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

David C. Tank1,2, Jonathan M. Eastman1,2, Matthew W. Pennell1,2, Pamela S. Soltis3, Douglas E. Soltis3,4, Cody E. Hinchliff5, Joseph W. Brown5, Emily B. Sessa4 and Luke J. Harmon1,2

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

- 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 & Barracloough, 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 & Field, 2010; Labandeira, 2010; Field 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 polyploidy – 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 polyploidy, 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, psbBTNH, 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 \( (c = \mu/\lambda) \), where \( \lambda \) is the speciation rate (birth) and \( \mu \) 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 \( (n = 1024; \text{Supporting Information Table S1}) \).

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 \( (\tau) \) and relative extinction \( (\kappa) \) 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 \( \epsilon \) event; Jiao et al., 2011) and the subsequent radiation of the core eudicots (the \( \gamma \) 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 \( \tau \) event). It is important to note that this event is separate from the previously characterized \( \rho \) and \( \sigma \) events in monocots (Paterson et al., 2012). However, because no noncommelinid monocot genomes have been sampled thus far, the phylogenetic placement of the \( \tau \) event is more uncertain. Jiao et al. (2014) hypothesized that \( \tau \) 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
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 Magnoliaceae and Laurales, and an additional duplication in Ranunculaceae. 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 Magnoliidae and six species of Laurales, initial plots of the divergence of duplicate gene pairs in terms of substitutions per synonymous site per year (Ks), indicate a shared WGD that is unique to these orders. Likewise, Ks analyses from seven species of Ranunculales revealed a shared WGD as suggested by Cui et al. (2006). However, because polyploidization 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 synteny and phylogenomic techniques (e.g. as in Jiao et al., 2014). Nevertheless, here we include a WGD event shared by Magnoliidae and Laurales, as well as a Ranunculaceae WGD (Table 1). For Ranunculaceae, additional evidence from phylogenetic analyses of MADS-box genes in Ranunculaceae supports a duplication early in the diversification of Ranunculaceae, 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 Brassicaeae (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; Brassicaeae, 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 + Brassicaeae 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

<table>
<thead>
<tr>
<th>WGD description</th>
<th>Reference</th>
<th>Nearest r increase</th>
<th>n-dist</th>
<th>Myr</th>
</tr>
</thead>
<tbody>
<tr>
<td>Angiospermae (c)</td>
<td>Jiao et al. (2011)</td>
<td>Mesangiospermae</td>
<td>3</td>
<td>49.2</td>
</tr>
<tr>
<td>Gunneridae (γ)</td>
<td>Jiao et al. (2012)</td>
<td>MRCA of (Superrosidae + Superasteridae)</td>
<td>1</td>
<td>0.6</td>
</tr>
<tr>
<td>Commelinidae (α)</td>
<td>Jiao et al. (2014)</td>
<td>MRCA of (Arecaceae) and (Commelinidae + Zingiberales)</td>
<td>1</td>
<td>2.3</td>
</tr>
<tr>
<td>Monocotyledonaceae (γ)</td>
<td>Jiao et al. (2014)</td>
<td>MRCA of (Commelinidae + Asparagales)</td>
<td>4</td>
<td>39.4</td>
</tr>
<tr>
<td>Nartheciidae (γ)</td>
<td>Jiao et al. (2014)</td>
<td>MRCA of (Commelinidae + Asparagales)</td>
<td>3</td>
<td>26.5</td>
</tr>
<tr>
<td>Petrosaviidae + Dioscoreaceae (γ)</td>
<td>Jiao et al. (2014)</td>
<td>MRCA of (Commelinidae + Asparagales)</td>
<td>2</td>
<td>5.7</td>
</tr>
<tr>
<td>Petrosaviidae (γ)</td>
<td>Jiao et al. (2014)</td>
<td>MRCA of (Commelinidae + Asparagales)</td>
<td>1</td>
<td>3.0</td>
</tr>
<tr>
<td>Commelinidae + Asparagales (γ)</td>
<td>Jiao et al. (2014)</td>
<td>MRCA of (Commelinidae + Asparagales)</td>
<td>0</td>
<td>0.0</td>
</tr>
<tr>
<td>Poaceae (ρ, σ)</td>
<td>Paterson et al. (2012); Jiao et al. (2014)</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Magnoliaceae + Laurales</td>
<td>Cui et al. (2006)</td>
<td>Lauraceae</td>
<td>4</td>
<td>26.8</td>
</tr>
<tr>
<td>Papaveraceae + Ranunculaceae</td>
<td>Pabón-Mora et al. (2013)</td>
<td>None</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Asteraceae</td>
<td>Barker et al. (2008)</td>
<td>Asteraceae</td>
<td>0</td>
<td>0.0</td>
</tr>
<tr>
<td>Brassicaeae (α)</td>
<td>Haudry et al. (2013); Kagale et al. (2014)</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Core Brassicaeae</td>
<td>Barker et al. (2009)</td>
<td>MRCA of (Capparaceae + Brassicaeae)</td>
<td>3</td>
<td>35.9</td>
</tr>
<tr>
<td>[Resedaceae + Brassicaeae] (β)</td>
<td>Barker et al. (2009)</td>
<td>MRCA of (Capparaceae + Brassicaeae)</td>
<td>1</td>
<td>19.9</td>
</tr>
<tr>
<td>Brassicaeae + Bataceae (β)</td>
<td>Schranz et al. (2011)</td>
<td>MRCA of (Capparaceae + Brassicaeae)</td>
<td>2</td>
<td>28.8</td>
</tr>
</tbody>
</table>
Research 5

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 × 2 × 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 delay is 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 through 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 e) 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 upicks (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

<table>
<thead>
<tr>
<th>Description</th>
<th>Label</th>
<th>Support</th>
<th>$Ma$</th>
<th>$r$</th>
<th>$e$</th>
<th>$\Delta_r$</th>
<th>$\Delta_e$</th>
</tr>
</thead>
<tbody>
<tr>
<td>MRCA of (Capparaceae + Brassicaceae)</td>
<td>26</td>
<td>0.94</td>
<td>34.7</td>
<td>0.2187</td>
<td>0.00</td>
<td>0.1614</td>
<td>-0.31</td>
</tr>
<tr>
<td>Cactaceae</td>
<td>27</td>
<td>1.00</td>
<td>29.1</td>
<td>0.2626</td>
<td>NA</td>
<td>0.1359</td>
<td>NA</td>
</tr>
<tr>
<td>Asteraceae</td>
<td>25</td>
<td>1.00</td>
<td>48.8</td>
<td>0.2101</td>
<td>NA</td>
<td>0.1292</td>
<td>NA</td>
</tr>
<tr>
<td>MRCA of (Caryophyllaceae) and (Stegnospomataceae)</td>
<td>23</td>
<td>1.00</td>
<td>62.2</td>
<td>0.1267</td>
<td>0.00</td>
<td>0.0910</td>
<td>-0.03</td>
</tr>
<tr>
<td>MRCA of (Gentianales + Solanales) and (Boraginaceae + Lamiaceae)</td>
<td>19</td>
<td>0.90</td>
<td>76.3</td>
<td>0.1287</td>
<td>0.52</td>
<td>0.0821</td>
<td>0.61</td>
</tr>
<tr>
<td>Fabaceae</td>
<td>21</td>
<td>1.00</td>
<td>69.3</td>
<td>0.1452</td>
<td>NA</td>
<td>0.0809</td>
<td>NA</td>
</tr>
<tr>
<td>Piperaceae</td>
<td>20</td>
<td>0.98</td>
<td>74.8</td>
<td>0.1012</td>
<td>NA</td>
<td>0.0534</td>
<td>NA</td>
</tr>
<tr>
<td>MRCA of (Araliaceae) and (Myodocarpaceae + Apiaceae)</td>
<td>18</td>
<td>0.99</td>
<td>76.5</td>
<td>0.0996</td>
<td>0.00</td>
<td>0.0469</td>
<td>0.00</td>
</tr>
<tr>
<td>MRCA of (Vochysiaceae + Myrtaceae) and (Melastomataceae)</td>
<td>15</td>
<td>0.76</td>
<td>88.2</td>
<td>0.1102</td>
<td>0.00</td>
<td>0.0451</td>
<td>-0.44</td>
</tr>
<tr>
<td>Annonaceae</td>
<td>17</td>
<td>0.99</td>
<td>81.2</td>
<td>0.0927</td>
<td>NA</td>
<td>0.0451</td>
<td>NA</td>
</tr>
<tr>
<td>Lauraceae</td>
<td>13</td>
<td>0.98</td>
<td>93.0</td>
<td>0.0857</td>
<td>NA</td>
<td>0.0379</td>
<td>NA</td>
</tr>
<tr>
<td>Mesangiospermae</td>
<td>1</td>
<td>1.00</td>
<td>218.8</td>
<td>0.0473</td>
<td>0.22</td>
<td>0.0349</td>
<td>-0.66</td>
</tr>
<tr>
<td>MRCA of (Superrosidae + Superasteridae)</td>
<td>7</td>
<td>0.98</td>
<td>119.2</td>
<td>0.0645</td>
<td>0.54</td>
<td>0.0174</td>
<td>0.35</td>
</tr>
<tr>
<td>MRCA of (Commelinidae + Asparagales)</td>
<td>4</td>
<td>0.92</td>
<td>135.6</td>
<td>0.0626</td>
<td>0.94</td>
<td>0.0154</td>
<td>0.71</td>
</tr>
<tr>
<td>MRCA of (Arecaceae) and (Commelinales + Zingiberales)</td>
<td>5</td>
<td>0.92</td>
<td>123.8</td>
<td>0.0629</td>
<td>0.00</td>
<td>0.0003</td>
<td>-0.94</td>
</tr>
<tr>
<td>MRCA of (Campanulidae + Lamidae)</td>
<td>8</td>
<td>1.00</td>
<td>106.2</td>
<td>0.0612</td>
<td>0.01</td>
<td>-0.0303</td>
<td>-0.54</td>
</tr>
<tr>
<td>MRCA of (Paeoniaceae) and (Altingiaceae)</td>
<td>9</td>
<td>0.98</td>
<td>103.6</td>
<td>0.0377</td>
<td>0.00</td>
<td>-0.0277</td>
<td>-0.50</td>
</tr>
<tr>
<td>MRCA of (Rhabdodendraceae) and (Simmondsiaceae)</td>
<td>16</td>
<td>0.93</td>
<td>87.5</td>
<td>0.0293</td>
<td>0.00</td>
<td>-0.0343</td>
<td>-0.65</td>
</tr>
<tr>
<td>Ceratophyllaceae</td>
<td>2</td>
<td>0.95</td>
<td>186.1</td>
<td>0.0075</td>
<td>NA</td>
<td>-0.0401</td>
<td>NA</td>
</tr>
<tr>
<td>Huerteales</td>
<td>14</td>
<td>0.78</td>
<td>90.5</td>
<td>0.0223</td>
<td>0.00</td>
<td>-0.0409</td>
<td>-0.67</td>
</tr>
<tr>
<td>Acoraceae</td>
<td>3</td>
<td>0.97</td>
<td>172.4</td>
<td>0.0040</td>
<td>NA</td>
<td>-0.0435</td>
<td>NA</td>
</tr>
<tr>
<td>Blandfordiaceae</td>
<td>6</td>
<td>1.00</td>
<td>119.6</td>
<td>0.0116</td>
<td>NA</td>
<td>-0.0505</td>
<td>NA</td>
</tr>
<tr>
<td>MRCA of (Aphloioaceae + Strasburgeriaceae)</td>
<td>12</td>
<td>1.00</td>
<td>93.9</td>
<td>0.0074</td>
<td>0.00</td>
<td>-0.0538</td>
<td>-0.63</td>
</tr>
<tr>
<td>Berberidopsidales</td>
<td>10</td>
<td>1.00</td>
<td>97.9</td>
<td>0.0043</td>
<td>0.00</td>
<td>-0.0599</td>
<td>-0.55</td>
</tr>
<tr>
<td>Curtisiaceae</td>
<td>11</td>
<td>0.96</td>
<td>96.1</td>
<td>0.0000</td>
<td>NA</td>
<td>-0.0645</td>
<td>NA</td>
</tr>
<tr>
<td>MRCA of (Montiniaeae) and (Sphenocleaceae + Hydroeleaceae)</td>
<td>22</td>
<td>1.00</td>
<td>63.1</td>
<td>0.0324</td>
<td>0.00</td>
<td>-0.0984</td>
<td>-0.68</td>
</tr>
<tr>
<td>Plocospermataceae</td>
<td>24</td>
<td>0.92</td>
<td>61.5</td>
<td>0.0000</td>
<td>NA</td>
<td>-0.1288</td>
<td>NA</td>
</tr>
</tbody>
</table>

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; e, maximum likelihood estimate of relative extinction; $\Delta_r$, mean magnitude of change in net diversification relative to immediate ancestor (see Fig. 1a for further details); $\Delta_e$, mean magnitude of change in relative extinction relative to immediate ancestor relative to immediate ancestor (see Fig. 1a for further details). Because diversification rate analyses were collapsed to the family level, relative extinction ($e$ and $\Delta_e$) cannot be calculated and is noted as NA (not applicable) for those clades. MRCA, most recent common ancestor.

diversification upicks (Fig. 3; Table 1; Angiospermae, Gunneridae, the monocot $\tau$ event in all positions except 3b, and all placements of the Brassicales $\beta$ 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 polyplody 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 & Barracough, 2004; Magallón & Castillo, 2009; Smith et al., 2011; Fitz-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 clad-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 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 ($\varepsilon$; 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 $\varepsilon$) for the MLE; colors within circles reflect magnitudes of change in $r$ (a) or $\varepsilon$ (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 $\varepsilon$ or with greater magnitudes of shifts ($\Delta r$ and $\Delta \varepsilon$). 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. 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 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 relative extinction. This 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 relative extinction. This 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 relative extinction. This period coincides with a nonrandom association of WGD events (Vanneste et al., 2014).
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; Vamosi & Vamosi, 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 (Magallon & Sanderson, 2001; Magallon & 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; Solits 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 Solits 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; Brassicaceae: 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 (Moore & 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 Brassicaeae + Capparaceae 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 rates 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


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.