New Phytologist

What explains high plant richness in East Asia? Time and diversification in the tribe Lysimachieae (Primulaceae)

Hai-Fei Yan[1,2], Cai-Yun Zhang[3], Arne A. Anderberg[4], Gang Hao[5], Xue-Jun Ge[1,2] and John J. Wiens[6]

[1]Key Laboratory of Plant Resources Conservation and Sustainable Utilization, South China Botanical Garden, Chinese Academy of Sciences, Guangzhou 510650, China; [2]Guangdong Provincial Key Laboratory of Applied Botany, South China Botanical Garden, Chinese Academy of Sciences, Guangzhou 510650, China; [3]Guangdong Institute of Chinese Materia Medica, Guangdong Food and Drug Vocational College, Guangzhou 510520, China; [4]Department of Botany, Swedish Museum of Natural History, PO Box 50007, SE-104 05 Stockholm, Sweden; [5]College of Life Sciences, South China Agricultural University, Guangzhou 510642, China; [6]Department of Ecology and Evolutionary Biology, University of Arizona, Tucson, AZ 85721, USA

Authors for correspondence:
Xue-Jun Ge
Tel: +86 20 37252551
Email: xjge@scbg.ac.cn

John J. Wiens
Tel: +1 520 6210337
Email: wiensj@email.arizona.edu

Received: *24 December 2017*
Accepted: *5 March 2018*

New Phytologist (2018) **219:** 436–448
doi: 10.1111/nph.15144

Key words: biogeography, diversification rate, Lysimachieae, phylogeny, species richness.

Summary

- What causes the disparity in biodiversity among regions is a fundamental question in biogeography, ecology, and evolutionary biology. Evolutionary and biogeographic processes (speciation, extinction, dispersal) directly determine species richness patterns, and can be studied using integrative phylogenetic approaches. However, the strikingly high richness of East Asia relative to other Northern Hemisphere regions remains poorly understood from this perspective. Here, for the first time, we test two general hypotheses (older colonization time, faster diversification rate) to explain this pattern, using the plant tribe Lysimachieae (Primulaceae) as a model system.
- We generated a new time-calibrated phylogeny for Lysimachieae (13 genes, 126 species), to estimate colonization times and diversification rates for each region and to test the relative importance of these two factors for explaining regional richness patterns.
- We find that neither time nor diversification rates alone explain richness patterns among regions in Lysimachieae. Instead, a new index that combines both factors explains global richness patterns in the group and their high East Asian biodiversity.
- Based on our results from Lysimachieae, we suggest that the high richness of plants in East Asia may be explained by a combination of older colonization times and faster diversification rates in this region.

Introduction

Why do some regions have higher species diversity than others? Explaining richness patterns is a major goal of ecology, biogeography, and evolutionary biology. For example, 25 biodiversity hotspots cover *c.* 1.4% of Earth's land surface, but include *c.* 44% of vascular plant species and *c.* 35% of species in four vertebrate groups (Myers *et al.*, 2000). However, the underlying causes of many diversity patterns remain unclear. For example, despite many studies on the latitudinal diversity gradient (e.g. Willig *et al.*, 2003; Mittelbach *et al.*, 2007) and elevational richness gradients (e.g. Hutter *et al.*, 2013), other richness patterns remain insufficiently studied and poorly understood.

Many factors may ultimately impact species richness, but only three processes can directly change species numbers in a region: speciation, extinction, and dispersal (Ricklefs, 1987). Given this perspective, two major hypotheses can explain why certain regions have more species than others (review in Wiens, 2011): (1) more time for richness to accumulate through speciation in regions that have been colonized longer (e.g. Willis, 1922); and/or (2) faster rates of net diversification (speciation–extinction) in lineages in high-diversity regions (e.g. Fischer, 1960), presumably due to ecological conditions there that increase speciation, reduce extinction, or both. These hypotheses can be tested using phylogenetic approaches for estimating biogeographic history and diversification rates. We do not treat 'carrying capacity' as a separate hypothesis since it can influence both colonization times and diversification rates (Pontarp & Wiens, 2017), but (like climate) does not directly change species numbers itself.

A striking but still unexplained richness pattern involves the remarkable species diversity of plants in East Asia (China, Japan, and Korea) relative to similar latitudes in North America and Europe (Latham & Ricklefs, 1993; Qian & Ricklefs, 2000; Qian, 2001, 2002; Adams, 2009). For example, China is among the world's most megadiverse countries (> 32 000 vascular plant species, *c.* 1.5 times more than the USA and Canada combined; Hong & Blackmore, 2013). Several explanations have been proposed to explain this diversity anomaly in terms of speciation, extinction, and colonization times. These hypotheses suggest that lineages in East Asia have lower extinction rates (i.e. during Quaternary Ice Ages; Adams, 2009), higher speciation rates (Xiang *et al.*, 2004), higher net diversification rates (Axelrod *et al.*, 1996;

Qian & Ricklefs, 2000), or were present in East Asia longer than in other regions (Latham & Ricklefs, 1993). However, few studies have used a phylogenetic approach to directly test these hypotheses. Xiang *et al.* (2004) inferred higher speciation rates in East Asia than in eastern North America by comparing species richness and substitution rates between 10 sister clades with species in both regions. Nevertheless, to our knowledge, no previous studies have directly tested both the diversification-rate and time hypotheses to explain East Asia's exceptional species richness.

Here, we use the tribe Lysimachieae (Primulaceae) to test whether high richness in East Asia is explained by older colonization times, faster diversification rates, or both. Lysimachieae offers an excellent model system. They are distributed globally, but most species occur in East Asia (171 spp.), with fewer species in other regions (< 30 each; Table 1).

We test the time and diversification-rate hypotheses using phylogenetic approaches. We first combine new molecular data with previously published data to estimate a time-calibrated phylogeny for Lysimachieae. We next perform biogeographic analyses to estimate their time of colonization in each region. We then examine the relationship between colonization time and current richness of regions to test the time hypothesis. Finally, we estimate diversification rates of Lysimachieae clades in each region and test for relationships between overall diversification rates of regions and their current richness. We also develop a new index that combines the time and diversification-rate hypotheses.

Materials and Methods

Taxonomy

We generally follow the taxonomy established by Pax & Knuth (1905) and followed in recent studies (e.g. Hu, 1994; Ståhl & Anderberg, 2004). Thus, Lysimachieae includes six genera: *Lysimachia* (211 species), *Anagallis* (31), *Trientalis* (3), *Glaux* (1), *Asterolinon* (2), and *Pelletiera* (2). Species and their regional distributions (based on previous literature) are listed in Supporting Information Table S1. However, some traditionally recognized genera have been merged into *Lysimachia* in recent studies (e.g. Banfi *et al.*, 2005; Manns & Anderberg, 2009). We followed the traditional classification, but our results also support expansion of *Lysimachia*. We therefore place these genera in quotation marks.

In total, 126 Lysimachieae species were sampled in the phylogeny, including *c.* 50% of the 250 described species (Ray, 1956; Hu & Kelso, 1996; Cholewa *et al.*, 2009; Manns & Anderberg, 2009). Our sampling includes all genera, all infrageneric taxa, and all major biogeographic regions. Five outgroup taxa (from Myrsinaceae) were also included (*Myrsine semiserrata*, *Myrsine faberi*, *Myrsine seguinii*, *Ardisia verbascifolia*, and *Embelia ribes*). These taxa are all closely related to Lysimachieae (e.g. Källersjö *et al.*, 2000; Yesson *et al.*, 2009).

Details on taxon sampling and species distributions are provided in Table S2. Species sampled per region are given in Table 1. There was a very strong relationship between the total richness of each region and the richness sampled in the tree from each region ($r^2 = 0.99$; $P < 0.0001$). Among regions, the proportion of species sampled ranged from 0.42 to 1.00 (mean: 0.73). Details on estimating regional richness (and delimiting regions) are provided later. There was also a very strong relationship between each clade's overall richness and its richness in our tree ($r^2 = 0.94$; $P < 0.0001$; Table S3).

Phylogenetic analysis

A total of 1111 new sequences were generated (84.62% of 13 genes for 126 species). Sequences included 10 plastid loci (i.e. *mat*K, *rbc*L, *trn*L-F, *trn*H-*psb*A, *rpl*20-*rps*12, *atp*F-*atp*H, *atp*B-*rbc*L, *rps*16, *trn*S-*trn*G, *rpl*32-*trn*L) and one nuclear locus (internal transcribed spacer region of nuclear ribosomal DNA; ITS, including ITS1, 5.8S ribosomal RNA gene, and ITS2). Total genomic DNA was isolated from silica-gel-dried fresh material or herbarium materials using a modified cetyltrimethylammonium bromide protocol (Doyle & Doyle, 1987). Primers for amplification and sequencing are described in Table S4. PCR amplification followed Zhang *et al.* (2012). New sequences were deposited in GenBank (Table S5).

To increase species sampling, we included sequences from 14 '*Anagallis*' species (Manns & Anderberg, 2005, 2011) and seven

Table 1 Summary of species richness, timing of colonization and weighted diversification rate of Lysimachieae in each region

Regions	Species richness (Sampled species)	First colonization time (Myr)	Summed ages (Myr)	Diversification rate			First colonization time × Diversification rate			Summed ages × Diversification rate ($\varepsilon = 0.45$)
				$\varepsilon = 0$	$\varepsilon = 0.45$	$\varepsilon = 0.90$	$\varepsilon = 0$	$\varepsilon = 0.45$	$\varepsilon = 0.90$	
East Asia (A)	171 (74)	22.2562	27.9167	0.2517	0.2147	0.1157	5.6024	4.7783	2.5740	5.9935
Europe (B)	22 (18)	28.9201	49.3975	0.0715	0.0513	0.0150	2.0681	1.4825	0.4324	2.5322
Africa (C)	28 (17)	9.5613	18.4416	0.2935	0.2322	0.0962	2.8063	2.2201	0.9200	4.2821
North America (D)	19 (13)	26.8263	44.4665	0.0793	0.0580	0.0175	2.1266	1.5546	0.4688	2.5769
Central–South America (E)	10 (8)	9.9528	16.3157	0.0578	0.0399	0.0103	0.5756	0.3968	0.1023	0.6504
Hawaii (F)	18 (14)	1.8608	3.9718	0.3171	0.2544	0.1048	0.5901	0.4735	0.1950	1.0106
Australasia (G)	6 (6)	2.3310	7.8067	0	0	0	0	0	0	0

ε, the relative extinction fraction (ε = speciation rate/extinction rate); diversification rates are in species per millions of years (Myr).

from *Lysimachia* subgenus *Lysimachiopsis* (Oh *et al.*, 2013). We used data from GenBank for all 11 genes sequenced here (with available data), and two additional plastid loci (*rpl*16 *ndh*F; Table S5). These new taxa lacked data for some markers. However, the limited missing data in our matrix (mean: 25.58% per species for 131 species) should have little impact on topology estimation or divergence dating, based on simulations and empirical analyses (e.g. Wiens & Morrill, 2011; Zheng & Wiens, 2015). Furthermore, the placements of these incomplete taxa in our tree and in previous studies (Manns & Anderberg, 2005, 2011; Oh *et al.*, 2013) were congruent.

Sequences were initially aligned using MUSCLE (Edgar, 2004), with subsequent manual adjustments using BIOEDIT (Hall, 1999). Aligned sequences were combined into a single matrix. Most data (12 759 of 13 462 base pairs) were from the chloroplast genome, which behaves as a single linked region (e.g. Son & Park, 2016). Indels were treated as missing data.

We used PARTITIONFINDER v.1.1.1 (Lanfear *et al.*, 2012) to select the optimal combination of partitioning schemes and DNA substitution models for the concatenated matrix, using the greedy algorithm. The Bayesian information criterion was used to compare model fit. The best-fitting set of partitions and models was then applied to divergence-time estimation (Tables S6–S7). The main goal here was to estimate a time-calibrated phylogeny with BEAST. Additional likelihood and Bayesian analyses (details in Methods S1) gave similar topologies to each other (Fig. S1) and to those from BEAST.

Divergence-time estimation

We used the Bayesian uncorrelated-lognormal relaxed-clock approach (Drummond *et al.*, 2006) in BEAST v.2.3.2 (Bouckaert *et al.*, 2014) on XSEDE (CIPRES Science Gateway; Miller *et al.*, 2010). We used both Yule speciation and birth–death tree priors (Drummond & Bouckaert, 2015). Both gave similar results (Table S8). We ran four independent analyses (50 million generations each, sampled every 5000). Run results were combined using LOGCOMBINER (Bouckaert *et al.*, 2014), after discarding the first 25% of generations from each analysis as burn-in. Convergence was assessed using TRACER v.1.6 (Rambaut *et al.*, 2014), based on effective sample sizes on likelihoods > 200 (Drummond *et al.*, 2006). The maximum clade credibility (MCC) tree was summarized using TREEANNOTATOR (Bouckaert *et al.*, 2014). Means and 95% highest posterior densities (HPDs) of age estimates were visualized using FIGTREE v.1.4.2 (Rambaut, 2009).

Appropriate calibrations are crucial for divergence dating (Ho & Phillips, 2009), but few reliable fossils are known for Lysimachieae. Therefore, we used one fossil calibration and two secondary calibration points (from other dating analyses). Fossil seeds similar to *Lysimachia vulgaris* are known from the latest part of the Middle Miocene (12–16 Ma; Friis, 1985). Oh *et al.* (2008) found that seeds of *L. vulgaris* and its close relatives (*Lysimachia terrestris*, *Lysimachia thyrsiflora*) are very similar (all species in Clade IV in our phylogeny). These species all have a poroid–alveolate seed-coat surface, a sponge-like outer layer, and columnar cells in the outer seed-coat layer (Oh *et al.*, 2008). We therefore used this seed fossil to set the age prior for Clade IV. Some species in the clade presently lack data on seed morphology, but the clade is strongly supported by molecular results (Hao *et al.*, 2004; this study), pollen type (Bennell & Hu, 1983), and chromosome number ($n = 21$; Löve & Löve, 1982; Probatova *et al.*, 2006). We assigned this fossil to the crown group of Clade IV (including *L. vulgaris* and relatives) using the standard lognormal prior distribution with a mean of 1, a deviation of 2, and an offset of 12 Ma. This combination yields a 95% credible interval (CI) on the lognormal distribution from 12 to 15.6 Ma, congruent with the hypothesized age of the seed fossil (Friis, 1985). The lognormal distribution assumes the clade's age will likely be close to the fossil's age but could be older (i.e. clades can be older than their oldest fossil representatives). The lognormal distribution is considered the most generally appropriate prior for fossil calibrations (Ho & Phillips, 2009).

Two secondary calibration points were also used. First, we used the crown age of Clades I–VIII (including core *Lysimachia* and '*Anagallis*'), with an estimated mean age of 28.47 Ma (95% HPD = 23.03–37.67 Ma) from Strijk *et al.* (2014). For the crown age of this node (Clade I–VIII) we used a normal prior distribution with mean of 28.5 Ma and standard deviation (SD) of 4 Ma (95% HPD = 21.9–35.1 Ma).

The other secondary calibration point was the stem age of Lysimachieae. However, previous studies estimated somewhat different ages for this clade. As explained later, we used three different ages in separate analyses, which gave similar results (Table S8). We therefore used the intermediate age for all subsequent analyses. First, Yesson *et al.* (2009) estimated the Lysimachieae stem at *c.* 30 Ma based on two datasets, but without confidence intervals. Similarly, Zanne *et al.* (2014) estimated the Lysimachieae stem at 28.2 Ma (no confidence interval). However, Strijk *et al.* (2014) estimated 37.88 Ma (95% HPD = 26.84–50.46 Ma). Renner & Schaefer (2010) estimated 41 Ma (28–52 Ma). We used three sets of priors for the Lysimachieae stem in three separate analyses, each with a normal distribution. First, we used 30 Ma (SD = 6 Ma), yielding a 95% HPD of 20–40, following Yesson *et al.* (2009) and Zanne *et al.* (2014). Second, we used 37.9 Ma (SD = 7 Ma (26.4–49.4 Ma)) following Strijk *et al.* (2014). Third, we used 41 Ma (SD = 7 Ma (29.5–52.5 Ma)) following Renner & Schaefer (2010).

In order to test the effect of the secondary calibration points, we also performed analyses using only the fossil calibration. However, all analyses yielded similar clade ages overall (see the Results section; Tables S8–S10), and subsequent analyses used the intermediate calibration point. Importantly, biogeographic and diversification analyses were also conducted on these alternative trees. The data matrix and all MCC trees files generated here are available as supporting information (Notes S1).

Incomplete species sampling might affect dating results (Stadler, 2009). We used two approaches to address these effects. First, we used the 'birth–death skyline contemporary' prior in BEAST v.2.3.2 to account for incomplete sampling, following Stadler *et al.* (2013). However, the major clades consistently estimated by other analyses were not recovered. Therefore, we did not consider these results further. Second, we subsampled species

from our tree, to see if age estimates were robust to incomplete sampling (if they are, it would suggest that our original subsampling should not strongly impact age estimates). We randomly sampled species from each major clade (using R 3.2.3; R Core Team, 2015) and performed BEAST analyses with the calibrations and tree priors described earlier. We randomly selected two or three species per large clade (≥ 10 species sampled) using R. For smaller clades, we included a single, randomly selected species. We included 26 species in total, including one outgroup, representing 10% of all Lysimachieae species. We then performed a regression between the stem ages of the 11 major clades from the subsampled analyses and the full analyses.

Biogeographic analyses

We compiled distributional data (Table S1) for all species described using literature data, assigning each species to one or more geographic regions. We generally used standard geographic definitions of continental boundaries to delineate regions, but with some differences based on patterns of endemism in Lysimachieae. Specifically, we used seven regions (Fig. 1): (A) East Asia, (B) Mediterranean region of Europe, (C) Africa, including Madagascar but excluding North Africa, (D) North America, (E) Central–South America, (F) Hawaii, and (G) Australasia. Detailed limits of regions (and their justifications) are given in Methods S2. Species could be assigned to more than one region using the biogeographic method used here (see below).

Colonization events were inferred using the dispersal–extinction–cladogenesis (DEC) model in RASP v.3.2 (Yu *et al.*, 2015) based on LAGRANGE (Ree & Smith, 2008). Analyses were performed using the MCC tree from BEAST. All ingroup taxa were included. We used a transition matrix (*Q*-matrix) to reflect changes in paleogeographic events (land bridges between continents) or dispersal ability between those areas (e.g. stepping stone dispersal or long-distance dispersal across oceans) in different time slices (Table S11). We obtained effectively identical results without these assumptions (Table S12). We found that 90.48% (114/126) of ingroup species are restricted to one region, whereas only 12 species occur in two to five regions (Table S2). Therefore, we initially ran RASP allowing ancestral species to occur in a maximum of two areas. We also used likelihood ratio, χ^2, and Akaike information criterion tests to compare the fit of different assumptions about the maximum number of regions. AIC weights (AICw) for each main clade (its crown and stem node) were interpreted as the relative probabilities of all possible ancestral states (Yu *et al.*, 2015). The most likely state for each node (region with highest proportional likelihood) was used to test the time hypothesis. The analyses were also performed using all alternative trees from BEAST (except the subsampling analyses).

Time hypothesis

We examined the relationship between regional species richness and the oldest colonization time for each region using least-squares linear regression (following Stephens & Wiens, 2003). We did not correct for phylogeny in these analyses because they were based on regions (which lack a phylogeny). We integrated the time-calibrated phylogeny with the ancestral areas estimated for each node to estimate when Lysimachieae colonized each region for the first time. However, it is impossible to infer when exactly on a given branch a region was colonized. Therefore, we simply assumed that biogeographic changes were on the middle of the branch (i.e. midpoint of crown and stem ages). Deviations from this assumption should have little impact, since the relative times are of interest here and most branches are short. Some clades had two ancestral areas (e.g. AB for Clades I–V; Fig. 1), indicating that the ancestor occurred in two regions simultaneously. In these cases, we used the clade's age as the colonization time for both regions. In other cases, an extant species represented a region's oldest colonization. We calculated the colonization time as half the species' age in these cases.

The primary analyses were based on the oldest colonization time for each region. However, multiple invasions of the same region occurred in some cases. Therefore, we examined the relationship between species richness and summed ages of colonization of each region (e.g. Stephens & Wiens, 2003; Hutter *et al.*, 2013). Thus, regions that have been colonized many times (and less recently) were given added weight. We included widespread, extant species in these calculations also.

Finally, the species in our phylogeny do not encompass all Lysimachieae species, nor all species in all regions (Table 1). Even though our species sampling in each region is proportional to its total richness (see above), and all regions have > 40% sampling (Table 1), some colonizations may still have been missed. However, unsampled colonizations may be particularly unlikely to have high richness in a region (all other things being equal, more species-rich clades are more likely to be sampled). Therefore, unsampled colonizations seem unlikely to be the oldest or most influential in a region. Furthermore, our primary analyses used the oldest colonization time, so failure to include later colonization events will have no impact on the results.

Diversification hypothesis

We examined the relationship between species richness of each region and the diversification rates of clades that occurred there. We first divided the species in each region into previously recognized, nonoverlapping, endemic higher taxa (e.g. genera, subgenera, species groups). We selected higher taxa to represent the largest monophyletic groups in each region (Table S13). If a clade had only one species in the region (regardless of its global richness), we assigned the clade a rate of zero. Our concern is only with the clade's richness in that region (and alternative approaches should have very little impact, given our weighting scheme). We then calculated the species richness for each named clade, assuming that species unsampled in our tree also belonged to these higher taxa, given their apparent monophyly based on sampled species. We estimated net diversification rates for each clade using the method-of-moments estimator for stem-group ages (Magallón & Sanderson, 2001) using GEIGER version 2.0.6 (Harmon *et al.*, 2008). The stem-group estimator is more accurate than the crown-group estimator, especially when taxon

New Phytologist

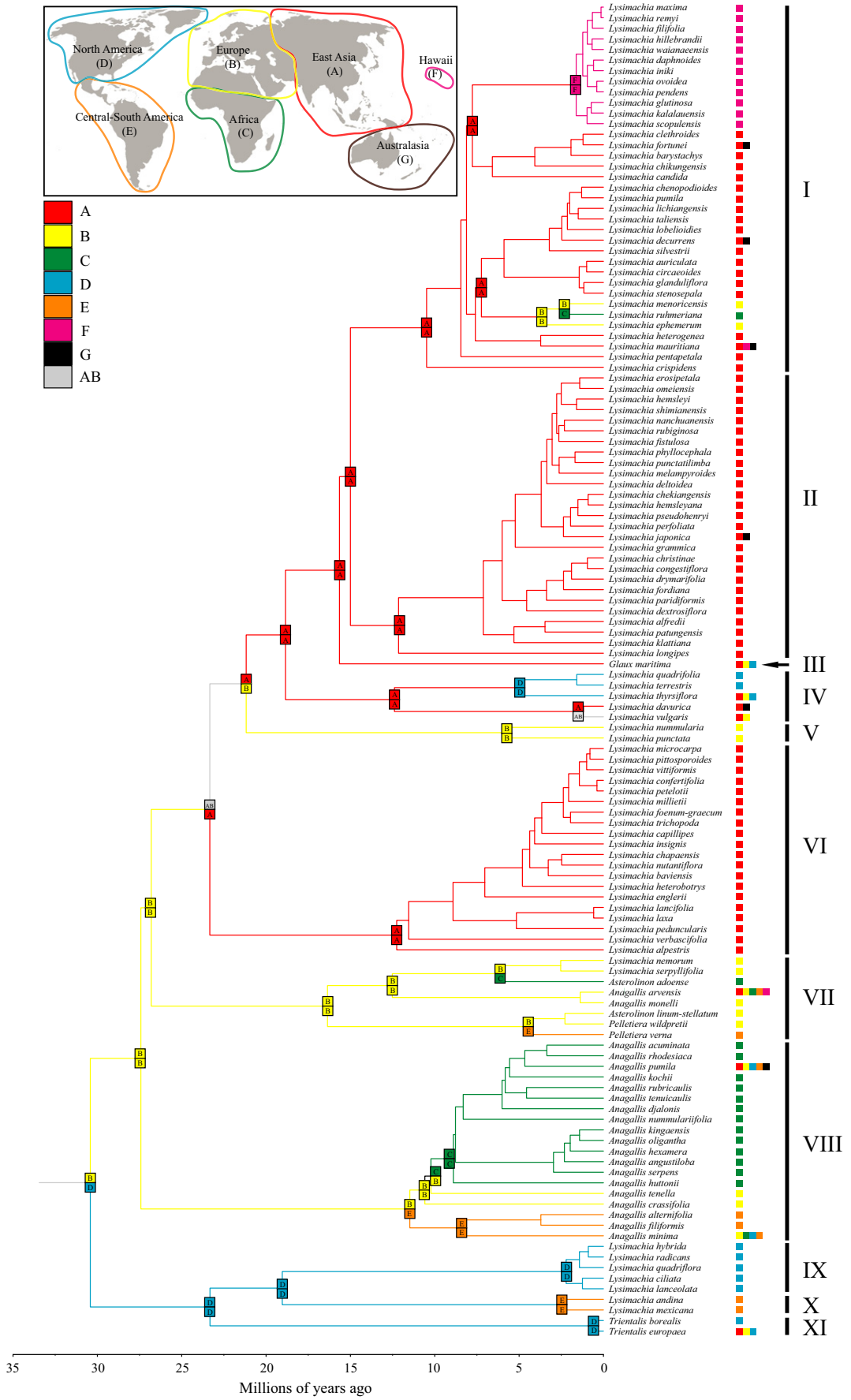

Fig. 1 Phylogeny and biogeographic reconstructions for Lysimachieae. The tree is maximum clade-credibility tree (from BEAST), estimated with the root date based on Strijk *et al.* (2014). The inset map indicates the geographic regions used in the biogeographic analyses. Colors indicate regions only. Ancestral areas were inferred by the dispersal–extinction–cladogenesis (DEC) model (in RASP 3.2). Colored branches and squares on nodes indicate ancestral areas. Colored squares at tips indicate current distributions of each species (Supporting Information Table S2). Main clades were represented with Latin numbers that are also shown in Fig. S1.

sampling is incomplete (Meyer & Wiens, 2018). This method uses clade ages, species richness, and a correction for failing to sample extinct clades (relative extinction fraction ε). We initially used an intermediate ε of 0.45, but found that very low and high values ($\varepsilon = 0$, $\varepsilon = 0.90$) gave similar results. The diversification rate for Australasia was treated as zero since there was no *in situ* diversification there (no endemic species or sister species). Importantly, the approach used here to estimate diversification rates incorporates all species in each clade (included in the phylogeny or not), and only requires sampling one species per clade.

To estimate a single diversification rate for regions with multiple clades, we weighted the diversification rate for each clade by multiplying it by its proportional richness in the region (the clade's richness in that region divided by the region's total richness). Thus, clades with more species in a region had more weight than clades with fewer species in that region. These weighted rates were then summed across clades to yield the weighted diversification rate for the region. Finally, we tested the relationship between the current richness of each region and its weighted diversification rate using linear regression in R. Again, we did not correct for phylogeny in these region-based analyses.

Some studies have stated that the rate estimator used here requires constant rates within clades and a positive relationship between richness and clade ages (e.g. Rabosky *et al.*, 2012). However, these studies did not address this estimator's accuracy. New simulations show that it can be reasonably accurate regardless of the relationship between clade age and richness (Kozak & Wiens, 2016) and regardless of whether rates are homogeneous or heterogeneous within the clade (Meyer & Wiens, 2018). Furthermore, the net diversification rate depends only on the age and richness of clades. Thus, a young clade with many species will have a high net diversification rate (and an older clade with few species a low rate), regardless of variation in instantaneous rates within the clade (either over time or among subclades). Simulations that include diversity dependence (i.e. rates changing over time) show that diversification rates still strongly predict richness patterns (Pontarp & Wiens, 2017). Nevertheless, the method-of-moments estimator does not estimate separate speciation and extinction rates.

To estimate speciation and extinction rates in each region, we used the Multiple State Speciation Extinction (MuSSE) model implemented in the R package DIVERSITREE (FitzJohn, 2012). We did not use GEOSSE (Goldberg *et al.*, 2011) here because it only allows analysis of two regions, and we analyzed six here. We used one state to represent each region. Australasia was excluded because it lacks endemic species. Species present in more than one region were coded based on their inferred ancestral region. Three parameters were included for each region (state): a speciation rate λ, an extinction rate μ, and a transition rate between different regions q. We compared the relative fit of the data with eight likelihood models, each with different combinations of parameters that were either free to vary among regions or constrained to be equal among regions: (1) λ free, μ free, and q free (i.e. separate values for each parameter estimated for each region); (2) λ equal between all regions, μ free, and q free; (3) λ equal, μ equal, and q free; (4) λ equal, μ equal, and q equal; (5) λ free, μ equal, and q equal; (6) λ equal, μ free, and q equal; (7) λ free, μ free, and q equal; and (8) λ free, μ equal, and q free. We compared their relative fit using the size-corrected AIC (AICc) (Burnham & Anderson, 2002). The Chi-squared distribution (ChiSq) and their significance (Pr) were calculated by comparison with the minimal model. We accounted for incomplete sampling in our tree for each region using a sampling fraction for each region (species in our tree in each region divided by the total number known in each region; Table 2). We estimated the posterior density distribution of parameters for the best-fitting model with Bayesian Markov chain Monte Carlo analyses (100 000 steps) to estimate speciation, extinction, and dispersal rates.

We also used these estimated speciation and extinction rates from MuSSE as an alternative approach to estimating the overall

Table 2 Summary of species richness and diversification rate (based on MuSSE analyses) of Lysimachieae in each region

Regions	Species richness (sampled species)	Sampling fraction (%)	Diversification rate		First colonization time × Diversification rate	
			MuSSE model 4	MuSSE model 6	MuSSE model 4	MuSSE model 6
East Asia (A)	171 (74)	43.27	0.2866	0.2704	6.3793	6.0191
Europe (B)	22 (18)	81.82	0.0851	0.0656	2.4620	1.8960
Africa (C)	28 (17)	60.71	0.1987	0.1958	1.9000	1.8722
North America (D)	19 (13)	68.42	0.1508	0.0832	4.0454	2.2309
Central–South America (E)	10 (8)	80.00	0.0678	0.0578	0.6749	0.5751
Hawaii (F)	18 (14)	77.78	0.8073	0.6691	1.5022	1.2450
Australasia (G)	6 (6)	100.00	0	0	0	0

diversification rate for each region. Importantly, this did not require assigning species to clades, nor calculating a weighted index among clades within a region. Australasia was assigned zero (see earlier).

Although some concerns have been raised about SSE methods (e.g. Maddison & FitzJohn, 2015; Rabosky & Goldberg, 2015), we show that our diversification rate estimates from MuSSE are similar to those from an alternative method (method-of-moments estimator). Therefore, it seems unlikely that issues of overall model fit are relevant here.

Combining diversification and time

We also developed a simple index that combines the effects of time and diversification rate. For this index, we simply multiplied the time (first colonization age or summed ages) by the net diversification rate (weighted net diversification rate or regional diversification rate from MuSSE) for each region.

Finally, in order to test the effects of different tree priors, calibration points, and epsilon values on our hypotheses (i.e. time, diversification-rate and combined-effect hypotheses), we performed these analyses on all alternative trees with full sampling and with different epsilon values. For brevity, results in the main text used only the primary tree and the intermediate epsilon.

Results

Phylogenetic relationships

The combined matrix consisted of 13 462 aligned base pairs for 131 species (including five outgroups). The likelihood, Bayesian, and BEAST topologies were similar and are summarized in Figs 1 and S1–S3. We identified 11 major clades in Lysimachieae with strong support (Fig. 1), which were concordant with major clades found in other recent studies (Hao *et al.*, 2004; Anderberg *et al.*, 2007; Manns & Anderberg, 2011). Our results strongly resolved the relationships of these major clades except for two nodes (Fig. 1). The first weakly supported relationship (bootstrap/posterior probability values: 65/0.86; Fig. S1) is among the monotypic genus '*Glaux*' (Clade III) and two other *Lysimachia* clades (Clades I and II). Previous studies found similarly weak support for these relationships (Hao *et al.*, 2004; Anderberg *et al.*, 2007). The second weak relationship (bootstrap/posterior probability values: 74/0.69; Fig. S1) is among Clade VII, Clade VIII, and Clades I–VI. Clade VII is sister to other '*Anagallis*' taxa (Clade VI) in Manns & Anderberg (2005). However, we placed Clade VII with Clades I–VI ('*Glaux*' and all *Lyismachia* except the endemic *Lysimachia* group in the New World), but with only moderate support.

Divergence-time estimation

Similar time estimates and topologies were obtained using three different calibration ages and different tree priors for the stem age of Lysimachieae (Figs 1, S2–S8; Tables S8, S14). We also tested the effect of incomplete sampling on divergence-time estimation using a tree with only 10% of all species included (26 species). Dating results based on these different strategies were all very similar to each other (Tables S8, S9). Our primary analyses (Fig. 1) used the full tree with the Yule prior, all calibration points, and the stem-age calibration with the intermediate age (from Strijk *et al.*, 2014). Stem ages of the 11 major clades estimated using different calibrations, reduced taxon sampling, and alternate tree priors were similar and tightly related to those from this baseline tree ($r^2 = 0.93$–0.99; Table S10).

Biogeographic reconstructions

Biogeographic analyses primarily used the tree with the intermediate root age and Yule speciation tree prior. We compared the fit of different constraints on the maximum number of ancestral regions for each node using likelihood ratio, χ^2, and AIC tests (Table S15). The best-fitting model in the DEC analysis (maximum two areas per ancestral species; Table S15) was used in the main analyses. Biogeographic inferences using a highly constrained transition matrix (Table S14) were basically identical to those without these assumptions (Table S12). However, the model utilizing these constraints had somewhat better fit ($\log_e L = -123.774$ vs $\log_e L = -126.619$; Tables S12, S14). Biogeographic inferences are illustrated in Fig. 1. These analyses suggest that the ancestor of Lysimachieae most likely occurred in both North America and Europe, and then dispersed to other regions (i.e. Africa, East Asia and Central–South America).

Time, diversification rate, and species richness

Results here are based on the tree using the intermediate root age and Yule speciation tree prior. Other trees yielded similar results (Table S16). For each region, the current species richness, oldest colonization time, and summed colonization times are listed in Table 1. There was no significant relationship between current species richness and the oldest colonization time among regions ($r^2 = 0.121$, $P = 0.446$; Fig. 2a), nor between richness and summed colonization ages ($r^2 = 0.022$, $P = 0.750$; Fig. 2b). Importantly, the ancestral regions (Europe and North America) for Lysimachieae have only 22 and 19 species respectively (Table 1). Other regions (Africa, Hawaii, South America, and Australia) have younger colonization times and low species richness. By contrast, East Asia has a relatively old colonization time and the highest species richness (Table 1). Therefore, Europe and North America act as outliers in these analyses, weakening any positive relationship between age and richness. The time–richness relationship is relatively strong when these two areas are removed ($r^2 = 0.809$, $P = 0.038$, using oldest colonization times), but not significant using summed colonization ages ($r^2 = 0.637$, $P = 0.106$; Fig. S9). Thus, time seems to contribute to the high species richness of East Asia, despite the lack of a time–richness relationship among regions overall.

The relationship between the weighted diversification rate of each region (stem-group estimator, $\varepsilon = 0.45$) and its regional species richness was also nonsignificant ($r^2 = 0.207$, $P = 0.306$, Fig. 2c). The lack of a significant relationship may be explained by two regions (Hawaii and Africa) having low richness but

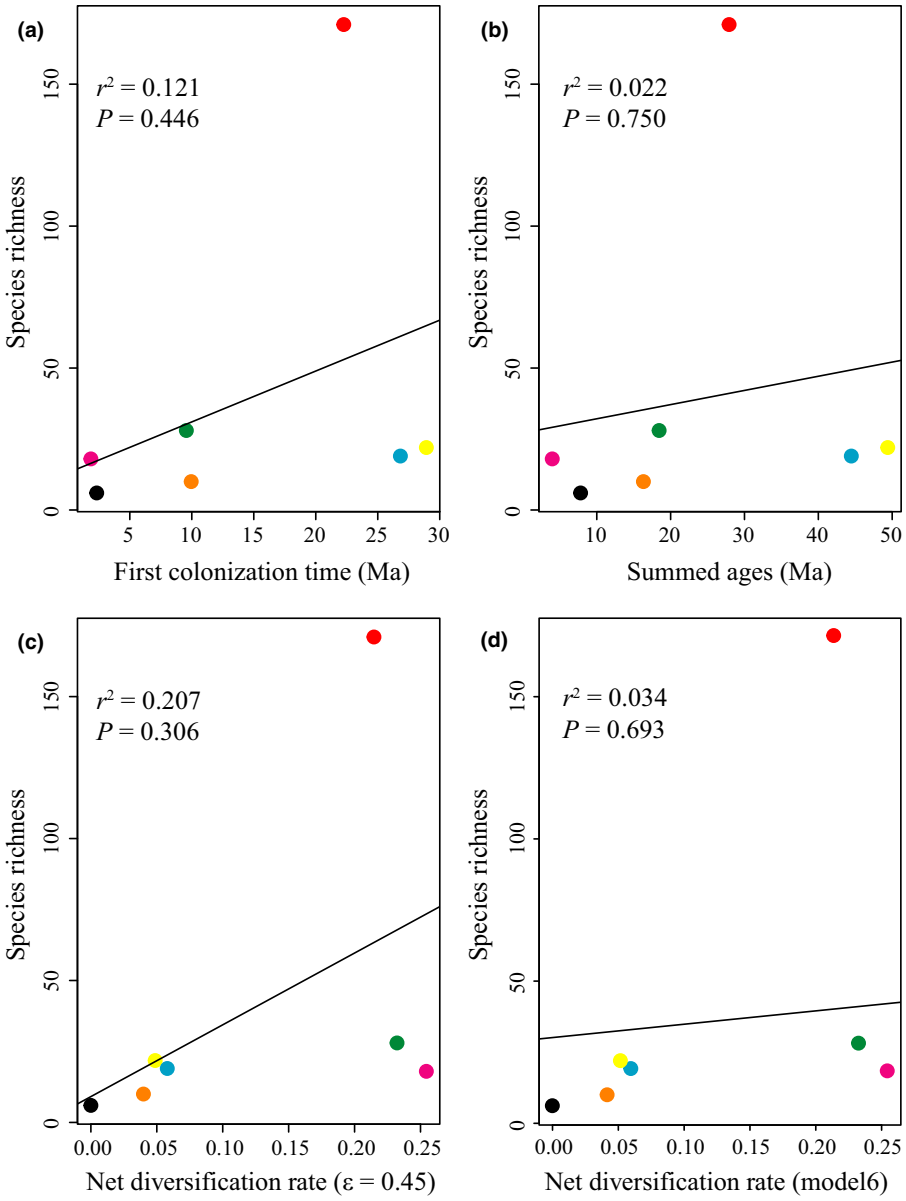

Fig. 2 Time and diversification separately fail to explain species richness of Lysimachieae across seven regions. Colored dots in each graph correspond to the colored regions as shown in Fig. 1. (a) Relationship between the time of first colonization and species richness for each region ($r^2 = 0.121$, $P = 0.446$). (b) Relationship between the summed age and species richness for each region ($r^2 = 0.022$, $P = 0.750$). (c) Relationship between the weighted diversification rate (based on $\varepsilon = 0.45$) and species richness for each region ($r^2 = 0.207$, $P = 0.306$). (d) Relationship between the diversification rate (based on the result of MuSSE: model 6) and species richness for each region ($r^2 = 0.034$, $P = 0.693$). Ma, millions of years ago.

relatively high diversification rates. For example, Hawaii has only 18 Lysimachieae species (mainly *Lysimachia* subgenus *Lysimachiopsis*) but a high diversification rate (0.2544 species/Myr). The rate and richness are similar in the African clade (0.2322, Table 1). A strong relationship between regional richness and diversification rate was supported after removing Hawaii and Africa ($r^2 = 0.960$, $P = 0.003$; Fig. S9). East Asia has a relatively high diversification rate (0.2147), which seemingly helps explain its high species richness. Relationships were similar using diversification rates estimated with ε of 0 and 0.90 (Fig. S10).

We also tested the hypothesis that time and diversification rates act together to explain richness patterns among regions. To test this combined effect, we simply multiplied the weighted diversification rate ($\varepsilon = 0.45$) for each region by its oldest colonization time. We found a strong relationship between regional species richness and this combined effect ($r^2 = 0.855$, $P = 0.003$, Fig. 3a). This relationship remains strong when East Asia is removed ($r^2 = 0.809$, $P = 0.006$).

The best-fitting MuSSE model supported different speciation rates but equal extinction and transition rates among regions (AICc = 788.788, $\log_e L = -386.394$, Table S17). Under this model, speciation rates in each region are as follows: East Asia, 0.3520; Europe, 0.1505; Africa, 0.2641; North America, 0.2162; Central–South America, 0.1332; and Hawaii, 0.8727 (Table S18). Thus, the highest speciation rates are in Hawaii and East Asia. However, the second-best model (model 6, AICc = 793.577, $\log_e L = -383.788$) has a similar fit to the best model (dAICc = 4.789), and allows both speciation and extinction rates to vary (Tables S17, S19). This difference is close to the standard cutoff for model selection (Burnham & Anderson, 2002). Speciation rates under the two models (best-fitting and second-best) are very similar ($r^2 = 0.979$, $P = 0.0002$), but extinction rates vary under the second-best model. Using this latter model

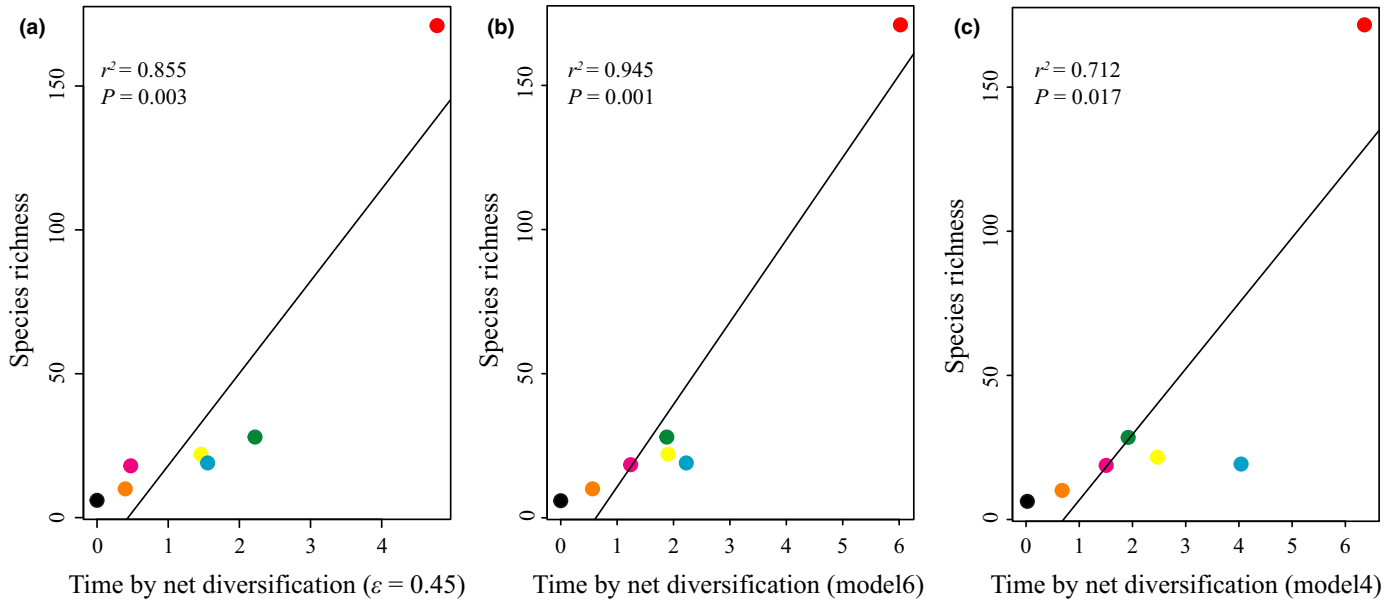

Fig. 3 Time and diversification rates together explain species richness of Lysimachieae across seven regions. Colored dots in each graph show the colored regions in Fig. 1. (a) Strong relationship between the combined effect of time (first colonization of each region) and weighted net diversification rate (based on $\varepsilon = 0.45$) and species richness of regions ($r^2 = 0.855$, $P = 0.003$). (b) Strong relationship between the combined effect of time (first colonization of each region) and diversification rate (based on MuSSE: model 6) and species richness of regions ($r^2 = 0.945$, $P = 0.001$). (c) Strong relationship between the combined effect of time (first colonization of each region) and diversification rate (based on MuSSE: model 4) and species richness of regions ($r^2 = 0.712$, $P = 0.017$).

(Table S19), Europe (0.1515) and North America (0.2665) have higher extinction rates than East Asia (0.1228) and Africa (0.0794). MuSSE analyses with alternative trees showed model 4 (best-fitting model) and model 6 (second-best) had a similar fit (Table S20). Net diversification rates ($\lambda - \mu$) for each region calculated by MuSSE (using the model with variable speciation and extinction rates among regions) were broadly similar to the weighted net diversification rates estimated for these regions (using $\varepsilon = 0.45$; $r^2 = 0.673$, $P = 0.024$; Tables 1, 2). Hawaii had an extremely high net diversification rate using MuSSE (0.6691) relative to the weighted net diversification rate (0.2544; Tables 1, 2). When Hawaii is removed, the relationship is very strong ($r^2 = 0.863$, $P = 0.022$). There was no significant relationship between the net diversification rate for each region from MuSSE and each region's species richness ($r^2 = 0.034$, $P = 0.693$, Fig. 2d). Results were similar under the model in which only speciation rates differed ($r^2 = 0.001$, $P = 0.944$). However, there was a significant relationship between the species richness of each region and the combined index of time and MuSSE-estimated diversification rates (under the model with different speciation and extinction rates; $r^2 = 0.945$, $P = 0.001$, Fig. 3b). Results were significant but weaker with East Asia removed ($r^2 = 0.767$, $P = 0.022$). Results were also similar but somewhat weaker ($r^2 = 0.712$, $P = 0.017$) using the model with only speciation rates that differ among regions (Fig. 3c).

Discussion

What explains the high species richness of plants in East Asia relative to other temperate regions? This is a long-standing question, but no studies have tested both time and diversification rates as potential explanations. Our results for Lysimachieae suggest that this pattern is explained by the combination of time and diversification rates, rather than one or the other. Specifically, East Asia has been colonized for a relatively long period of time by Lysimachieae, and also has relatively high net diversification rates compared with other regions. However, in our analyses, neither factor alone explains high richness in East Asia. In the paragraphs that follow, we discuss the importance of this combined explanation. We then discuss future work, which should address two major questions: (1) What ecological factors explain the evolutionary and biogeographic patterns found? (2) Do results for Lysimachieae apply to other plant groups in East Asia?

Our results suggest that high species richness in East Asia in Lysimachieae is explained by a combination of time and relatively high diversification rates. Neither variable alone explained richness patterns among regions. Specifically, even though East Asia has been colonized for a relatively long period of time relative to most other regions, Europe and North America have been colonized longer and yet are relatively species poor. Therefore, based on time alone, Europe and North America should have the highest richness. Similarly, even though East Asia has a relatively high weighted net diversification rate (much higher than Europe or North America), both Hawaii and Africa have higher rates. However, Hawaii and Africa were both colonized relatively recently, and therefore their high net diversification rates have not yet yielded high species richness. Overall, neither time nor diversification rates explained richness patterns among regions. Yet, a weighted measure that incorporates both time and diversification rates provides a strong explanation for richness patterns among

regions, especially the high richness of East Asia relative to other regions. To our knowledge, this is the first study to use such a combined measure. Interestingly, many previous studies on richness patterns tend to favor time over diversification rates (e.g. Hutter *et al.*, 2013; Wiens *et al.*, 2013), whereas others favor diversification rates over time (e.g. Pyron & Wiens, 2013; Wiens, 2015). Few have found that both explain richness patterns (e.g. Smith *et al.*, 2007), although such a pattern is not unexpected based on theory (e.g. Pontarp & Wiens, 2017). Our results here strongly reinforce the need to test the influence of both time and diversification rates on spatial richness patterns. Moreover, our results suggest that the combined effects of these variables should also be considered in cases (like ours) in which neither variable explains richness patterns on its own.

It is important to note that these combined effects of time and diversification rates are not an inevitable explanation for richness patterns. Specifically, if we were attempting to explain richness patterns among clades only (e.g. genera), then it is true that multiplying the age of the clade by its net diversification rate should yield that clade's observed number of species. However, this is not so straightforward for spatial richness patterns among regions. For example, the weighted net diversification rate for the region is based on the age of each named clade (not how long they have been present in each region), weighted by the number of clades in each region, the number of species from each clade in the region, and the global diversification rate of each clade. Conversely, the age of a clade in a region (i.e. oldest colonization time) is not necessarily the same as the age used to estimate the diversification rate. Therefore, we strongly emphasize that the strong relationship observed here between species richness and our combined measure of colonization time and diversification rate is not inevitable.

Our results help identify the combination of evolutionary and biogeographic processes (i.e. speciation, extinction, colonization time) that explain high species richness in East Asia (Axelrod *et al.*, 1996; Xiang *et al.*, 2004). We suggest that a high priority for future research should be to identify the specific ecological factors that act on these processes (e.g. Guo *et al.*, 1998; Qian & Ricklefs, 2000; Qian, 2002; Wang *et al.*, 2009, 2011). For example, our results suggest that patterns of species richness in Lysimachieae are explained by relatively high speciation rates in East Asia, the highest of any region except Hawaii (Tables S10, S16). Therefore, future work should seek to explain what causes high speciation rates in these regions. We suggest that the radiation of Lysimachieae across diverse elevations and associated climatic regimes in both regions may be important (Chen & Hu, 1979; Oh *et al.*, 2013). The high diversification rate of Hawaii may also result (at least in part) from ecological opportunity associated with the absence of some potentially competing mainland plant clades, as suggested by the ecological theory of adaptive radiation (Schluter, 2000). However, this hypothesis will require additional analyses to elucidate (e.g. has there been greater ecological diversification in Hawaii? are there fewer relevant clades that might compete with them?). Our second-best MuSSE model (Table S17) also suggests that Lysimacheae in North America experienced relatively high extinction rates, which substantially lowered their net diversification rates. By contrast, Europe and East Asia share relatively modest extinction rates. One potential explanation for the high extinction rates in North America is the combined effects of glaciation in northern regions and aridification in western regions (Latham & Ricklefs, 1993; Qian & Ricklefs, 2000; Eronen *et al.*, 2012; Ricklefs & He, 2016). Surprisingly, Hawaii appears to have the highest extinction rates of any region, but its very high speciation rate seems to compensate for the dampening effects of this high extinction on net diversification rates.

Our results also show that area and latitude may not be the most important factors driving richness patterns. In Lysimacheae, species richness in East Asia is higher than the combined richness of Europe, North America, Central–South America, and Australasia (Table 1), despite the much greater area of these combined regions. Furthermore, these species-poor regions include low-latitude tropical regions (e.g. Africa, Central–South America). We suggest that area and latitude may only impact richness if there is sufficient time to build up richness in each region.

A second major question from our study is whether our results for Lysimachieae are representative of all plant groups in East Asia. Our underlying goal is to understand why East Asia has such high plant richness overall (e.g. Qian & Ricklefs, 2000). We addressed this question using Lysimachieae as a model system. However, we recognize that this tribe may not be fully representative of all groups of plants that contribute to high East Asian species richness. That being said, we do not know of any potentially relevant way in which Lysimachieae is different from other species-rich plant groups in the region. For example, the tribe has been considered a typical group with high species richness in East Asia (Chen & Hu, 1979; Li, 1996; Ståhl & Anderberg, 2004). In addition, the tribe has many characteristics that are widespread in other large Chinese plant groups (e.g. *Primula*, *Gentiana*, *Saxifraga*, and *Pedicularis*; Li, 1996). These characteristics include herbaceous growth form, insect pollination, no long-distance dispersal ability, and mesic habitat (Ståhl & Anderberg, 2004). Nevertheless, we think that the best way to address whether our results are representative or not is to do similar analyses in other groups with high species richness in East Asia, and see whether they have patterns similar to those found here for Lysimachieae.

In summary, our study helps reveal the underlying causes of high plant species richness in East Asia. In the tribe Lysimachieae, this pattern is caused by both relatively long occurrence in the region (time-for-speciation effect) and relatively high net diversification rates in the clades that occur there. Importantly, neither variable alone fully explains this pattern. More broadly, our results demonstrate the importance of testing the combined effects of time and diversification, and the need to test both hypotheses (and possibly their interaction) to explain large-scale patterns of species richness.

Acknowledgements

We thank the Plant DNA Bank in Korea, Herbario Nacional de México, and Ki-Joong Kim, Qing Liu, and Lian-Ming Gao for samples. We thank Chi-Ming Hu for species identification. We thank Cristian Román-Palacios, Shea Lambert, Elizabeth Miller, and Tania Hernández-Hernández for help with data analyses.

This research was supported by the National Natural Science Foundation of China (31570222), the National Key Basic Research Program of China (2014CB954100), and the Strategic Priority Research Program of the Chinese Academy of Sciences (XDPB020303).

Author contributions

H-F.Y., G.H., X-J.G., and J.J.W. designed the study; C-Y.Z. performed the experiments; A.A.A. provided all other data; H-F.Y. analyzed data; and H-F.Y. and J.J.W. wrote the paper, with significant contributions from G.H. and X-J.G.

ORCID

Hai-Fei Yan http://orcid.org/0000-0002-5692-9391

References

Adams J. 2009. *Species richness: patterns in the diversity of life*. Chichester, UK: Praxis Publishing.

Anderberg AA, Manns U, Källersjö M. 2007. Phylogeny and floral evolution of the Lysimachieae (Ericales, Myrsinaceae): evidence from *ndh*F sequence data. *Willdenowia* **37**: 407–421.

Axelrod DI, Al-Shehbaz I, Raven PH. 1996. History of the modern flora of China. In: Zhang AL, Wu SG, eds. *Floristic characteristics and diversity of East Asian plants*. Beijing, China: China Higher Education Press, 43–55.

Banfi E, Galasso G, Soldano A. 2005. Notes on systematics and taxonomy for the Italian vascular flora 1. *Atti della Società italiana di scienze naturali e del Museo civico di storia naturale di Milano* **146**: 219–244.

Bennell AP, Hu CM. 1983. The pollen morphology and taxonomy of *Lysimachia*. *Notes from the Royal Botanic Garden Edinburgh* **40**: 425–458.

Bouckaert R, Heled J, Kühnert D, Vaughan T, Wu CH, Xie D, Suchard MA, Rambaut A, Drummond AJ. 2014. BEAST 2: a software platform for Bayesian evolutionary analysis. *PLoS Computational Biology* **10**: e1003537.

Burnham KP, Anderson DR. 2002. *Model selection and multimodel inference: a practical information-theoretic approach*, 2nd edn. New York, NY, USA: Springer-Verlag.

Chen FH, Hu CM. 1979. Taxonomic and phytogeographic studies on Chinese species of *Lysimachia*. *Acta Phytotaxonomica Sinica* **17**: 21–56.

Cholewa AF, Pipoly JJ, Ricketson JM. 2009. Myrsinaceae. In: Flora of North America Editorial Committee, ed. *Flora of North America north of Mexico, volume 8: Magnoliophyta: Paeoniaceae to Ericaceae*. New York, NY, USA: Oxford University Press, 302–318.

Doyle JJ, Doyle JL. 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochemical Bulletin* **19**: 11–15.

Drummond AJ, Bouckaert RR. 2015. *Bayesian evolutionary analysis with BEAST*. Cambridge, UK: Cambridge University Press.

Drummond AJ, Ho SY, Phillips MJ, Rambaut A. 2006. Relaxed phylogenetics and dating with confidence. *PLoS Biology* **4**: e88.

Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research* **32**: 1792–1797.

Eronen JT, Fortelius M, Micheels A, Portmann FT, Puolamäki K, Janis CM. 2012. Neogene aridification of the Northern Hemisphere. *Geology* **40**: 823–826.

Fischer AG. 1960. Latitudinal variations in organic diversity. *Evolution* **14**: 64–81.

FitzJohn RG. 2012. Diversitree: comparative phylogenetic analyses of diversification in R. *Methods in Ecology and Evolution* **3**: 1084–1092.

Friis EM. 1985. *Angiosperms fruits and seeds from the Middle Miocene of Jutland (Denmark). Det Kongelige Danske Videnskabernes Selskabs Biologiske Skrifter*24:3. (The Society of the Royal Danish Scientific Society Biological Reports). Copenhagen: Munksgaard.

Goldberg EE, Lancaster LT, Ree RH. 2011. Phylogenetic inference of reciprocal effects between geographic range evolution and diversification. *Systematic Biology* **60**: 451–465.

Guo Q, Ricklefs RE, Cody ML. 1998. Vascular plant diversity in eastern Asia and North America: historical and ecological explanations. *Botanical Journal of the Linnean Society* **128**: 123–136.

Hall TA. 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series* **41**: 95–98.

Hao G, Yuan YM, Hu CM, Ge XJ, Zhao NX. 2004. Molecular phylogeny of *Lysimachia* (Myrsinaceae) based on chloroplast *trn*L-F and nuclear ribosomal ITS sequences. *Molecular Phylogenetics and Evolution* **31**: 323–339.

Harmon LJ, Weir JT, Brock CD, Glor RE, Challenger W. 2008. GEIGER: investigating evolutionary radiations. *Bioinformatics* **24**: 129–131.

Ho SY, Phillips MJ. 2009. Accounting for calibration uncertainty in phylogenetic estimation of evolutionary divergence times. *Systematic Biology* **58**: 367–380.

Hong DY, Blackmore S. 2013. *Plants of China: a companion of the Flora of China*. Cambridge, UK: Cambridge University Press.

Hu CM. 1994. On the geographical distribution of the Primulaceae. *Journal of Tropical and Subtropical Botany* **2**: 1–14.

Hu CM, Kelso S. 1996. Primulaceae. In: Wu ZY, Raven PH, eds. *Flora of China 15*. St Louis, USA/Beijing, China: Missouri Botanical Garden Press/Science Press, 39–189.

Hutter CR, Guayasamin JM, Wiens JJ. 2013. Explaining Andean megadiversity: the evolutionary and ecological causes of glassfrog elevational richness patterns. *Ecology Letters* **16**: 1135–1144.

Källersjö M, Bergqvist G, Anderberg AA. 2000. Generic realignment in Primuloid families of the Ericales s.l.: a phylogenetic analysis based on DNA sequences from three chloroplast genes and morphology. *American Journal of Botany* **87**: 1325–1341.

Kozak KH, Wiens JJ. 2016. Testing the relationships between diversification, species richness, and trait evolution. *Systematic Biology* **65**: 975–988.

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

Latham RE, Ricklefs RE. 1993. Continental comparisons of temperate-zone tree species diversity. In: Ricklefs RE, Schluter D, eds. *Species diversity in ecological communities: historical and geographical perspectives*. Chicago, IL, USA: University of Chicago Press, 294–314.

Li XW. 1996. Floristic statistics and analyses of seed plants from China. *Acta Botanica Yunnanica* **18**: 363–384.

Löve A, Löve D. 1982. IOPB chromosome number reports LXXVII. *Taxon* **31**: 766–768.

Maddison WP, FitzJohn RG. 2015. The unsolved challenge to phylogenetic correlation tests for categorical characters. *Systematic Biology* **64**: 127–136.

Magallón S, Sanderson MJ. 2001. Absolute diversification rates in angiosperm clades. *Evolution* **55**: 1762–1780.

Manns U, Anderberg AA. 2005. Molecular phylogeny of *Anagallis* (Myrsinaceae) based on ITS, *trn*L-F, and *ndh*F sequence data. *International Journal of Plant Sciences* **166**: 1019–1028.

Manns U, Anderberg AA. 2009. New combinations and names in *Lysimachia* (Myrsinaceae) for species of *Anagallis*, *Pelletiera* and *Trientalis*. *Willdenowia* **39**: 49–54.

Manns U, Anderberg AA. 2011. Biogeography of 'tropical *Anagallis*' (Myrsinaceae) inferred from nuclear and plastid DNA sequence data. *Journal of Biogeography* **38**: 950–961.

Meyer ALS, Wiens JJ. 2018. Estimating diversification rates for higher taxa: BAMM can give problematic estimates of rates and rate shifts. *Evolution* **72**: 39–53.

Miller MA, Pfeiffer W, Schwartz T. 2010. Creating the CIPRES Science Gateway for inference of large phylogenetic trees. In: *Proceedings of the Gateway Computing Environments Workshop (GCE)*. New Orleans, LA, USA, 1–8.

Mittelbach GG, Schemske DW, Cornell HV, Allen AP, Brown JM, Bush MB, Harrison SP, Hurlbert AH, Knowlton N, Lessios HA *et al.* 2007. Evolution

and the latitudinal diversity gradient: speciation, extinction and biogeography. *Ecology Letters* **10**: 315–331.

Myers N, Mittermeier RA, Mittermeier CG, Da Fonseca GA, Kent J. 2000. Biodiversity hotspots for conservation priorities. *Nature* **403**: 853–858.

Oh IC, Anderberg AL, Schönenberger J, Anderberg AA. 2008. Comparative seed morphology and character evolution in the genus *Lysimachia* (Myrsinaceae) and related taxa. *Plant Systematics and Evolution* **271**: 177–197.

Oh IC, Schönenberger J, Motley TJ, Myrenås M, Anderberg AA. 2013. Phylogenetic relationships among endemic Hawaiian *Lysimachia* (Ericales: Primulaceae): insights from nuclear and chloroplast DNA sequence data. *Pacific Science* **67**: 237–251.

Pax F, Knuth R. 1905. *Primulaceae*. Leipzig, Germany: Verlag von Wilhelm Engelmann.

Pontarp M, Wiens JJ. 2017. The origin of species richness patterns along environmental gradients: uniting explanations based on time, diversification rate, and carrying capacity. *Journal of Biogeography* **44**: 722–735.

Probatova NS, Rudyka EG, Barkalov VY, Nesterova IA, Kudrin SG, Chubar EA. 2006. Chromosome numbers of vascular plants from nature reserves of the Primorsky Territory and the Amur River basin. *Botanicheskii Zhurnal (Moscow & Leningrad)* **91**: 1117–1134.

Pyron RA, Wiens JJ. 2013. Large-scale phylogenetic analyses reveal the causes of high tropical amphibian diversity. *Proceedings of the Royal Society of London B: Biological Sciences* **280**: 20131622.

Qian H. 2001. A comparison of generic endemism of vascular plants between East Asia and North America. *International Journal of Plant Sciences* **162**: 191–199.

Qian H. 2002. A comparison of the taxonomic richness of temperate plants in East Asia and North America. *American Journal of Botany* **89**: 1818–1825.

Qian H, Ricklefs RE. 2000. Large-scale processes and the Asian bias in species diversity of temperate plants. *Nature* **407**: 180–182.

R Core Team. 2015. *R: a language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. [WWW document] URL https://www.R-project.org/. [accessed 1 March 2016]

Rabosky DL, Goldberg EE. 2015. Model inadequacy and mistaken inference of trait-dependent speciation. *Systematic Biology* **64**: 340–355.

Rabosky DL, Slater GJ, Alfaro ME. 2012. Clade age and species richness are decoupled across the eukaryotic tree of life. *PLoS Biology* **10**: e1001381.

Rambaut A. 2009. *FigTree v1.4.2*. [WWW document] URL http://tree.bio.ed.ac.uk/software/figtree/ [accessed 1 March 2016]

Rambaut A, Suchard MA, Xie D, Drummond AJ. 2014. *Tracer v1.6*. [WWW document] URL http://beast.bio.ed.ac.uk/Tracer [accessed 1 March 2016]

Ray JD. 1956. *The genus Lysimachia in the New World*. Champaign, IL, USA: The University of Illinois Press.

Ree RH, Smith SA. 2008. Maximum likelihood inference of geographic range evolution by dispersal, local extinction, and cladogenesis. *Systematic Biology* **57**: 4–14.

Renner SS, Schaefer H. 2010. The evolution and loss of oil-offering flowers: new insights from dated phylogenies for angiosperms and bees. *Philosophical Transactions of the Royal Society B: Biological Sciences* **365**: 423–435.

Ricklefs RE. 1987. Community diversity: relative roles of local and regional processes. *Science* **235**: 167–171.

Ricklefs RE, He F. 2016. Region effects influence local tree species diversity. *Proceedings of the National Academy of Sciences, USA* **113**: 674–679.

Schluter D. 2000. *The ecology of adaptive radiation*. Oxford, UK: Oxford University Press.

Smith SA, Nieto Montes, de Oca A, Reeder TW, Wiens JJ. 2007. A phylogenetic perspective on elevational species richness patterns in Middle American treefrogs: why so few species in lowland tropical rainforests? *Evolution* **61**: 1188–1207.

Son O, Park SJ. 2016. Complete chloroplast genome sequence of *Lysimachia coreana* (Primulaceae). *Mitochondrial DNA Part A* **27**: 2263–2265.

Stadler T. 2009. On incomplete sampling under birth–death models and connections to the sampling-based coalescent. *Journal of Theoretical Biology* **261**: 58–66.

Stadler T, Kühnert D, Bonhoeffer S, Drummond AJ. 2013. Birth–death skyline plot reveals temporal changes of epidemic spread in HIV and hepatitis C virus (HCV). *Proceedings of the National Academy of Sciences, USA* **110**: 228–233.

Ståhl B, Anderberg AA. 2004. Myrsinaceae. In: Kubitzki K, ed. *The families and genera of vascular plants 6*. Heidelberg, Germany: Springer-Verlag, 266–281.

Stephens PR, Wiens JJ. 2003. Explaining species richness from continents to communities: the time-for-speciation effect in emydid turtles. *American Naturalist* **161**: 112–128.

Strijk JS, Bone RE, Thébaud C, Buerki S, Fritsch PW, Hodkinson TR, Strasberg D. 2014. Timing and tempo of evolutionary diversification in a biodiversity hotspot: Primulaceae on Indian Ocean islands. *Journal of Biogeography* **41**: 810–822.

Wang Z, Brown JH, Tang Z, Fang J. 2009. Temperature dependence, spatial scale, and tree species diversity in eastern Asia and North America. *Proceedings of the National Academy of Sciences, USA* **106**: 13388–13392.

Wang Z, Fang J, Tang Z, Lin X. 2011. Patterns, determinants and models of woody plant diversity in China. *Proceedings of the Royal Society of London B: Biological Sciences* **278**: 2122–2132.

Wiens JJ. 2011. The causes of species richness patterns across space, time, and clades and the role of "ecological limits". *The Quarterly Review of Biology* **86**: 75–96.

Wiens JJ. 2015. Faster diversification on land than sea helps explain global biodiversity patterns among habitats and animal phyla. *Ecology Letters* **18**: 1234–1241.

Wiens JJ, Kozak KH, Silva N. 2013. Diversity and niche evolution along aridity gradients in North American lizards (Phrynosomatidae). *Evolution* **67**: 1715–1728.

Wiens JJ, Morrill MC. 2011. Missing data in phylogenetic analysis: reconciling results from simulations and empirical data. *Systematic Biology* **60**: 719–731.

Willig MR, Kaufman DM, Stevens RD. 2003. Latitudinal gradients of biodiversity: pattern, process, scale, and synthesis. *Annual Review of Ecology, Evolution, and Systematics* **34**: 273–309.

Willis JC. 1922. *Age and area*. Cambridge, UK: Cambridge University Press.

Xiang QY, Zhang WH, Ricklefs RE, Qian H, Chen ZD, Wen J, Li JH. 2004. Regional differences in rates of plant speciation and molecular evolution: a comparison between eastern Asia and eastern North America. *Evolution* **58**: 2175–2184.

Yesson C, Toomey NH, Culham A. 2009. *Cyclamen*: time, sea and speciation biogeography using a temporally calibrated phylogeny. *Journal of Biogeography* **36**: 1234–1252.

Yu Y, Harris AJ, Blair C, He X. 2015. RASP (Reconstruct Ancestral State in Phylogenies): a tool for historical biogeography. *Molecular Phylogenetics and Evolution* **87**: 46–49.

Zanne AE, Tank DC, Cornwell WK, Eastman JM, Smith SA, FitzJohn RG, McGlinn 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.

Zhang CY, Wang FY, Yan HF, Hao G, Hu CM, Ge XJ. 2012. Testing DNA barcoding in closely related groups of *Lysimachia* L. (Myrsinaceae). *Molecular Ecology Resources* **12**: 98–108.

Zheng Y, Wiens JJ. 2015. Do missing data influence the accuracy of divergence-time estimation with BEAST? *Molecular Phylogenetics and Evolution* **85**: 41–49.

Supporting Information

Additional Supporting Information may be found online in the Supporting Information tab for this article:

Fig. S1 Maximum likelihood tree of Lysimachieae based on plastid and nuclear DNA sequence data.

Fig. S2 Maximum clade credibility tree of Lysimachieae from BEAST with its root calibrated following Renner & Schaefer (2010), and Yule speciation tree prior.

Fig. S3 Maximum clade credibility tree of Lysimachieae from BEAST with its root calibrated following Yesson *et al.* (2009) and Zanne *et al.* (2014), and a Yule speciation tree prior.

Fig. S4 Maximum clade credibility tree of Lysimachieae from BEAST with its root calibrated following Strijk *et al.* (2014), and a Birth-Death tree prior.

Fig. S5 Maximum clade credibility tree of Lysimachieae from BEAST with its root calibrated following Renner & Schaefer (2010), and a Birth-Death tree prior.

Fig. S6 Maximum clade credibility tree of Lysimachieae from BEAST with its root calibrated following Yesson *et al.* (2009) and Zanne *et al.* (2014), and a Birth-Death tree prior.

Fig. S7 Maximum clade credibility tree of Lysimachieae from BEAST under a Yule speciation tree prior and only including the fossil calibration.

Fig. S8 Maximum clade credibility tree of Lysimachieae from BEAST under Birth-Death tree prior and only fossil calibration.

Fig. S9 The relationships between the species richness of regions and their colonization times and diversification rates after removing outlier regions.

Fig. S10 The relationships between the current species richness of regions and weighted diversification rates using alternative relative extinction fractions (ε).

Table S1 List of all Lysimachieae species and their distributions

Table S2 Species of Lysimachieae sampled in this analysis, and their distributions

Table S3 Sampling information for 11 major clades, and their stem ages and diversification rates under each epsilon value

Table S4 Primers used in this study

Table S5 GenBank accession numbers of sequences used in this study

Table S6 The best-fitting set of partitions and models as estimated by PARTITIONFINDER

Table S7 The positions of each partition in the concatenated matrix

Table S8 Divergence times of major clades based on different tree priors (Yule, Birth-Death) and calibration points

Table S9 Divergence times of major clades based on two different tree priors (Yule, Birth-Death) and only 10% sampling of species in the tribe

Table S10 Regression results between the stem ages of the 11 major clades from the primary tree and the alternative trees

Table S11 A transition matrix (Q-matrix) to reflect changes in palaeogeographic events and potential dispersal between pairs of regions during different time periods

Table S12 The ancestral regions and divergence times inferred by the DEC algorithm for major clades without using assumptions about dispersal at different time periods

Table S13 Higher taxa, stem ages, species richness, diversification rates for higher taxa, and diversification rate of each region

Table S14 The ancestral regions and divergence times inferred by the DEC algorithm for major clades

Table S15 Comparison of model fit for different maximum numbers of regions per node in DEC analyses

Table S16 Summary of the relationships between species richness and time, net diversification rate, and the combined index of time and diversification rates, using different time-calibrated phylogenies

Table S17 Comparison of different MuSSE models

Table S18 Parameter estimates from Bayesian MCMC analyses under the best-fitting MuSSE model (Model4)

Table S19 Parameter estimates from Bayesian MCMC analyses under the second-best MuSSE model (Model6)

Table S20 Comparision of different MuSSE models for alternative trees

Methods S1 Additional phylogenetic analyses.

Methods S2 Detailed description of biogeographic regions.

Notes S1 The concatenated matrix and all MCC trees used in this study.

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.