bioRxiv preprint doi: https://doi.org/10.1101/619064; this version posted January 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Angiosperm speciation speeds up near the poles

J. IGEA[1*], A. J. TANENTZAP[1]

[1] *Department of Plant Sciences, University of Cambridge, Downing St, Cambridge, CB2 3EA, UK*

[*] Correspondence: ji247@cam.ac.uk

Authorship:

J.I. conceived the study and performed the analysis; J.I. and A.J.T interpreted the analysis and wrote the manuscript.

Short running title: (less than 45 characters excluding spaces) Angiosperm speciation cools down in the tropics

Keywords: latitudinal diversity gradient, speciation, angiosperms, macroevolution, biogeography, biodiversity

Type of article: Letter

Number of words in the abstract: 139

Number of words in the main text (excluding abstract, acknowledgements, references, table and figure legends): 4,744

Number of words in each text box: NA

Number of figures, tables, and text boxes: 3 Figures, 1 Table

bioRxiv preprint doi: https://doi.org/10.1101/619064; this version posted January 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Abstract

Recent evidence has questioned whether the Latitudinal Diversity Gradient (LDG), whereby species richness increases towards the Equator, results from higher rates of speciation in the tropics. Allowing for time heterogeneity in speciation rate estimates for over 60,000 angiosperm species, we found that the LDG does not arise from variation in speciation rates because lineages do not speciate faster in the tropics. These results were consistently retrieved using two other methods to test the association between occupancy of tropical habitats and speciation rates. Our speciation rate estimates were robust to the effects of both undescribed species and missing taxa. Overall, our results show that speciation rates follow an opposite pattern to global variation in species richness. Greater ecological opportunity in the temperate zones, stemming from less saturated communities, higher species turnover or greater environmental change, may ultimately explain these results.

bioRxiv preprint doi: https://doi.org/10.1101/619064; this version posted January 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Introduction

Biodiversity on Earth is very unevenly distributed. The Latitudinal Diversity Gradient (LDG), whereby species richness increases towards the Equator, is the most prominent example of this unevenness. This pervasive pattern has fascinated biologists for more than two centuries (von Humboldt 1808; Wallace 1878; Pianka 1966; Hawkins 2001), but its underlying causes remain largely unknown.

A major class of explanations proposes that elevated rates of speciation generate the higher diversity in the tropics and so primarily create the LDG (Mittelbach *et al.* 2007). This increase in speciation is generally attributed to higher environmental energy in the tropics, which in turn can hasten evolution through shorter generation times and higher mutation rates (Dowle *et al.* 2013), and is known as the "evolutionary speed hypothesis" (Rensch 1959; Rohde 1992). Other explanations propose that elevated speciation rates near the Equator can stem from higher chances of allopatric speciation caused by narrower thermal niches (Ghalambor *et al.* 2006), or from stronger biotic interactions (e.g., faster coevolution) – the 'Red Queen running faster when she is hot' (Brown 2014). However, while some studies on mammals (Rolland *et al.* 2014), amphibians (Pyron & Wiens 2013) and butterflies (Condamine *et al.* 2012) have shown that tropical lineages speciate faster, other studies have found no latitudinal differences (Rabosky & Huang 2015; Economo *et al.* 2018) or higher rates of speciation in temperate areas (Weir & Schluter 2007), particularly when focusing on recent rates of speciation (Schluter 2016; Schluter & Pennell 2017; Rabosky *et al.* 2018). Recent speciation rates should be less affected by extinction (Nee *et al.* 1994) and more accurate than deep-time estimates (Louca & Pennell 2019), even in the absence of paleontological evidence (Marshall 2017). They should

bioRxiv preprint doi: https://doi.org/10.1101/619064; this version posted January 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

also reflect any variable that historically influenced speciation and also varied latitudinally, such as environmental energy (Title & Rabosky 2019).

Flowering plants are one of the largest eukaryotic radiations with roughly 350,000 species and show a marked LDG (Francis & Currie 2003). Despite this pattern, latitudinal differences in macroevolutionary rates in angiosperms have received comparatively little attention at a species-level. Previous efforts have shown that angiosperm diversification is faster in tropical families (Jansson & Davies 2008) and in tropical lineages within a limited number of clades (Svenning *et al.* 2008; Jansson *et al.* 2013). The only species-level analysis found no differences in speciation, but did so by using a method that assumed that all rate differences in the angiosperm Tree of Life (ToL) were due to geographic distribution alone (Antonelli *et al.* 2015). This method ignores any other sources of rate variation and has been shown to be inadequate, especially for large phylogenies where substantial heterogeneity in diversification rates is expected (Caetano *et al.* 2018). One reason for the lack of attention to angiosperms is that their vast diversity makes it difficult to obtain well-sampled phylogenetic and distributional data at a species-level, which is required to estimate speciation rates accurately. Here we tested how speciation rates varied with latitudinal distribution in over 60,000 angiosperm species using different methods that accommodated macroevolutionary rate heterogeneity and incomplete taxon sampling. Our analyses used the largest phylogenetic trees presently available for angiosperms (Zanne *et al.* 2014; Smith & Brown 2018), alongside all geographic distribution data obtainable for these species (Mounce *et al.* 2017) from the Global Biodiversity Information Facility (GBIF). We show that latitudinal variation in recent speciation rates cannot generate the LDG because speciation was not faster in the tropics.

bioRxiv preprint doi: https://doi.org/10.1101/619064; this version posted January 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Material and Methods

Phylogenetic datasets and macroevolutionary rate estimates

We estimated speciation rates in the angiosperm radiation with Bayesian Analysis of Macroevolutionary Mixtures (BAMM) v.2.5.0 (Rabosky 2014). BAMM detects heterogeneity in speciation and extinction rates through time and across lineages. The input for BAMM was the most comprehensive time-calibrated phylogenetic tree for seed plants that was built using a hierarchical clustering analysis of Genbank sequence data (Smith & Brown 2018). On average, sister branches in this phylogeny had an overlap of 1792 base pairs, which corresponds to roughly one or two gene regions. We used a modified version of the tree containing GenBank taxa with a backbone provided by Open Tree of Life (i.e., GBOTB) for 72,986 flowering plants. Species names were standardised to The Plant List v.1.1 (Jin & Qian 2019). Following Igea *et al.* 2017, convergence of such a large BAMM analysis was achieved by dividing the initial tree into clades of less than 6000 species. Doing so resulted in 17 clades (16 monophyletic clades and 1 clade that contained the backbone of the tree plus the unassigned species), which we then used to run independent BAMM speciation/extinction analyses. Within each clade, non-random incomplete taxon sampling was incorporated into the estimation process by calculating the number of species sampled in each family. Prior settings were obtained using the *setBAMMpriors* function in the R package *BAMMtools* (Rabosky *et al.* 2014) and *expectedNumberOfShifts* was set at 50. Analyses were run for a minimum of 100 million generations and ≥20% of each run was discarded as a burn-in when necessary. Effective samples sizes of the log-likelihood and the number of rate shifts were both

bioRxiv preprint doi: https://doi.org/10.1101/619064; this version posted January 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

above 200. The event files for all 17 clades were then combined to generate a single BAMM result file that was used for all downstream analyses.

Although the reliability of BAMM estimates has recently been questioned (Moore *et al.* 2016), simulations have shown that accurate estimates of speciation can be obtained even in the absence of fossil data (Rabosky *et al.* 2017; Mitchell *et al.* 2018). Tip-based estimates of speciation like the ones we use here should also be robust to the effect of extinction and are increasingly used when analysing the effect of traits on lineage diversification and summarising geographic patterns of speciation rate variation (Title & Rabosky 2019). BAMM extinction estimates are thought to be more unreliable (Rabosky 2010; Marshall 2017; Mitchell *et al.* 2019), so we do not discuss them here. BAMM estimates of diversification rates, which are calculated as the net outcome of speciation minus extinction (Title & Rabosky 2019), will therefore also be more unreliable than speciation rates. Hence, and following other similar papers (Rabosky *et al.* 2018), we only discuss BAMM estimates of speciation. Furthermore, we also estimated recent speciation rates using the Diversification Rate metric (DR; Jetz *et al.* 2012, Supplementary Note 2), which only considers branch lengths and splitting events and does not accommodate for incomplete sampling. We also focus on estimates of speciation and extinction rates from other methods in subsequent sections.

Latitudinal datasets

We obtained latitudinal data for 236,894 angiosperms from a study that estimated the median, maximum and minimum latitude for each species (Mounce *et al.* 2017) using observations from GBIF. Since our goal was to assess whether speciation rates drive

bioRxiv preprint doi: https://doi.org/10.1101/619064; this version posted January 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

the LDG, we focused on latitudinal differences and did not consider regional and longitudinal differences in diversity, which may be important in some cases (Reyes *et al*. 2015; Xing & Ree 2017; Igea & Tanentzap 2019). We then intersected these data with our BAMM dataset (n = 72,986) to obtain estimates of speciation rate and latitude for as many angiosperms as possible, hereafter referred to as the 'full' dataset (n = 60,990). The species names in the GBIF datasets were standardised with The Plant List (TPL) using the *Taxonstand* (Cayuela *et al*. 2012) and *Taxonlookup* (Pennell *et al*. 2016) packages before collation.

Tropical and temperate species were defined as those where the absolute median latitude was below and above 23.5°, respectively. Species were also grouped into equal-width latitudinal bins according to their median latitude to define latitudinal bands. To ensure that these latitudinal bands were completely tropical or temperate, the width of each band was set at 9.4° (i.e., starting from a band centered in the Equator from -4.7° to 4.7°, then from 4.7° to 14.1°, 14.1° to 23.5°, etc).

BAMM-based correlation of tropicality and speciation

We correlated species-specific (i.e., tip-based) speciation rates from BAMM with the latitudinal data using Structured Rate Permutations on Phylogenies (STRAPP) (Rabosky & Huang 2016). STRAPP assesses the significance of the empirical association of macroevolutionary rates and phenotypic traits by comparing it to a null distribution generated by permuting the speciation rates across the phylogeny while maintaining the position of the rate shifts in the phylogenetic tree. Crucially, this test does not require that all rate variation is caused by the focal trait and thus shows a reduced Type I error rate (Rabosky & Huang 2016). We performed all STRAPP two-

bioRxiv preprint doi: https://doi.org/10.1101/619064; this version posted January 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

tailed tests using the *traitdependentBAMM* function of the *BAMMtools* package (Rabosky *et al.* 2014) with 1000 replicates and logging the rates. We assessed whether *i*) tropical and temperate species had different speciation rates using a Mann Whitney U test; *ii*) speciation rates for each of the latitudinal bands were different using the Kruskal-Wallis rank sum statistic; and *iii*) species absolute median latitude was correlated with speciation rate using Spearman's rank correlation coefficient.

We repeated the analyses discarding sets of species that may bias our results. First, we used only densely sampled species (i.e., with five or more data points in the GBIF dataset). Second, we discarded widespread species (i.e., neither strictly tropical nor strictly temperate) for analyses comparing strict tropical and temperate taxa. Strictly tropical and temperate species were defined as species with median, maximum and minimum latitude all occurring either inside or outside the tropics, respectively. Third, we discarded species with median latitudes in the highest latitudinal bands (>50° and <-50°), which contained few species with extreme estimates of λ.

Assessing the reliability of the BAMM results

We assessed the consistency of the BAMM results in the face of geographic biases, low sampling and topological uncertainty.

First, the full dataset had more temperate than tropical species and so was inconsistent with the LDG. Poorly sampled clades may have average longer branches and lower λ, and so our speciation rate estimates may be affected by latitudinal sampling bias. We therefore tested whether clades with lower sampling fractions were found at lower latitudes and had lower λ rates. We employed two additional strategies to analyse the

bioRxiv preprint doi: https://doi.org/10.1101/619064; this version posted January 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

effect of the disproportionate amount of temperate species in our BAMM analyses. We generated 100 random samples of our full dataset (herein 'unbiased' datasets), each with 30,000 species that maintained the proportion of tropical (57.5%) and temperate (42.5%) species present in GBIF. We also generated 100 random samples (herein 'extreme tropical' datasets) that assumed all species absent from the GBIF dataset of Mounce *et al.* (2017) were tropical, so the proportions of tropical and temperate species were 70.9% and 29.1%, respectively. We derived these values by estimating that 109,461 species were absent from GBIF by subtracting the number of species in the GBIF dataset from the number of species in The Plant List according to *TaxonLookup* (Pennell *et al.* 2016). Although both the unbiased and extreme tropical datasets were generated to assess the effect of excess temperate species, the extreme tropical dataset also represented the most extreme scenario of undescribed biodiversity. The sizes of these datasets was fixed at 30,000 species so that the tropical species present in each subsample were not excessively redundant when obtained from a total pool of 26,873 species by random draws of 17,130 and 21,273 species for the unbiased and extreme tropical datasets, respectively. Using the two datasets, we then assessed whether the full dataset λ estimates were affected by poorer sampling for tropical species. We did so by selecting 10 random replicates of the 30,000-species subsampled datasets and rerunning the BAMM speciation analyses as detailed above. We determined whether the relationship of the full with both the unbiased and extreme tropical λ estimates was different for tropical and temperate species. To do this, we predicted the full λ estimates using the subsampled λ estimates, 'tropicality' (i.e. a binary variable indicating if a species was exclusively tropical based on our aforementioned definition) and the interaction of the λ estimates and tropicality. We fitted this model separately for each of the 10 replicates in the

bioRxiv preprint doi: https://doi.org/10.1101/619064; this version posted January 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

unbiased and extreme tropical datasets, incorporating the effect of the phylogeny using the *phylolm* package (Tung Ho & Ané 2014). We also tested whether the association between speciation and latitude in the full dataset stemmed from the excessive proportion of temperate species by repeating the STRAPP analyses with the 100 replicates in the unbiased and extreme tropical datasets.

Second, our full dataset contained λ estimates for 60,990 species, which represent 17.6% of described angiosperm species (total n = 346,365 calculated with *Taxonlookup*). To assess the reliability of λ estimates given the large number of missing taxa, we randomly generated 10 subsamples ('small' datasets) with the same proportion (i.e., 17.6%) of the species in our full dataset (n = 10,739). As above, we ran BAMM for each of the 10 subsamples and predicted the full λ estimates using the small λ estimates.

Third, we analysed a single phylogenetic tree, which may only represent one of many potential hypotheses for the angiosperm ToL. To assess the effect of the unaccounted topological and branch length uncertainty, we tested the relationship between speciation rates and latitude in a smaller dataset of 28,057 species obtained from an independently estimated phylogenetic tree. The tree was estimated with GenBank sequences for seven gene regions with a median per-taxon proportion of missing data of 0.60 (Zanne *et al.* 2014; Qian & Jin 2016). We obtained BAMM λ rates for this dataset as described elsewhere (Igea *et al.* 2017). We then tested the correlation between the BAMM estimates of our full dataset and this reduced dataset. We also used the reduced dataset to determine whether there were differences in BAMM λ rates between tropical and temperate species and among different latitudinal bands;

bioRxiv preprint doi: https://doi.org/10.1101/619064; this version posted January 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

and whether absolute median latitude and speciation rates were associated. All of these analyses were performed as described above.

Geography-dependent speciation

We used the Geographic Hidden-State Speciation and Extinction (GeoHiSSE) framework to test for further differences in speciation between tropical and temperate lineages. GeoHiSSE estimates speciation, extinction and transition rates for two geographic areas while incorporating rate heterogeneity that is independent of geography by including unobserved characters known as "hidden states" (Caetano *et al.* 2018). Including these hidden states is crucial to create biologically meaningful null models where diversification varies independently of the focal traits. We considered any species occurring strictly below and above 23.5° as tropical and temperate respectively, while species whose range span both >23.5° and <23.5° were classified as widespread. Sampling fractions for each geographic state in the phylogenetic tree were estimated to be 0.12, 0.30, and 0.24 for strictly tropical, strictly temperate, and widespread species, respectively. Following Caetano *et al.* 2018, we fitted 12 different models (numbers 1-12 in Caetano *et al.* 2018, Table 1). Some of these were "null models" where no diversification rate variation is caused by geography. Other models included area-dependent and area-independent diversification with two to five hidden states, and some also separated range contraction from lineage extinction (i.e., "+extirpation models"). We then selected the best fitting model using the sample-sized corrected Akaike Information Criterion (AICc). We used the "fast version" of the GeoHiSSE model as implemented in *hisse* v.1.9.6 (Beaulieu & O'Meara 2016) to make analyses feasible in such a large phylogeny. The size of the dataset meant that some of the models took a very long

bioRxiv preprint doi: https://doi.org/10.1101/619064; this version posted January 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

time to fit (2 months on 2.5 GHz processor for a single fit). We could therefore not obtain rate estimates across the nodes and tips using the marginal reconstruction algorithm, which required estimating probabilities for all possible combinations of states at each node in the tree given all combinations at nodes in the rest of the tree. It was also not feasible to model-average the "effective state proportion" (i.e., the estimated proportion of each geographic area and hidden state combination), which would have allowed us to test if a only a small proportion of the lineages were associated with a geographic-dependent diversification process (Caetano *et al.* 2018).

Clade-based analyses

We estimated clade-level measures of speciation across the angiosperm phylogenetic tree. We used a set of 4 million-year-wide time intervals from 0 to 24 million years (myr) to define the ages of the clades. Clades were delimited as the largest non-overlapping monophyletic groups of four or more species where sampling was larger than 30%. This criterion ensures that the probability of retrieving the correct crown age of the clade is above 70% (Rabosky 2016). At least 50% of the species in the clades also had to have latitudinal data. We used RPANDA (R: Phylogenetic ANalyses of DiversificAtion, Morlon *et al.* 2016) to fit six different models of diversification: *i*) constant speciation (λ) and extinction (μ) fixed at 0; *ii*) constant λ and μ; *iii*) exponential λ and μ fixed at 0; *iv*) exponential λ and constant μ; *v*) constant λ and exponential μ; and *vi*) exponential λ and μ. Incomplete sampling within clades was incorporated into the estimates by calculating the proportion of sampled species within the genera in each clade. Speciation rates were estimated using model averaging with AIC weights. The correlation of the clade-level estimates of speciation and the proportion of temperate and tropical species in each clade was assessed using

bioRxiv preprint doi: https://doi.org/10.1101/619064; this version posted January 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

phylogenetic least squares regression as implemented in the R package *nlme* (Pinheiro *et al.* 2014). The lambda parameter was optimised using a maximum likelihood approach. We weighted the regressions by the proportion of GBIF data points in each clade to account for different sampling levels. We also assessed if potential latitudinal differences in sampling proportions could explain our results. We did so by testing if the sampling proportions of largely tropical and largely temperate clades (defined as clades where >50% of the species were tropical and temperate, respectively) were different across the 6 time intervals using Pearson's chi-squared test.

Results

We found that tip-specific speciation rates were consistently smaller in the tropics for 60,990 species across the whole angiosperm radiation. These results were repeatable with several ways of classifying species as tropical or temperate. Using correlation tests that controlled for phylogenetic pseudoreplication (Rabosky & Huang 2016), we found that tropical species had smaller mean speciation rates (λ) than temperate species (Fig 1a; $\lambda_{temperate}$ = 0.73 species/myr, $\lambda_{tropical}$ = 0.65 species/myr, p-value = 0.043). Angiosperms in latitudinal bands closer to the poles (i.e. absolute median latitude >42.3°) speciated at faster rates than in any tropical latitudinal band (Fig. 1b). The absolute median latitude occupied by individual species was also positively correlated with their speciation rate (Fig. 1c; Spearman's ρ = 0.086, p-value = 0.009). These results were maintained when we: *i*) analysed only those species that were densely sampled in GBIF (n = 53,344 species) and *ii*) discarded the highest latitudinal bands (>50° N and S), which contained fewer species (n = 60,034) (Table S1). Speciation rates were also different when comparing strictly tropical and strictly temperate species by discarding species that occurred both inside and outside the

bioRxiv preprint doi: https://doi.org/10.1101/619064; this version posted January 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

tropics (n = 46,426), but this difference was not significant (p-value = 0.064; Table S1). .

Our full dataset contained a disproportionate amount of temperate species and so did not hold a LDG (Fig. S1), but this bias could not explain the consistent association between latitude and speciation rates. First, we generated 100 "unbiased" datasets that maintained the proportion of tropical and temperate species present in GBIF and were consistent with the LDG. Second, we generated another set of 100 "extreme tropical" datasets, where the proportions of temperate and tropical species were 29.1% and 70.9%, respectively. Rerunning BAMM for 10 of the 100 unbiased and extreme tropical datasets, we found that the λ estimates were very positively correlated with those from the full dataset (median $\rho = 0.82$, median p-value < 0.001 for both the unbiased and the extreme tropical datasets; Figs 2, S2). This correlation between λ of the full and subsampled datasets did not differ between tropical and temperate species, as the effect size of the interaction between tropicality and the subsampled λ estimates was not different from zero across all the 10 replicates for the unbiased (mean = -0.0005, standard deviation = 0.0017) and the extreme tropical (mean = -0.0015, standard deviation = 0.0024) datasets. As with the full dataset, we also found that speciation rates in the tropics were consistently lower and that λ was positively correlated with absolute median latitude using STRAPP across the 100 replicates of both the unbiased (Fig. 3) and extreme tropical (Fig. S3) datasets. Finally, we confirmed that our results could not be explained by better sampled temperate species having higher λ. Densely sampled clades were only slightly more frequent at higher latitudes (Fig. S4), and they had smaller estimates of λ (Fig. S5), contrary to what we might expect if our results were a sampling artefact.

bioRxiv preprint doi: https://doi.org/10.1101/619064; this version posted January 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

We also confirmed that the speciation rate estimates were robust to the low levels of taxon sampling. We compared the BAMM λ estimates from the full dataset with estimates generated by randomly sampling 17.6% of the species in our full dataset 10 times and rerunning BAMM. This level of sampling reflected the proportion of angiosperm species included in our full dataset from those described in The Plant List. The resulting λ estimates from the full and the small datasets were positively correlated (median $\rho = 0.67$, median p-value < 0.001). This association was slightly weaker for tropical than for temperate species, but this difference was very small and not statistically significant across the 10 replicates (mean = -0.04, standard deviation = 0.03; Fig. S6).

Our results were based on a single phylogenetic tree that represented one of many possible topologies in the angiosperm ToL, and so we confirmed that the results were consistent with another independently derived phylogeny (Zanne *et al.* 2014; Qian & Jin 2016). Using BAMM λ estimates and latitudinal information for this 28,057 species tree, we found that λ was higher outside the tropics and that speciation rates increased towards the poles (Fig. S7). The BAMM λ estimates of this topology and our full dataset were positively correlated (Spearman's $\rho = 0.46$, p-value < 0.001).

Elevated speciation rates in the tropics were also rejected as the driving force of the LDG using a method that explicitly tested the influence of geography on species diversification. The best fitting model with GeoHiSSE (AICc weight = 1; Table 1) had five different hidden states that accommodated rate variation across lineages and no differences in turnover and extinction fraction between tropical and temperate

bioRxiv preprint doi: https://doi.org/10.1101/619064; this version posted January 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

regions. This result suggests that most of the variation in diversification rates across the angiosperm radiation is not linked to latitude, though some angiosperm lineages can show range-dependent diversification. We were precluded from exploring this latter process using marginal reconstructions by the sheer size of the dataset (see Methods). The confidence intervals around the maximum-likelihood estimates were too wide to identify any differences in rate estimates, so we do not discuss them further.

Finally, we found that tropical species did not speciate faster across clades of different ages in the angiosperm phylogeny. Using clades in six 4 million-year-wide time slices, from 0 to 24 million years, where we were confident in speciation rates and GBIF coverage, we fitted models that implemented time-variable and time-constant speciation and extinction. Mirroring the BAMM results, clades with a larger proportion of temperate species had higher rates of speciation in the best fitting model (Fig S8). This result was not totally explained by sampling biases. Temperate clades had significantly denser average sampling only in the 8-12 and 12-16 myr intervals ($\chi2 = 5.22$, df = 1, p-value = 0.020 and $\chi2 = 12.19$, df = 1 p-value < 0.001, respectively), and these differences were small (0.65 vs 0.62 and 0.65 vs 0.61 for temperate vs tropical in each period, respectively).

Discussion

Using data for over 60,000 flowering plants, our findings rejected the long-standing notion that the greater species diversity of the tropics can be explained by higher recent speciation rates (Mittelbach *et al.* 2007), such as from higher environmental energy (Dowle *et al.* 2013). We instead found higher rates of recent speciation closer

bioRxiv preprint doi: https://doi.org/10.1101/619064; this version posted January 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

to the poles. An explanation for this difference is that temperate biotas have fewer species, so their niche space may be less saturated, thereby increasing opportunities for lineage divergence (Simpson 1953; Schluter 2016). Similarly, reproductive isolation may be elevated at higher latitudes by the greater ecological opportunity generated from recurrent environmental change and climate instability (Cutter & Gray 2016). Other explanations may be related to the ecological and life history traits of temperate species. For example, faster speciation rates at higher latitudes could stem from the higher frequency of small-seeded species (Moles *et al.* 2007). Small seed size is positively correlated with angiosperm diversification (Igea *et al.* 2017), and temperate species did have smaller seeds in our full dataset (Supplementary Note 1). Therefore, our results suggest that latitudinal differences in recent speciation rates do not shape the LDG but are shaped by differences in species diversity and traits (Weir & Price 2011; Schluter & Pennell 2017).

Rejecting the role of speciation rate as the cause of the LDG calls for alternative explanations (Antonelli *et al.* 2015). Many theories have been proposed to explain the LDG, but they can all be divided into four broad categories arising from differences in: *i)* diversification rates; *ii*) dispersal rates; *iii)* time for species accumulation and *iv)* ecological limits to species diversity (Etienne *et al.* 2019). Hypotheses in the different categories often invoke multiple causal mechanisms and so may not be mutually exclusive (Pontarp *et al.* 2019). First, the elevated speciation near the poles could be coupled with higher extinction rates to produce a scenario where temperate zones act as both a “cradle” and “grave” of biodiversity (Cutter & Gray 2016). Net diversification might therefore still be higher in tropical areas despite the absolute rates of speciation and extinction each being smaller than closer to the poles. Our

bioRxiv preprint doi: https://doi.org/10.1101/619064; this version posted January 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

GeoHiSSE analyses do not, however, support this scenario as we found no such differences in turnover (Table 1). Estimates of extinction rate from molecular phylogenies in the absence of fossil data are often unreliable (Marshall 2017; Mitchell *et al.* 2018), which further complicates direct tests of this hypothesis. Second, the tropical conservatism hypothesis proposes that lower net migration from the tropics to temperate zones contributes to the LDG (Wiens & Donoghue 2004), because most lineages originated in the tropics and have been unable to disperse into harsher, temperate environments. For example, in New World woody angiosperms, transitions between tropical and temperate environments are rare and most temperate lineages are relatively young (Kerkhoff *et al.* 2014). This scenario was also not supported with GeoHiSSE as we detected no differences in transition rates between the tropics and the temperate zones and vice versa. Third, if rates of diversification and migration have no latitudinal variation, the greater diversity of the tropics may stem from their larger area through time, i.e. the age-and-area hypothesis (Fine & Ree 2006). For instance, current species richness across the world's 32 bioregions is best predicted by area and productivity over geological time (Jetz & Fine 2012). Area-through-time differences have also been proposed to underlie differences in diversity among tropical regions (Couvreur 2015). Fourth, a whole class of hypotheses posits that there are ecological limits to the number of species that a region can hold, and that these limits vary with latitude (Rabosky 2009). For example, the larger diversity in the tropics could be explained by larger niche spaces and greater niche packing (Pellissier *et al.* 2018). Such a mechanism could interact with differential diversification rates to amplify spatial differences in diversity (Etienne et al. 2019). As ecological and macroevolutionary processes can interact, integrating all of them into a mechanistic eco-evolutionary framework can ultimately advance the search for explanations to the

bioRxiv preprint doi: https://doi.org/10.1101/619064; this version posted January 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

LDG (Pontarp *et al*. 2019). Other ecological interactions, like negative conspecific density dependence arising from herbivores and pathogens, have also been proposed to explain the high diversity in the tropics (LaManna *et al*. 2017).

Recent findings have suggested that rates of species origination are highest in species-poor areas (Quintero & Jetz 2018; Rabosky *et al.* 2018). Our results are consistent with these findings and therefore suggest that latitudinal variation in recent speciation rates is not a major engine of global diversity patterns in one of the largest eukaryotic radiations. Using time-heterogeneous estimates of speciation, we were able to exclude faster rates of evolution, such as arising from more favourable environmental conditions or more intense biotic interactions, as responsible for shaping the angiosperm LDG. Future work must now resolve whether recent rapid speciation nearer the poles is widespread across the ToL and its underlying causes.

bioRxiv preprint doi: https://doi.org/10.1101/619064; this version posted January 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

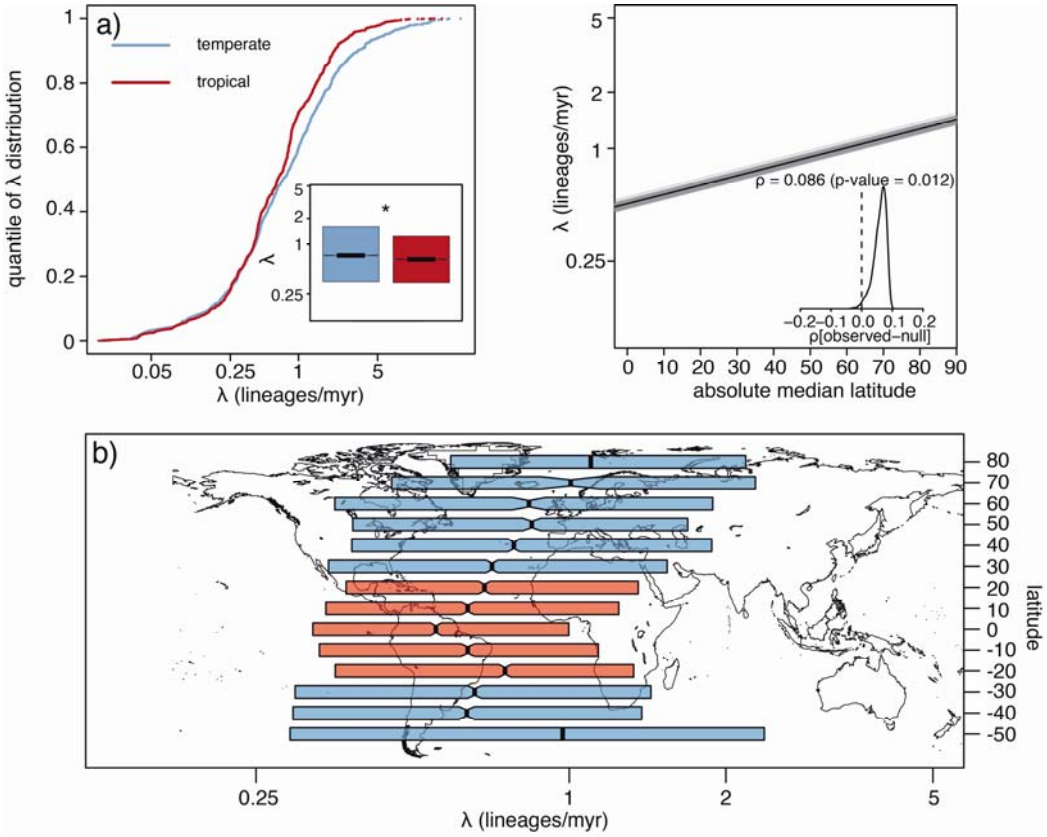

Figure 1. Tropical species have smaller speciation rates (λ) than temperate species. a) Rank-ordered distribution and boxplot (inset) of tip λ inferred with BAMM for tropical (red) and temperate (blue) species. * indicates significant difference (p-value = 0.043) between the two groups, assessed with a Mann-Whitney test implemented in STRAPP. **b)** λ grouped by latitudinal band of each species. Notches in boxplots indicate 95% confidence intervals around median, denoted by thick vertical lines and boxes span the interquartile ranges. **c)** Spearman's ρ correlation between species absolute median latitude and λ as estimated with STRAPP. Grey lines are correlations across 1000 samples of the posterior distribution estimated with BAMM. Solid black line indicates median correlation ± 95% confidence interval across the posterior distribution. Inset shows the difference between the empirical and null correlations estimated with 1000 permutations of evolutionary rates across the phylogeny.

bioRxiv preprint doi: https://doi.org/10.1101/619064; this version posted January 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

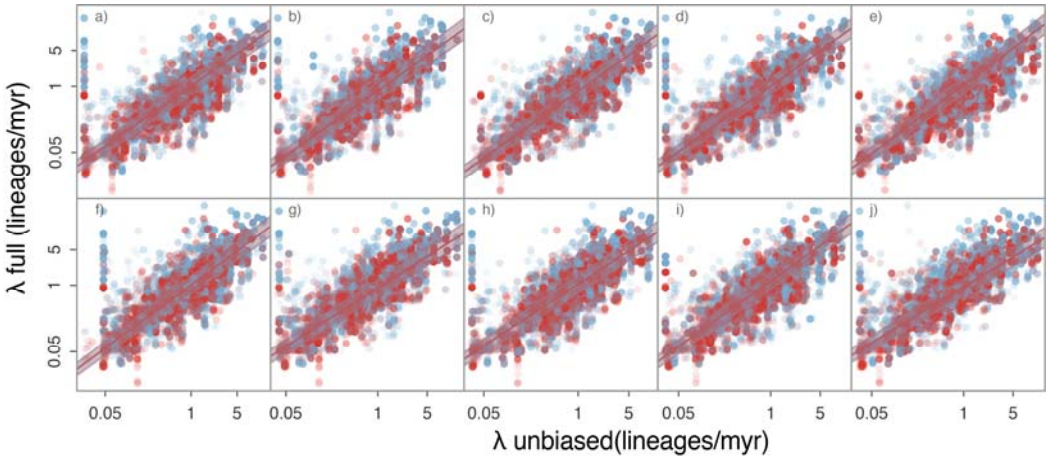

Figure 2. BAMM λ estimates were positively correlated in the full and 10 unbiased datasets (a to j) with no effect of tropicality. Solid lines are phylogenetic linear regressions predicting λ in the full tree (n = 60,990 species) with λ in the unbiased tree (n = 30,000 species) in temperate (shown in blue) and tropical (shown in red) species. Shaded areas indicate the 95% confidence intervals.

bioRxiv preprint doi: https://doi.org/10.1101/619064; this version posted January 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

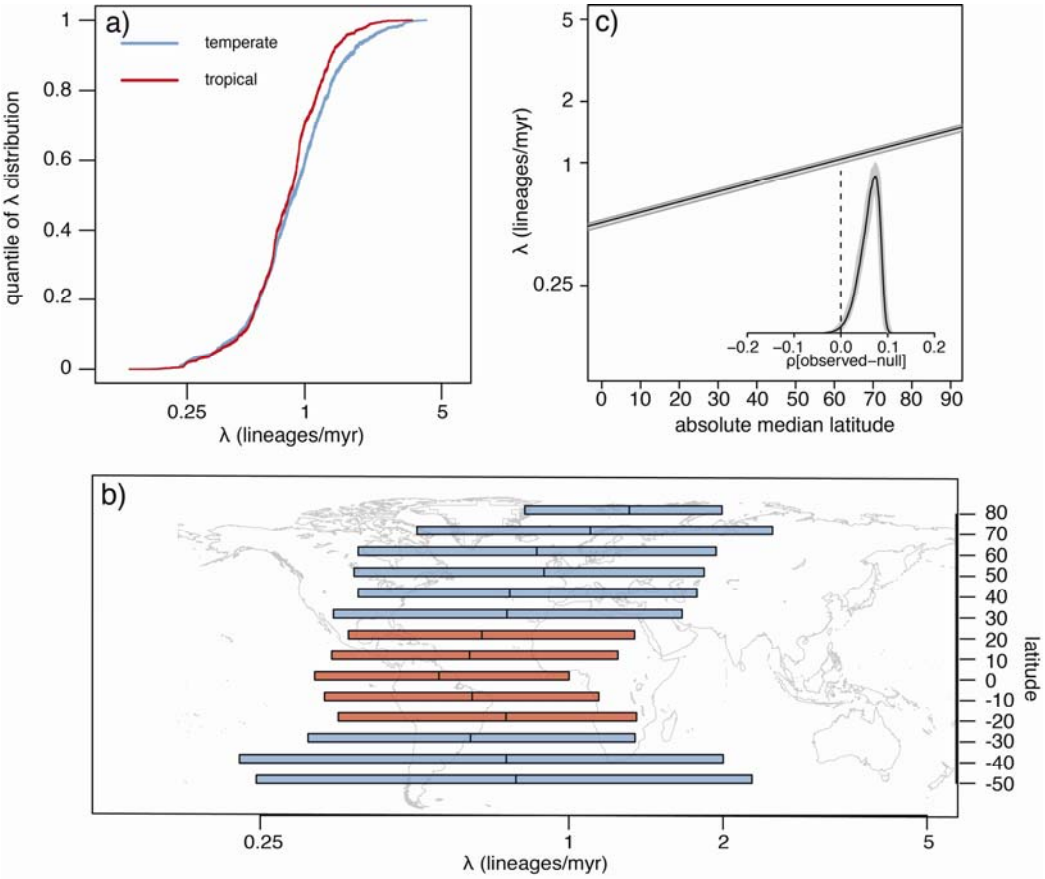

Figure 3. Temperate species have higher speciation rates (λ) than tropical species for 100 subsampled "unbiased" datasets. Datasets were generated whereby 42.5% and 57.5% of species were temperate and tropical, respectively. **a)** Rank-ordered distribution of λ inferred with BAMM for tropical (red) and temperate (blue) species. **b)** λ grouped by latitudinal band of each species. Boxes show the average interquartile range across the 100 subsampled datasets, and lines show the average medians. **c)** Spearman's ρ correlation of species absolute median latitude and λ as estimated with STRAPP. Solid black line indicates median correlation ± 95% confidence interval across the posterior distribution. Inset shows the median (black line) ± 95% confidence interval for the difference between the empirical and null correlations estimated with 1000 permutations of evolutionary rates across the phylogeny in each of 100 subsampled datasets.

bioRxiv preprint doi: https://doi.org/10.1101/619064; this version posted January 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Tables

Table 1. Model comparison of 12 GeoHiSSE models with up to five hidden rate categories. Model numbers and descriptions refer to Caetano *et al* (2018). Null models have no differences in diversification associated with geography and diversification parameters are constrained to be equal between areas in the same hidden state category. Area independent models (CID) have no geographic-dependent differences in diversification. Full models have no constrained parameters. Extremely high AICc values for the three simplest models suggest that they were a very poor fit. Original GeoSSE refers to the formulation in Goldberg *et al.* (2011).

model	description	# free parameters	AICc
5	CID-GeoHiSSE, 5 hidden rate classes, null model	13	418090
11	CID-GeoHiSSE+extirpation, 5 hidden rate classes, null model	15	423888
10	GeoHiSSE+extirpation, 2 hidden rate classes, full model	19	425927
9	CID-GeoHiSSE+extirpation, 3 hidden rate classes, null model	11	428085
12	CID-GeoHiSSE+extirpation, 2 hidden rate classes	9	429549
4	GeoHiSSE, 2 rate classes, full model	15	430692
3	CID-GeoHiSSE, 3 hidden rate classes, null model	9	442036
8	GeoSSE+extirpation, full model	9	458691
2	Original GeoSSE, full model	7	470420
1	CID-original GeoSSE	4	1.56E+19
6	CID-GeoHiSSE, 2 hidden rate classes	7	1.56E+19
7	CID-GeoSSE+extirpation	6	1.56E+19

bioRxiv preprint doi: https://doi.org/10.1101/619064; this version posted January 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

SUPPLEMENTARY MATERIAL

dataset name	$\lambda_{temperate}$	$\lambda_{tropical}$	p-value	dataset size
more.GBIF.data	0.702	0.641	0.050	53,344
no.widespread	0.760	0.668	0.064	46,426
no.high.latitude	0.719	0.652	0.044	60,034

Table S1. Temperate species have mean higher speciation rates (λ, estimated with STRAPP) than tropical species across different datasets: "more.GBIF.data" excludes species with less than five GBIF data points; "no.widespread" excludes species occurring both in and outside the tropics; "no.high.latitude" excludes species from poorly sampled latitudinal bands with absolute median latitude ≥50°. Units for λ and dataset size are lineages/myr and species, respectively.

Supplementary Note 1

Smaller seeded species have been shown to speciate faster than large seeded species (Igea *et al.* 2017). We assessed whether seed size differed between temperate and tropical lineages as it has previously been shown to do so (Moles *et al.* 2007), and so latitudinal variation in seed size could explain the differences in speciation between tropical and temperate species. We obtained seed size measurements for 13,178 species in our full dataset from a previous study (Igea *et al.* 2017) and we found that seed size was larger in tropical species (mean seed $size_{temperate}$ = 0.0018 g; mean seed $size_{tropical}$ = 0.0125 g; t = -28.239; df = 4300.7; p-value < 0.001). This difference remained when phylogeny was considered (phylANOVA: t = 33.933; p value = 0.001; significance assessed with 1,000 random simulations using phytools (Revell 2012))

bioRxiv preprint doi: https://doi.org/10.1101/619064; this version posted January 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Supplementary Note 2

BAMM is a model-based approach to estimate tip speciation rates. To compare the BAMM results with a non-model based approach (i.e., that only relies on branch lengths and splitting events), we calculated the Diversification Rate metric (DR; Jetz *et al.* 2012) for the species in our full dataset (n = 60,990). This measure of recent speciation rate incorporates the number of nodes and the internode distances separating a species from the root and gives greater weight to branches closer to the present (Title & Rabosky 2019). We then compared the DR values of tropical and temperate species and assessed the variation of DR across latitudinal bands. Mirroring the BAMM results, we found that tropical species had smaller values of DR and that latitudinal bands closer to the poles had higher values of DR (Fig. S9).

bioRxiv preprint doi: https://doi.org/10.1101/619064; this version posted January 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

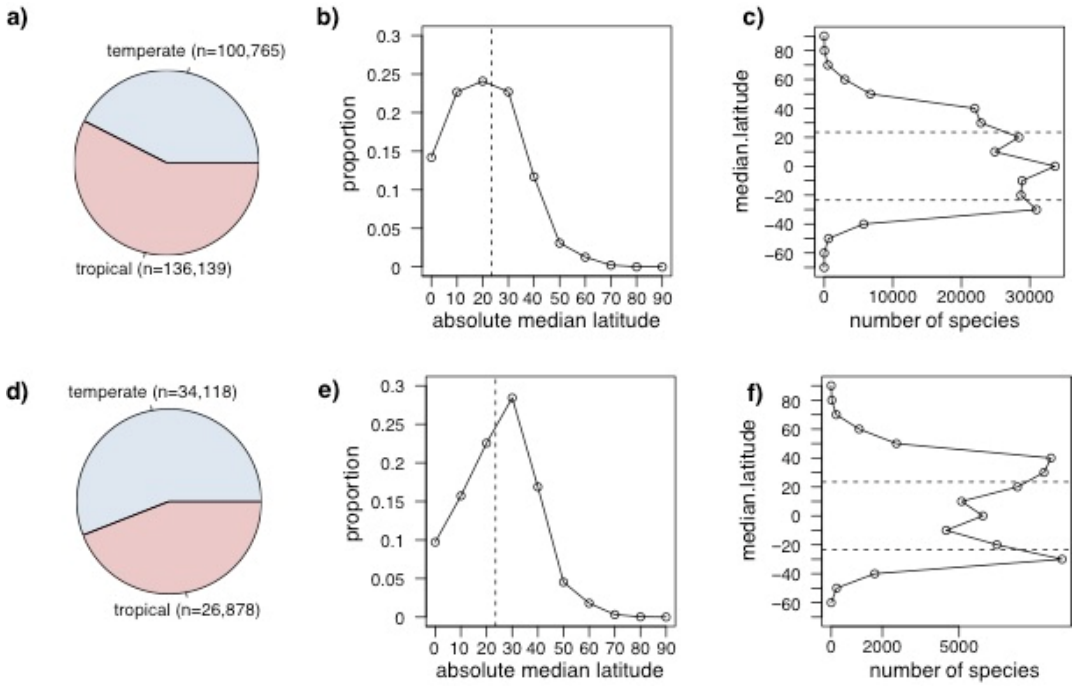

Figure S1. The full dataset had a disproportionate number of temperate species. **a)** Proportion of tropical and temperate species in the GBIF dataset (n = 236,894); **b)** proportion of species in each latitudinal band in the GBIF dataset; **c)** number of species in each latitudinal band in the GBIF dataset; **d)** proportion of tropical and temperate species in the full dataset (n = 60,990); **e)** proportion of species in each latitudinal band in the full dataset; **f)** number of species in each latitudinal band in the full dataset; in **b), c), e)** and **f)**, the dotted vertical line denotes the 23.5° threshold between tropical and temperate zones.

bioRxiv preprint doi: https://doi.org/10.1101/619064; this version posted January 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

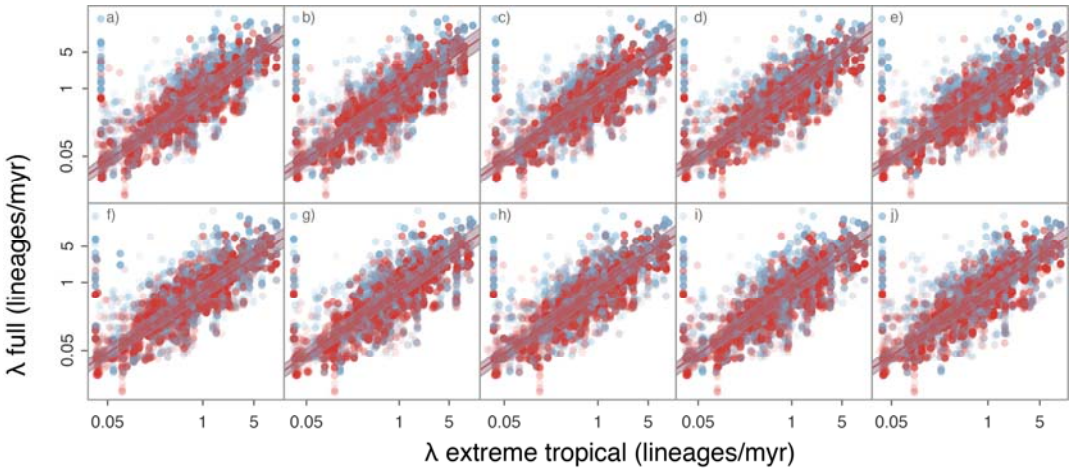

Figure 2. BAMM λ estimates were positively correlated in the full and 10 extreme tropical datasets (a to j) with no effect of tropicality on this correlation. Solid lines are phylogenetic linear regressions predicting λ in the full tree (n = 60,990 species) with λ in the extreme tropical tree (n = 30,000 species) in temperate (blue) and tropical (red) species. Shaded areas indicate the 95% confidence intervals.

bioRxiv preprint doi: https://doi.org/10.1101/619064; this version posted January 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

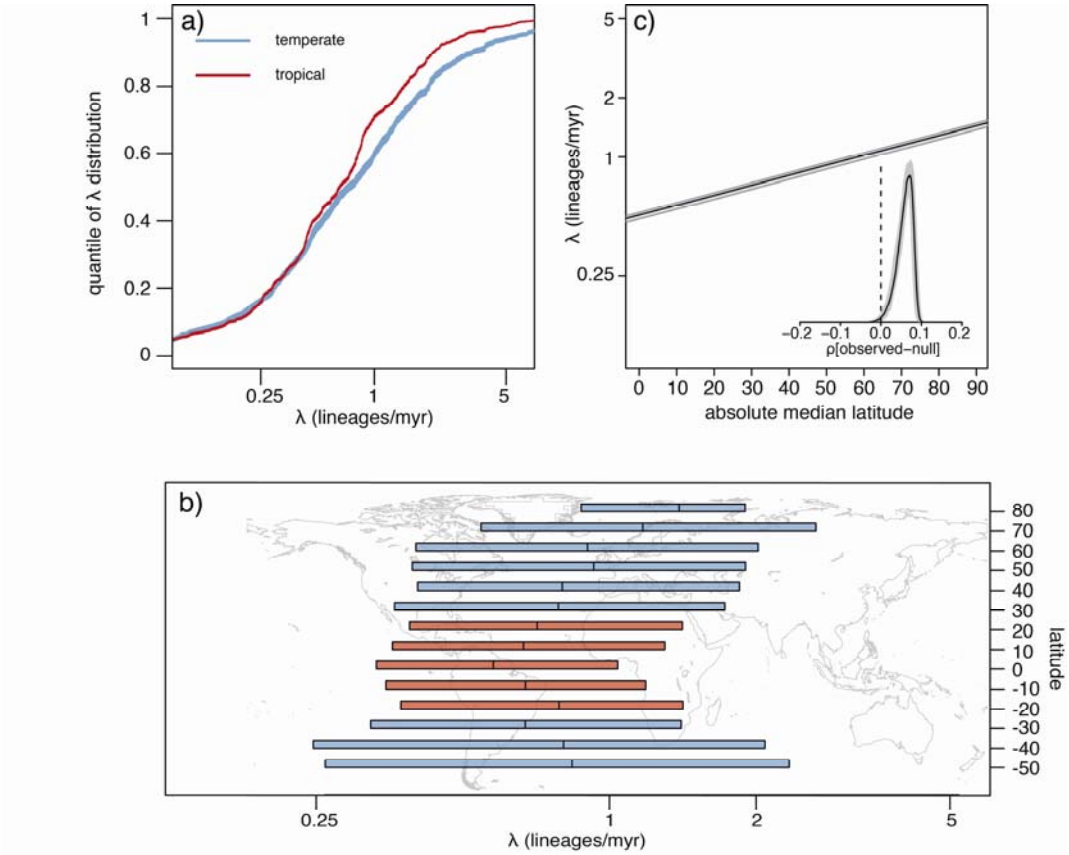

Figure S3. Temperate species have higher speciation rates (λ) than tropical species for 100 subsampled "extreme tropical" datasets with 29.1% temperate and 70.9% tropical species. **a)** Rank ordered distribution of λ inferred with BAMM for tropical (red) and temperate (blue) species. **b)** λ grouped by latitudinal band of each species. Boxes and lines are presented as in Fig. S4. **c)** Spearman's ρ correlation of species absolute median latitude and λ as estimated with STRAPP. . Solid black line indicates median correlation ± 95% confidence interval across the posterior distribution. The inset shows the median (black line) ± 95% confidence interval for the difference between the empirical and null correlations estimated with 1000 permutations of evolutionary rates across the phylogeny in each of 100 subsampled datasets.

bioRxiv preprint doi: https://doi.org/10.1101/619064; this version posted January 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

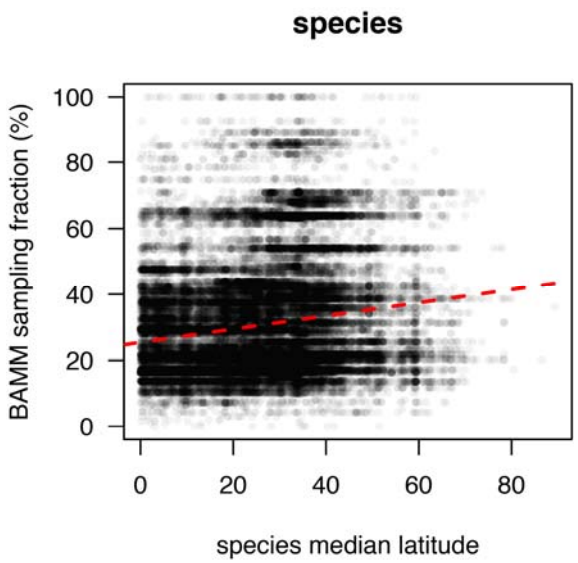

Figure S4. Relationship between the species absolute median latitude and the family-level sampling fraction used in the BAMM analyses. The red line is the slope of the linear regression (slope = 0.201, p-value < 0.0001).

bioRxiv preprint doi: https://doi.org/10.1101/619064; this version posted January 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

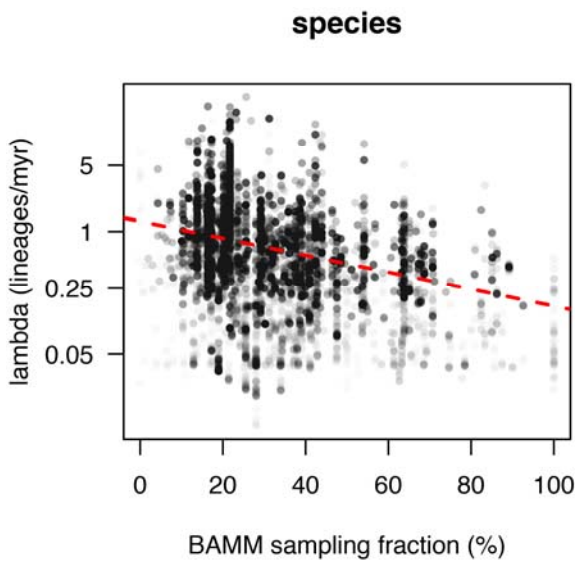

Figure S5. More densely sampled clades have smaller estimates of speciation rate (λ). The dotted line is the slope of the linear regression of the log(λ) and the BAMM sampling fraction (slope = -0.027 , p-value < 0.0001).

bioRxiv preprint doi: https://doi.org/10.1101/619064; this version posted January 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

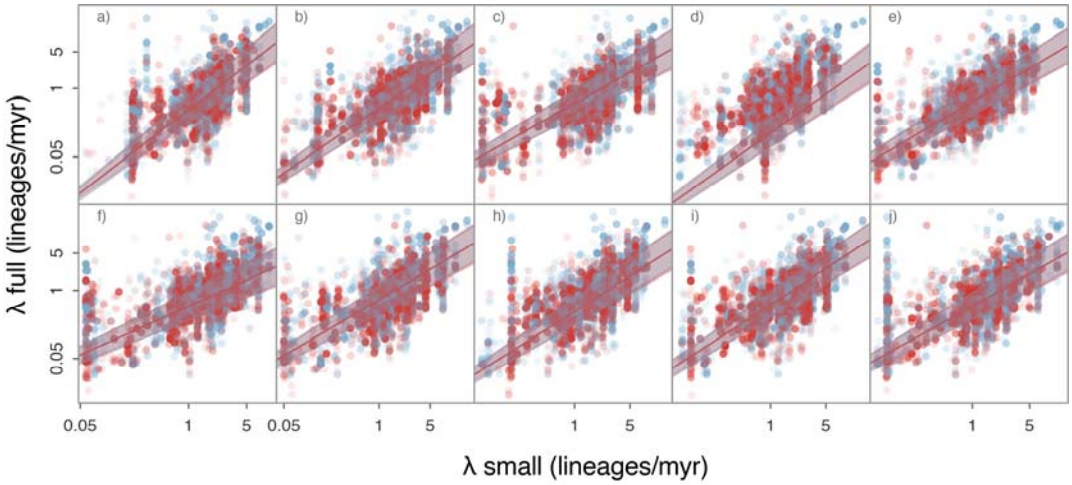

Figure S6. BAMM λ estimates are positively correlated in the full and 10 small datasets and tropicality has no effect on this correlation. **a) - j)** Solid lines are phylogenetic linear regressions predicting λ in the full tree (n = 60,990 species) with λ in the small tree (n = 10,739 species) in temperate (shown in blue) and tropical (shown in red) species. Shaded areas indicate the 95% confidence intervals.

bioRxiv preprint doi: https://doi.org/10.1101/619064; this version posted January 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

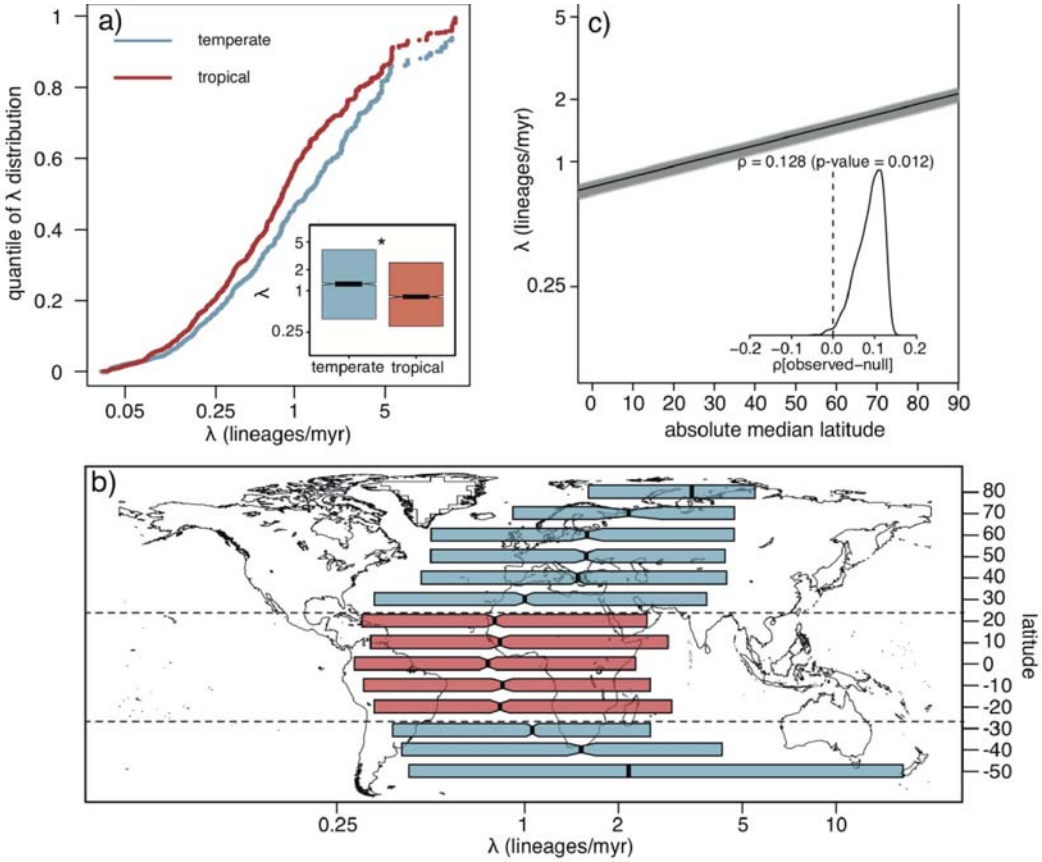

Figure S7. Tropical species have smaller speciation rates (λ) than temperate species with the Zanne phylogeny (n = 28,057 species). a) Rank ordered distribution and boxplot (inset) of λ inferred with BAMM for tropical (red) and temperate (blue) species; b) λ grouped by latitudinal band of each species; and c) Spearman's ρ correlation between species absolute median latitude and λ as estimated with STRAPP. Lines, boxes and symbols as in Fig. 1.

bioRxiv preprint doi: https://doi.org/10.1101/619064; this version posted January 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

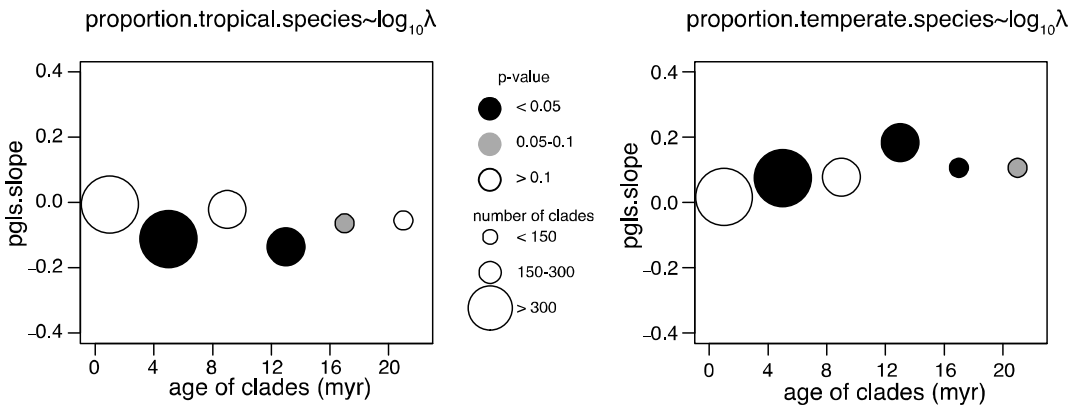

Figure S8. Temperate clades have higher speciation rates in the clade-based analysis. Correlation of a) the proportion of tropical and b) temperate species in each clade with the corresponding speciation rate (λ) estimated with RPANDA. The correlation coefficient is the phylogenetic generalised least squares (PGLS) slope. The size of the points is scaled to the number of clades in each time interval and their colours show the statistical significance of the slope.

bioRxiv preprint doi: https://doi.org/10.1101/619064; this version posted January 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

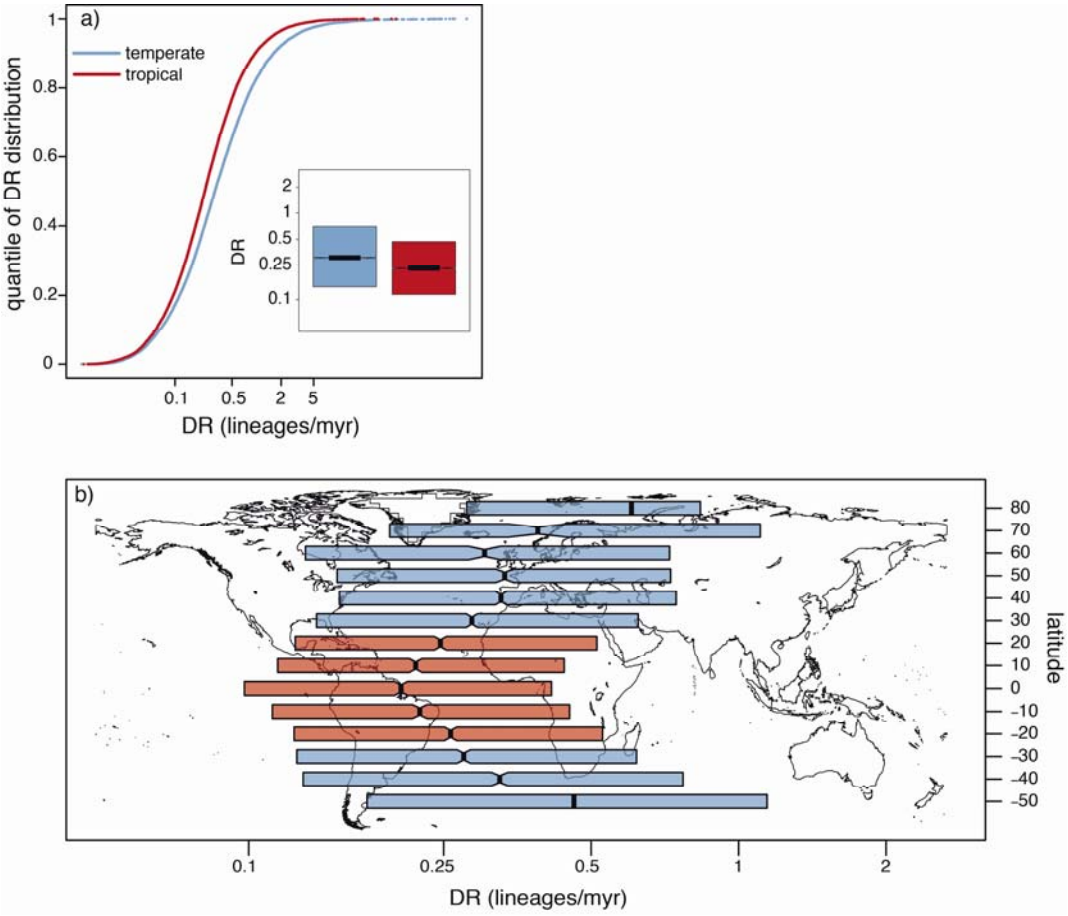

Figure S9. Tropical species have smaller DR values than temperate species. a) Rank-ordered distribution and boxplot (inset) of DR for tropical (red) and temperate (blue) species. **b)** DR grouped by latitudinal band of each species. Notches in boxplots indicate 95% confidence intervals around median, denoted by thick vertical lines and boxes span the interquartile ranges.

<u>Acknowledgments</u>

We thank the Wellcome Trust (Grant Number 105602/Z/14/Z), Gatsby Charitable Foundation (Grant Number GAT2962) and Isaac Newton Trust (Grant Number 17.24r) for funding. We thank the Handling Editor, Hervé Sauquet and two anonymous reviewers for comments that improved the manuscript.

bioRxiv preprint doi: https://doi.org/10.1101/619064; this version posted January 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Data accessibility:

Scripts and data to reproduce the analyses will be deposited upon acceptance in Figshare

Declaration of interests

The authors declare no competing interests.

References

Antonelli, A., Zizka, A., Silvestro, D., Scharn, R., Cascales-Miñana, B. & Bacon, C.D. (2015). An engine for global plant diversity: highest evolutionary turnover and emigration in the American tropics. *Front. Genet.*, 6.

Beaulieu, J.M. & O'Meara, B.C. (2016). Detecting Hidden Diversification Shifts in Models of Trait-Dependent Speciation and Extinction. *Syst. Biol.*, 65, 583–601.

Brown, J.H. (2014). Why are there so many species in the tropics? *J. Biogeogr.*, 41, 8–22.

Caetano, D.S., O'Meara, B.C. & Beaulieu, J.M. (2018). Hidden state models improve state-dependent diversification approaches, including biogeographical models. *Evolution*, 72, 2308–2324.

Cayuela, L., Cerda, Í.G. la, Albuquerque, F.S. & Golicher, D.J. (2012). taxonstand: An r package for species names standardisation in vegetation databases. *Methods Ecol. Evol.*, 3, 1078–1083.

Condamine, F.L., Sperling, F.A., Wahlberg, N., Rasplus, J.-Y. & Kergoat, G.J. (2012). What causes latitudinal gradients in species diversity? Evolutionary processes and ecological constraints on swallowtail biodiversity. *Ecol. Lett.*, 15, 267–277.

Couvreur, T.L.P. (2015). Odd man out: why are there fewer plant species in African rain forests? *Plant Syst. Evol.*, 301, 1299–1313.

Cutter, A.D. & Gray, J.C. (2016). Ephemeral ecological speciation and the latitudinal biodiversity gradient. *Evolution*, 70, 2171–2185.

Dowle, E.J., Morgan-Richards, M. & Trewick, S.A. (2013). Molecular evolution and the latitudinal biodiversity gradient. *Heredity*, 110, 501–510.

bioRxiv preprint doi: https://doi.org/10.1101/619064; this version posted January 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Economo, E.P., Narula, N., Friedman, N.R., Weiser, M.D. & Guénard, B. (2018). Macroecology and macroevolution of the latitudinal diversity gradient in ants. *Nat. Commun.*, 9, 1778.

Etienne, R.S., Cabral, J.S., Hagen, O., Hartig, F., Hurlbert, A.H., Pellissier, L., *et al.* (2019). A Minimal Model for the Latitudinal Diversity Gradient Suggests a Dominant Role for Ecological Limits. *Am. Nat.*, 194, E122–E133.

Fine, P.V.A. & Ree, R.H. (2006). Evidence for a time□Integrated species□area effect on the latitudinal gradient in tree diversity. *Am. Nat.*, 168, 796–804.

Francis, A.P. & Currie, D.J. (2003). A globally consistent richness□climate relationship for angiosperms. *Am. Nat.*, 161, 523–536.

Ghalambor, C.K., Huey, R.B., Martin, P.R., Tewksbury, J.J. & Wang, G. (2006). Are mountain passes higher in the tropics? Janzen's hypothesis revisited. *Integr. Comp. Biol.*, 46, 5–17.

Goldberg, E.E., Lancaster, L.T. & Ree, R.H. (2011). Phylogenetic Inference of Reciprocal Effects between Geographic Range Evolution and Diversification. *Syst. Biol.*, 60, 451–465.

Hawkins, B.A. (2001). Ecology's oldest pattern? *Trends Ecol. Evol.*, 16, 470.

von Humboldt, A. (1808). *Ansichten der Natur mit wissenschaftlichen Erlauterungen*. J.G. Cotta, Tubingen, Germany.

Igea, J., Miller, E.F., Papadopulos, A.S. & Tanentzap, A.J. (2017). Seed size and its rate of evolution correlate with species diversification across angiosperms. *PLoS Biol.*, 15, e2002792.

Igea, J. & Tanentzap, A.J. (2019). Multiple macroevolutionary routes to becoming a biodiversity hotspot. *Sci. Adv.*, 5, eaau8067.

Jansson, R. & Davies, T.J. (2008). Global variation in diversification rates of flowering plants: energy vs. climate change. *Ecol. Lett.*, 11, 173–183.

Jansson, R., Rodríguez-Castañeda, G. & Harding, L.E. (2013). What can multiple phylogenies say about the latitudinal diversity gradient? A new look at the tropical conservatism, out of the tropics, and diversification rate hypotheses. *Evolution*, 67, 1741–1755.

Jetz, W. & Fine, P.V.A. (2012). Global Gradients in Vertebrate Diversity Predicted by Historical Area-Productivity Dynamics and Contemporary Environment. *PLOS Biol.*, 10, e1001292.

Jetz, W., Thomas, G.H., Joy, J.B., Hartmann, K. & Mooers, A.O. (2012). The global diversity of birds in space and time. *Nature*, 491, 444–448.

bioRxiv preprint doi: https://doi.org/10.1101/619064; this version posted January 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Jin, Y. & Qian, H. (2019). V.PhyloMaker: an R package that can generate very large phylogenies for vascular plants. *Ecography*, 42, 1353–1359.

Kerkhoff, A.J., Moriarty, P.E. & Weiser, M.D. (2014). The latitudinal species richness gradient in New World woody angiosperms is consistent with the tropical conservatism hypothesis. *Proc. Natl. Acad. Sci.*, 111, 8125–8130.

LaManna, J.A., Mangan, S.A., Alonso, A., Bourg, N.A., Brockelman, W.Y., Bunyavejchewin, S., *et al.* (2017). Plant diversity increases with the strength of negative density dependence at the global scale. *Science*, 356, 1389–1392.

Louca, S. & Pennell, M.W. (2019). Phylogenies of extant species are consistent with an infinite array of diversification histories. *bioRxiv*, 719435.

Marshall, C.R. (2017). Five palaeobiological laws needed to understand the evolution of the living biota. *Nat. Ecol. Evol.*, 1, 0165.

Mitchell, J.S., Etienne, R.S. & Rabosky, D.L. (2019). Inferring diversification rate variation from phylogenies with fossils. *Syst. Biol.*, 68, 1–18.

Mittelbach, G.G., Schemske, D.W., Cornell, H.V., Allen, A.P., Brown, J.M., Bush, M.B., *et al.* (2007). Evolution and the latitudinal diversity gradient: speciation, extinction and biogeography. *Ecol. Lett.*, 10, 315–331.

Moles, A.T., Ackerly, D.D., Tweddle, J.C., Dickie, J.B., Smith, R., Leishman, M.R., *et al.* (2007). Global patterns in seed size. *Glob. Ecol. Biogeogr.*, 16, 109–116.

Moore, B.R., Höhna, S., May, M.R., Rannala, B. & Huelsenbeck, J.P. (2016). Critically evaluating the theory and performance of Bayesian analysis of macroevolutionary mixtures. *Proc. Natl. Acad. Sci.*, 113, 9569–9574.

Morlon, H., Lewitus, E., Condamine, F.L., Manceau, M., Clavel, J. & Drury, J. (2016). RPANDA: an R package for macroevolutionary analyses on phylogenetic trees. *Methods Ecol. Evol.*, 7, 589–597.

Mounce, R., Smith, P. & Brockington, S. (2017). Ex situ conservation of plant diversity in the world's botanic gardens. *Nat. Plants*, 3, 795.

Nee, S., Holmes, E.C., May, R.M. & Harvey, P.H. (1994). Extinction rates can be estimated from molecular phylogenies. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.*, 344, 77–82.

Pellissier, V., Barnagaud, J.-Y., Kissling, W.D., Şekercioğlu, Ç. & Svenning, J.-C. (2018). Niche packing and expansion account for species richness–productivity relationships in global bird assemblages. *Glob. Ecol. Biogeogr.*, 27, 604–615.

Pennell, M.W., FitzJohn, R.G. & Cornwell, W.K. (2016). A simple approach for maximizing the overlap of phylogenetic and comparative data. *Methods Ecol. Evol.*, 7, 751–758.

bioRxiv preprint doi: https://doi.org/10.1101/619064; this version posted January 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Pianka, E.R. (1966). Latitudinal gradients in species diversity: a review of concepts. *Am. Nat.*, 100, 33–46.

Pinheiro, J., Bates, D., DebRoy, S. & Sarkar, D. (2014). R Core Team (2014) nlme: linear and nonlinear mixed effects models. R package version 3.1-117. *Available https://CRAN.R-project.org/package=nlme.*

Pontarp, M., Bunnefeld, L., Cabral, J.S., Etienne, R.S., Fritz, S.A., Gillespie, R., *et al.* (2019). The Latitudinal Diversity Gradient: Novel Understanding through Mechanistic Eco-evolutionary Models. *Trends Ecol. Evol.*, 34, 211–223.

Pyron, R.A. & Wiens, J.J. (2013). Large-scale phylogenetic analyses reveal the causes of high tropical amphibian diversity. *Proc. R. Soc. Lond. B Biol. Sci.*, 280, 20131622.

Qian, H. & Jin, Y. (2016). An updated megaphylogeny of plants, a tool for generating plant phylogenies and an analysis of phylogenetic community structure. *J. Plant Ecol.*, 9, 233–239.

Quintero, I. & Jetz, W. (2018). Global elevational diversity and diversification of birds. *Nature*, 555, 246–250.

Rabosky, D.L. (2009). Ecological limits and diversification rate: alternative paradigms to explain the variation in species richness among clades and regions. *Ecol. Lett.*, 12, 735–743.

Rabosky, D.L. (2010). Extinction rates should not be estimated from molecular phylogenies. *Evolution*, 64, 1816–1824.

Rabosky, D.L. (2014). Automatic detection of key innovations, rate shifts, and diversity-dependence on phylogenetic trees. *PloS One*, 9, e89543.

Rabosky, D.L. (2016). Challenges in the estimation of extinction from molecular phylogenies: A response to Beaulieu and O'Meara. *Evolution*, 70, 218–228.

Rabosky, D.L., Chang, J., Title, P.O., Cowman, P.F., Sallan, L., Friedman, M., *et al.* (2018). An inverse latitudinal gradient in speciation rate for marine fishes. *Nature*, 559, 392–395.

Rabosky, D.L., Grundler, M., Anderson, C., Title, P., Shi, J.J., Brown, J.W., *et al.* (2014). BAMMtools: an R package for the analysis of evolutionary dynamics on phylogenetic trees. *Methods Ecol. Evol.*, 5, 701–707.

Rabosky, D.L. & Huang, H. (2015). Minimal effects of latitude on present-day speciation rates in New World birds. *Proc R Soc B*, 282, 20142889.

Rabosky, D.L. & Huang, H. (2016). A robust semi-parametric test for detecting trait-dependent diversification. *Syst. Biol.*, 65, 181–193.

Rabosky, D.L., Mitchell, J.S. & Chang, J. (2017). Is BAMM flawed? Theoretical and practical concerns in the analysis of multi-rate diversification models. *Syst. Biol.*, 66,

bioRxiv preprint doi: https://doi.org/10.1101/619064; this version posted January 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

477–498.

Rensch, B. (1959). *Evolution above the species level*. Methuen, London.

Reyes, E., Morlon, H. & Sauquet, H. (2015). Presence in Mediterranean hotspots and floral symmetry affect speciation and extinction rates in Proteaceae. *New Phytol.*, 207, 401–410.

Rohde, K. (1992). Latitudinal gradients in species diversity: the search for the primary cause. *Oikos*, 514–527.

Rolland, J., Condamine, F.L., Jiguet, F. & Morlon, H. (2014). Faster speciation and reduced extinction in the tropics contribute to the mammalian latitudinal diversity gradient. *PLoS Biol.*, 12, e1001775.

Schluter, D. (2016). Speciation, ecological opportunity, and latitude: (American Society of Naturalists address). *Am. Nat.*, 187, 1–18.

Schluter, D. & Pennell, M.W. (2017). Speciation gradients and the distribution of biodiversity. *Nature*, 546, 48.

Simpson, G. (1953). *The major features of evolution*. Columbia University Press, New York, USA.

Smith, S.A. & Brown, J.W. (2018). Constructing a broadly inclusive seed plant phylogeny. *Am. J. Bot.*, 105, 302–314.

Svenning, J.-C., Borchsenius, F., Bjorholm, S. & Balslev, H. (2008). High tropical net diversification drives the New World latitudinal gradient in palm (Arecaceae) species richness. *J. Biogeogr.*, 35, 394–406.

Title, P.O. & Rabosky, D.L. (2019). Tip rates, phylogenies and diversification: What are we estimating, and how good are the estimates? *Methods Ecol. Evol.*, 10, 821–834.

Tung Ho, L. si & Ané, C. (2014). A linear-time algorithm for gaussian and non-gaussian trait evolution models. *Syst. Biol.*, 63, 397–408.

Wallace, A.R. (1878). *Tropical nature and other essays*. Macmillan, London, United Kingdom.

Weir, J.T. & Price, T.D. (2011). Limits to speciation inferred from times to secondary sympatry and ages of hybridizing species along a latitudinal gradient. *Am. Nat.*, 177, 462–469.

Weir, J.T. & Schluter, D. (2007). The latitudinal gradient in recent speciation and extinction rates of birds and mammals. *Science*, 315, 1574–1576.

Wiens, J.J. & Donoghue, M.J. (2004). Historical biogeography, ecology and species richness. *Trends Ecol. Evol.*, 19, 639–644.

bioRxiv preprint doi: https://doi.org/10.1101/619064; this version posted January 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Xing, Y. & Ree, R.H. (2017). Uplift-driven diversification in the Hengduan Mountains, a temperate biodiversity hotspot. *Proc. Natl. Acad. Sci.*, 114, E3444–E3451.

Zanne, A.E., Tank, D.C., Cornwell, W.K., Eastman, J.M., Smith, S.A., FitzJohn, R.G., *et al.* (2014). Three keys to the radiation of angiosperms into freezing environments. *Nature*, 506, 89–92.