb TJ, Feild TS. 2010. Leaf hydraulic evolution led a surge in leaf photosynthetic capacity during early angiosperm diversification. *Ecology Letters* 13: 175–183.
- Burnham K, Anderson DR. 2002. *Model selection and multimodel inference: a practical information-theoretic approach*. New York, NY, USA: Springer.
- Cantino PD, Doyle JA, Graham SW, Judd WS, Olmstead RG, Soltis DE, Soltis PS, Donoghue MJ. 2007. Towards a phylogenetic nomenclature of Tracheophyta. *Taxon* 56: 822–846.
- Clarke JT, Warnock RCM, Donoghue PCJ. 2011. Establishing a time-scale for plant evolution. *New Phytologist* 192: 266–301.
- Clausen J, Keck DD, Hiesey WM. 1945. *Experimental studies on the nature of species*. II. Washington, DC, USA: Carnegie Institution of Washington.
- Crane PR, Friis EM, Pedersen KR. 1995. The origin and early diversification of angiosperms. *Nature* 374: 27–33.
- Cui L, Wall PK, Leebens-Mack JH, Lindsay BG, Soltis DE, Doyle JJ, Soltis PS, Carlson JE, Arumuganathan K, Barakat A *et al.* 2006. Widespread genome duplications throughout the history of flowering plants. *Genome Research* 16: 738–749.
- Cusimano N, Renner SS. 2010. Slowdowns in diversification rates from real phylogenies may not be real. *Systematic Biology* 59: 458–464.
- Darwin C. 1903. Letter 395. To J.D. Hooker. In: Darwin F, Seward AC, eds. *More letters of Charles Darwin, vol II*. London, UK: John Murray, 20–22.
- Davies T, Barraclough T. 2004. Darwin's abominable mystery: Insights from a supertree of the angiosperms. *Proceedings of the National Academy of Sciences, USA* 101: 1904–1909.
- Dilcher DL. 2001. Paleobotany: some aspects of non-flowering and flowering plant evolution. *Taxon* 50: 697–711.
- Feild TS, Brodribb TJ, Iglesias A, Chatelet DS, Baresch A, Upchurch GRJ, Gomez B, Mohr BAR, Coiffard C, Kvacek J *et al.* 2011. Fossil evidence for Cretaceous escalation in angiosperm leaf vein evolution. *Proceedings of the National Academy of Sciences, USA* 108: 8363–8366.
- Fiz-Palacios O, Schneider H, Heinrichs J, Savolainen V. 2011. Diversification of land plants: insights from a family-level phylogenetic analysis. *BMC Evolutionary Biology* 11: 341.
- Gorelick R. 2001. Did insect pollination cause increased seed plant diversity? *Biological Journal of the Linnean Society* 74: 407–427.
- Haudry A, Platts AE, Vello E, Hoen DR, Leclercq M, Williamson RJ, Forczek E, Joly-Lopez Z, Steffen JG, Hazzouri KM *et al.* 2013. An atlas of over 90 000 conserved noncoding sequences provides insight into crucifer regulatory regions. *Nature Genetics* 45: 891–898.
- Jiao Y, Leebens-Mack J, Ayyampalayam S, Bowers JE, McKain MR, McNeal J, Rolf M, Ruzicka DR, Wafula E, Wickett NJ *et al.* 2012. A genome triplication associated with early diversification of the core eudicots. *Genome Biology* 13: R3.
- Jiao Y, Li J, Tang H, Paterson AH. 2014. Integrated syntenic and phylogenomic analyses reveal an ancient genome duplication in monocots. *Plant Cell* 26: 2792–2802.
- Jiao Y, Wickett NJ, Ayyampalayam S, Chanderbali AS, Landherr L, Ralph PE, Tomsho LP, Hu Y, Liang H, Soltis PS *et al.* 2011. Ancestral polyploidy in seed plants and angiosperms. *Nature* 473: 97–100.
- Joppa LN, Roberts DL, Pimm SL. 2011. How many species of flowering plants are there? *Proceedings of the Royal Society B: Biological Sciences* 278: 554–559.

- Judd WS, Campbell CS, Kellogg EA, Stevens PF, Donoghue MJ. 2007. *Plant systematics: a phylogenetic approach*. Sunderland, MA, USA: Sinauer Associates.
- Kagale S, Robinson SJ, Nixon J, Xiao R, Huebert T, Condie J, Kessler D, Clarke WE, Edger PP, Links MG *et al.* 2014. Polyploid evolution of the Brassicaceae during the Cenozoic Era. *Plant Cell* 26: 2777–2791.
- Kim KJ. 2005. Two chloroplast DNA inversions originated simultaneously during the early evolution of the sunflower family (Asteraceae). *Molecular Biology and Evolution* 22: 1783–1792.
- Labandeira CC. 2010. The pollination of mid Mesozoic seed plants and the early history of long-proboscid insects. *Annals of the Missouri Botanical Garden* 97: 469–513.
- Levin DA. 1983. Polyploidy and novelty in flowering plants. *The American Naturalist* 122: 1–25.
- Lidgard S, Crane PR. 1990. Angiosperm diversification and Cretaceous floristic trends: a comparison of palynofloras and leaf macrofloras. *Paleobiology* 16: 77–93.
- Magallón S, Castillo A. 2009. Angiosperm diversification through time. *American Journal of Botany* 96: 349–365.
- Magallón S, Sanderson MJ. 2001. Absolute diversification rates in angiosperm clades. *Evolution* 55: 1762–1780.
- May MR, Moore BR. 2014. How well can we detect shifts in rates of lineage diversification? A simulation study of sequential AIC methods. *bioRxiv*. doi: 10.1101/011452.
- Mayrose I, Zhan SH, Rothfels CJ, Magnuson-Ford K, Barker MS, Rieseberg LH, Otto SP. 2011. Recently formed polyploid plants diversify at lower rates. *Science* 333: 1257.
- McElwain JC, Punyasena SW. 2007. Mass extinction events and the plant fossil record. *Trends in Ecology & Evolution* 22: 548–557.
- Mooers AO, Heard SB. 1997. Inferring evolutionary process from phylogenetic tree shape. *Quarterly Review of Biology* 72: 31–54.
- Morlon H. 2014. Phylogenetic approaches for studying diversification. *Ecology Letters* 17: 508–525.
- Morlon H, Parsons TL, Plotkin JB. 2011. Reconciling molecular phylogenies with the fossil record. *Proceedings of the National Academy of Sciences, USA* 108: 16327–16332.
- Morlon H, Potts MD, Plotkin JB. 2010. Inferring the dynamics of diversification: a coalescent approach (Harvey PH, ed.). *PLoS Biology* 8: e1000493.
- Mundry R, Nunn CL. 2009. Stepwise model fitting and statistical inference: turning noise into signal pollution. *The American Naturalist* 173: 119–123.
- Ott M, Zola J, Stamatakis A, Aluru S. 2007. Large-scale maximum likelihood-based phylogenetic analysis on the IBM BlueGene/L. *Proceedings of the 2007 ACM/IEEE Conference on Supercomputing* 4, 1–11.
- Pabón-Mora N, Hidalgo O, Gleissberg S, Litt A. 2013. Assessing duplication and loss of APETALA1/FRUITFULL homologs in Ranunculales. *Frontiers in Plant Science* 4: 1–14.
- Paterson AH, Wang X, Li J, Tang H. 2012. Ancient and recent polyploidy in monocots. In: Soltis DE, Soltis PS, eds. *Polyploidy and genome evolution*. Heidelberg, Germany: Springer, 93–108.
- Pennell MW, Eastman JM, Slater GJ, Brown JW, Uyeda JC, FitzJohn RG, Alfaro ME, Harmon LJ. 2014. geiger v2.0: an expanded suite of methods for fitting macroevolutionary models to phylogenetic trees. *Bioinformatics* 30: 2216–2218.
- Pennell MW, Harmon LJ. 2013. An integrative view of phylogenetic comparative methods: connections to population genetics, community ecology, and paleobiology. *Annals of the New York Academy of Sciences* 1289: 90–105.
- Phillimore AB, Price TD. 2008. Density-dependent cladogenesis in birds. *PLoS Biology* 6: e71.
- Pybus OG, Harvey PH. 2000. Testing macro-evolutionary models using incomplete molecular phylogenies. *Proceedings of the Royal Society B-Biological Sciences* 267: 2267–2272.
- Pyron RA, Burbrink FT. 2012. Extinction, ecological opportunity, and the origins of global snake diversity. *Evolution* 66: 163–178.
- Pyron RA, Burbrink FT. 2013. Phylogenetic estimates of speciation and extinction rates for testing ecological and evolutionary hypotheses. *Trends in Ecology & Evolution* 28: 729–736.
- Quental TB, Marshall CR. 2010. Diversity dynamics: molecular phylogenies need the fossil record. *Trends in Ecology & Evolution* 25: 434–441.
- Rabosky DL. 2006. LASER: a maximum likelihood toolkit for detecting temporal shifts in diversification rates from molecular phylogenies. *Evolutionary Bioinformatics Online* 2: 273–276.
- Rabosky DL. 2009a. Ecological limits on clade diversification in higher taxa. *The American Naturalist* 173: 662–674.
- Rabosky DL. 2009b. Ecological limits and diversification rate: alternative paradigms to explain the variation in species richness among clades and regions. *Ecology Letters* 12: 735–743.
- Rabosky DL. 2010. Primary controls on species richness in higher taxa. *Systematic Biology* 59: 634–645.
- Rabosky DL. 2012. Testing the time-for-speciation effect in the assembly of regional biotas. *Methods in Ecology and Evolution* 3: 224–233.
- Rabosky DL. 2014. Automatic detection of key innovations, rate shifts, and diversity-dependence on phylogenetic trees. *PLoS ONE* 9: e89543.
- Rabosky DL, Adams DC. 2012. Rates of morphological evolution are correlated with species richness in salamanders. *Evolution* 66: 1807–1818.
- Rabosky DL, Donnellan SC, Talaba AL, Lovette IJ. 2007. Exceptional among-lineage variation in diversification rates during the radiation of Australia's most diverse vertebrate clade. *Proceedings of the Royal Society B: Biological Sciences* 274: 2915–2923.
- Rabosky DL, Glor RE. 2010. Equilibrium speciation dynamics in a model adaptive radiation of island lizards. *Proceedings of the National Academy of Sciences, USA* 107: 22178–22183.
- Rabosky DL, Lovette IJ. 2008. Density-dependent diversification in North American wood warblers. *Proceedings of the Royal Society B: Biological Sciences* 275: 2363–2371.
- Rabosky DL, Slater GJ, Alfaro ME. 2012. Clade age and species richness are decoupled across the eukaryotic tree of life. *PLoS Biology* 10: e1001381.
- Ricklefs RE. 2006. Time, species, and the generation of trait variance in clades. *Systematic Biology* 55: 151–159.
- Ricklefs RE, Losos JB, Townsend TM. 2007. Evolutionary diversification of clades of squamate reptiles. *Journal of Evolutionary Biology* 20: 1751–1762.
- Sanderson MJ, Donoghue MJ. 1994. Shifts in diversification rate with the origin of angiosperms. *Science* 264: 1590–1593.
- Schranz ME, Edger PP, Pires JC, van Dam NM, Wheat CW. 2011. Comparative genomics in the Brassicales: ancient genome duplications, glucosinolate diversification and Pierinae herbivore radiation. In: Edwards D, Batley J, Parkin I, Kole C, eds. *Genetics, genomics and breeding of crop plants*. Enfield, NH, USA: Science Publishers, 206–218.
- Schranz ME, Mohammadin S, Edger PP. 2012. Ancient whole genome duplications, novelty and diversification: the WGD Radiation Lag-Time Model. *Current Opinion in Plant Biology* 15: 147–153.
- Smith SA, Beaulieu JM, Donoghue MJ. 2010. An uncorrelated relaxed-clock analysis suggests an earlier origin for flowering plants. *Proceedings of the National Academy of Sciences, USA* 107: 5897–5902.
- Smith SA, Beaulieu JM, Stamatakis A, Donoghue MJ. 2011. Understanding angiosperm diversification using small and large phylogenetic trees. *American Journal of Botany* 98: 404–414.
- Smith SA, O'Meara BC. 2012. treePL: divergence time estimation using penalized likelihood for large phylogenies. *Bioinformatics* 28: 2689–2690.
- Soltis DE, Albert VA, Leebens-Mack J, Bell CD, Paterson AH, Zheng C, Sankoff D, dePamphilis CW, Wall PK, Soltis PS. 2009. Polyploidy and angiosperm diversification. *American Journal of Botany* 96: 336–348.
- Soltis DE, Segovia-Salcedo MC, Jordon-Thaden I, Majure L, Miles NM, Mavrodiev EV, Mei W, Cortez MB, Soltis PS, Gitzendanner MA. 2014a. Are polyploids really evolutionary dead-ends (again)? A critical reappraisal of Mayrose *et al.* (2011). *New Phytologist* 202: 1105–1117.
- Soltis DE, Smith SA, Cellinese N, Wurdack KJ, Tank DC, Brockington SF, Refulio-Rodriguez NF, Walker JB, Moore MJ, Carlswald BS *et al.* 2011. Angiosperm phylogeny: 17 genes, 640 taxa. *American Journal of Botany* 98: 704–730.

- Soltis PS, Soltis DE, Chase MW, Endress P, Crane PR. 2004. The diversification of flowering plants. In: Cracraft J, Donoghue MJ, eds. *The tree of life*. New York, NY, USA: Oxford University Press, 154–167.
- Soltis DE, Visger CJ, Soltis PS. 2014b. The polyploidy revolution then...and now: Stebbins revisited. *American Journal of Botany* 101: 1057–1078.
- Soltis PS, Liu X, Marchant DB, Visger CJ, Soltis DE. 2014c. Polyploidy and novelty: Gottlieb's legacy. *Philosophical Transactions of the Royal Society of London Series B: Biological Sciences* 369: 20130351.
- Stadler T. 2011. Inferring speciation and extinction processes from extant species data. *Proceedings of the National Academy of Sciences, USA* 108: 16145–16146.
- Stadler T, Rabosky DL, Ricklefs RE, Bokma F. 2014. On age and species richness of higher taxa. *The American Naturalist* 184: 447–455.
- Stamatakis A. 2006. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* 22: 2688–2690.
- Stebbins GL. 1947. Types of polyploids; their classification and significance. *Advances in Genetics* 1: 403–429.
- Stebbins GL. 1950. *Variation and evolution in plants*. New York, NY, USA: Columbia University Press.
- Stebbins GL. 1970. Adaptive radiation of reproductive characteristics in angiosperms, I: pollination mechanisms. *Annual Review of Ecology and Systematics* 1: 307–326.
- Stebbins GL. 1971. Adaptive radiation of reproductive characteristics in angiosperms, II: seeds and seedlings. *Annual Review of Ecology and Systematics* 2: 237–260.
- Stebbins GL. 1974. *Flowering plants: evolution above the species level*. Cambridge, MA, USA: Harvard University Press.
- Stebbins GL. 1981. Why are there so many species of flowering plants? *BioScience* 31: 573–577.
- Stevens PF. (2001 onwards). *Angiosperm phylogeny website*. Version 12, July 2012 [and more or less continuously updated since]. [WWW document] URL <http://www.mobot.org/MOBOT/research/APweb/> [accessed 21 July 2014].
- The Angiosperm Phylogeny Group. 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.
- The Plant List. 2010. *Version 1.0*. [WWW document] URL <http://www.theplantlist.org/1/> [accessed 23 April 2012].
- Vamosi JC, Vamosi SM. 2011. Factors influencing diversification in angiosperms: at the crossroads of intrinsic and extrinsic traits. *American Journal of Botany* 98: 460–471.
- Van de Peer Y, Fawcett JA, Proost S, Sterck L, Vandepoele K. 2009. The flowering world: a tale of duplications. *Trends in Plant Science* 14: 680–688.
- Van de Peer Y, Maere S, Meyer A. 2010. 2R or not 2R is not the question anymore. *Nature Reviews Genetics* 11: 166.
- Vanneste K, Baele G, Maere S, Van de Peer Y. 2014. Analysis of 41 plant genomes supports a wave of successful genome duplications in association with the Cretaceous–Paleogene boundary. *Genome Research* 24: 1334–1347.
- Wilf P, Johnson KR. 2004. Land plant extinction at the end of the Cretaceous: a quantitative analysis of the North Dakota megafossil record. *Paleobiology* 30: 347–368.
- Zanne AE, Tank DC, Cornwell WK, Eastman JM, Smith SA, FitzJohn RG, McGlenn DJ, O'Meara BC, Moles AT, Reich PB *et al.* 2014. Three keys to the radiation of angiosperms into freezing environments. *Nature* 506: 89–92.
- Zhan SH, Glick L, Tsigenopoulos CS, Otto SP, Mayrose I. 2014. Comparative analysis reveals that polyploidy does not decelerate diversification in fish. *Journal of Evolutionary Biology* 27: 391–403.

Supporting Information

Additional supporting information may be found in the online version of this article.

Fig. S1 Observed relationship between lineage age and species richness for families and orders.

Table S1 MEDUSA estimates for shifts in diversification finding support in at least one analysis across the distribution of bootstrap replicate trees

Table S2 Stem ages for families retained for diversification analyses

Please note: Wiley Blackwell are not responsible for the content or functionality of any supporting information supplied by the authors. Any queries (other than missing material) should be directed to the *New Phytologist* Central Office.