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

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 28,000 angiosperm species, we found that the LDG does not arise from variation in speciation rates because lineages speciated faster outside 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 and greater environmental change, may ultimately explain these results.

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

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[1–4], 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[5]. This increase in speciation is generally attributed to higher environmental energy in the tropics, which in turn can hasten evolution[6], such as through shorter generation times, higher mutation rates[7] and foster more complex biotic interactions (e.g., faster coevolution or the 'Red Queen running faster when she is hot'[8]). All these factors make reproductive isolation more likely and favour speciation. However, while some studies on mammals[9], amphibians[10] and butterflies[11] have shown that tropical lineages speciate faster, other studies have found no latitudinal differences[12,13] or higher rates of speciation in temperate areas[14], particularly when focusing on recent rates of speciation[15–17]. Recent speciation rates should be less affected by extinction[18] and more accurate than deep-time estimates, even in the absence of paleontological evidence[19]. They should also reflect any variable that historically influenced speciation and also varied latitudinally, such as environmental energy[20].

Flowering plants are one of the largest eukaryotic radiations (roughly 350,000 species) and show a marked LDG[21]. Despite this pattern, latitudinal differences in

51 macroevolutionary rates in angiosperms have received comparatively little attention at
52 a species-level. Previous efforts have shown that angiosperm diversification is faster
53 in tropical families[22] and in tropical lineages within a limited number of clades[23,24].
54 The only species-level analysis found no differences in speciation, but did so by
55 assuming all rate differences in the angiosperm Tree of Life (ToL) were due to
56 geographic distribution alone[25]. One reason for the lack of attention to angiosperms is
57 that their vast diversity makes it difficult to obtain well-sampled phylogenetic and
58 distributional data at a species-level, which is required to estimate speciation rates
59 accurately.
60
61 **Results and discussion**
62 Here we use species-level data from the whole angiosperm ToL to assess the
63 latitudinal variation of speciation rates in plants. We first estimated species-specific
64 (i.e., tip-based) speciation rates for 29,703 species of flowering plants[26] using
65 Bayesian Analysis of Macroevolutionary Mixtures (BAMM[27]). BAMM allows for
66 rate heterogeneity through time and across lineages and accommodates non-random
67 incomplete taxon sampling. We then obtained latitudinal data for 236,894 angiosperm
68 species from a previous study[28] and classified species as tropical or temperate using
69 their median latitude estimated from the Global Biodiversity Information Database
70 (GBIF, see Methods for details). We collated the BAMM and GBIF datasets to obtain
71 a dataset (herein 'full') of 28,057 species.
72
73 We found that speciation rates consistently increased towards the poles using several
74 ways of delimiting tropical and temperate species. Using correlation tests that control
75 for phylogenetic pseudoreplication [24], we found that tropical species had smaller

mean speciation rates (λ) than temperate species (Fig 1a; $\lambda_{temperate}$ = 1.24 species/myr, $\lambda_{tropical}$ = 0.81 species/myr, p-value = 0.024). Angiosperms in all 10° latitudinal bands outside the tropics (i.e. absolute median latitude >23.5°) 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 = 0.128, p-value = 0.013). These results were maintained when we: *i*) analysed only those species that were densely sampled in GBIF; *ii*) compared strictly tropical and strictly temperate species (i.e., discarding broadly ranging species that occurred both inside and outside the tropics); and *iii*) discarded the highest latitudinal bands (>50° N and S), which contained very few species (Table S1).

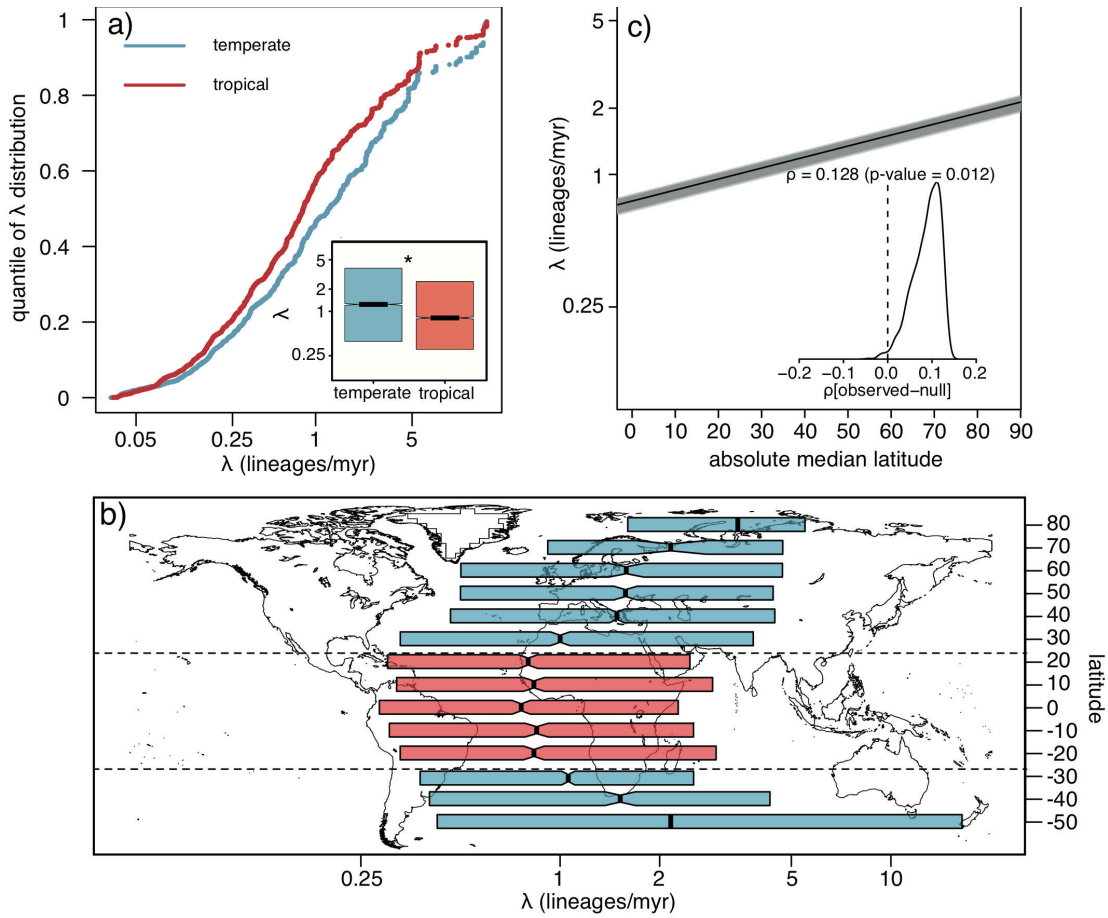


Figure 1. Temperate species have higher speciation rates (λ) than tropical species. a) Rank ordered distribution and boxplot (inset) of λ inferred with BAMM for tropical (red) and temperate (blue) species. * indicates significant difference (p-value = 0.024) between the two groups, assessed with a Mann-Whitney 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. Inset shows the difference between the empirical and null correlations estimated with 1000 permutations of evolutionary rates across the phylogeny.

Our full dataset contained a disproportionate amount of temperate species but this did not explain the consistent association between latitude and speciation rates. The initial GBIF dataset was consistent with the LDG, containing 42.5% and 57.5% temperate and tropical species, respectively (Fig S1a-c). However, the data were disproportionately represented by temperate species (61.7%) after merging GBIF distributions with the macroevolutionary rates (Fig. S1d-f). We assessed if this latitudinal bias could affect our results in two ways. First, we generated 100 random samples of our full dataset (herein 'unbiased' datasets), each with 10,000 species that maintained the proportion of tropical (57.5%) and temperate (42.5%) species present in GBIF. Second, we generated another set of 100 random samples (herein 'extreme tropical' datasets), where the proportions of temperate and tropical species were 29.1% and 70.9% respectively. These were calculated assuming that all species absent from GBIF were tropical and therefore represent the most extreme scenario of

113 latitudinal bias. We then assessed whether the λ estimates were affected by the poorer
114 sampling for tropical species, which may result in average longer branches and lower
115 λ. We reran BAMM for 10 of the 100 unbiased and extreme tropical datasets and
116 compared the resulting estimates with the ones from the full dataset (n = 28,057
117 species). We found that the λ estimates from the full and the subsampled datasets
118 were very positively correlated (median ρ = 0.82, median p-value < 0.001 and median
119 ρ = 0.83, median p-value < 0.001 for the unbiased and the extreme tropical datasets,
120 respectively), and that this association did not differ between tropical and temperate
121 species (Figs S2, S3). We also confirmed that the association of λ and latitude was not
122 a result of the excessive proportion of temperate species in the full dataset by
123 repeating the STRAPP analyses subsetted to the species present in each of the 100
124 unbiased and extreme tropical datasets. Our conclusions were identical between the
125 full and both the unbiased (Fig S4) and extreme tropical datasets (Fig S5). Finally, we
126 confirmed that our results could not be explained by different levels of sampling for
127 BAMM between temperate and tropical species. We found that the relationship of
128 median latitude and BAMM sampling fraction was positive but small (Fig S6).
129 Moreover, densely sampled clades, which occurred at higher latitudes, tended to have
130 smaller estimates of λ (Fig. S7), contrary to what we might expect if our results were
131 a sampling artefact. These analyses collectively exclude the conclusion that our
132 results stem from lower estimates of speciation rates in more sparsely sampled
133 tropical clades.
134
135 We also confirmed that the speciation rate estimates were robust despite the low
136 levels of sampling of the angiosperm ToL. Our full dataset contained 8% of the
137 angiosperm species described in The Plant List[30]. To assess the effect of missing taxa,

138 we randomly sampled 8% of the species (n = 2,273 species) in our full dataset ten
139 times and reran BAMM on these 'small' datasets. As above, the λ estimates from the
140 full and the small datasets were positively correlated (median ρ = 0.66, median p-
141 value < 0.001), and this association did not differ between tropical and temperate
142 species (Fig. S8).
143
144 The positive association between speciation rates and latitude was consistently
145 retrieved using two alternative methods for estimating macroevolutionary rates. First,
146 we used FiSSE (Fast, intuitive State-dependent Speciation-Extinction analysis[31]), a
147 non-parametric test that determines the effect of a binary trait on speciation rates. We
148 found a positive association of speciation rate estimated with FiSSE (DR) and
149 whether species were temperate with our full dataset (median $DR_{temperate}$ = 0.29
150 species/myr, median $DR_{tropical}$ = 0.18 species/myr, p-value = 0.001). To test whether
151 the excess of temperate species influenced this result, we repeated the analyses in the
152 100 unbiased datasets of 10,000 species and found no significant differences
153 ($DR_{temperate}$ = 0.15 species/myr, $DR_{tropical}$ = 0.13 species/myr, p-value = 0.168).
154 However, simulations suggested that this non-significant result may stem from a lack
155 of power with FiSSE for our given phylogenetic and trait data (Fig S9, see Methods).
156 FiSSE, unlike BAMM, also does not account for incomplete sampling in the
157 speciation rate estimates, so we caution against strong interpretations of the results
158 obtained with this method with the present levels of missing taxa. Second, we
159 analysed the association between latitude and speciation rate across clades of different
160 ages in the angiosperm ToL. Following ref. 26, we defined six 4 million-year-wide
161 time slices, from 0 to 24 million years (myr), which were used to select clades where
162 speciation rates could confidently be estimated (i.e., probability of recovering the

163 crown age of lineages was > 70% following ref. 30) and GBIF coverage was > 50%.
164 We then fitted models that implemented time-variable and time-constant functions for
165 speciation and extinction[32], and selected the best fitting model to obtain clade-based
166 estimates of speciation rates. Mirroring the BAMM results, we found that clades with
167 a larger proportion of temperate species had higher rates of speciation (Fig S10).
168
169 Overall, our findings do not support the long-standing notion that the greater species
170 diversity of the tropics can be explained by higher rates of speciation. We instead
171 found higher rates of recent speciation closer to the poles. As temperate biotas have
172 fewer species, their niche space may be less saturated, thereby increasing
173 opportunities for lineage divergence[15,33]. Similarly, reproductive isolation may be
174 elevated at higher latitudes by the greater ecological opportunity generated from
175 recurrent environmental change and climate instability[34]. Other explanations may be
176 related to the ecological and life history traits of temperate species. For example,
177 faster speciation rates at higher latitudes could stem from the higher frequency of
178 small-seeded species[35]. Small seed size is positively correlated with angiosperm
179 diversification[26], and temperate species did have smaller seeds in our full dataset (see
180 Methods, phylANOVA: p value = 0.001, significance assessed with 1,000 random
181 simulations with phytools[36]). Therefore, our results suggest that latitudinal
182 differences in speciation rate do not shape the LDG but are shaped by differences in
183 species diversity and traits[16,37].
184
185 Rejecting the role of speciation rate as the cause of the LDG calls for alternative
186 explanations[25]. From a macroevolutionary perspective, greater tropical species
187 diversity may arise for at least three reasons[38]: *i*) higher rates of diversification (i.e.,

188 speciation minus extinction), *ii*) higher rates of net migration, and *iii*) greater potential
189 for historical colonisation. First, the elevated speciation near the poles could be
190 coupled with higher extinction rates to produce a scenario where temperate zones act
191 as both a "cradle" and "grave" of biodiversity[34]. Net diversification might therefore
192 still be higher in tropical areas despite the absolute rates of speciation and extinction
193 each being smaller than closer to the poles. However, the lack of reliable extinction
194 rate estimates from molecular phylogenies in the absence of fossil data[19,39]
195 complicates a direct test of this hypothesis. Second, the tropical conservatism
196 hypothesis proposes that lower net migration from the tropics to temperate zones
197 contributes to the LDG[40]: most lineages originated in the tropics and have been unable
198 to disperse into harsher environments. Third, if rates of diversification and migration
199 have no latitudinal variation, the greater diversity of the tropics may stem from their
200 larger area through time (age-and-area hypothesis[41]).
201
202 An alternative to these ideas is that the LDG may simply have nothing to do with
203 macroevolution. A whole class of hypotheses posits that there are ecological limits to
204 the number of species that a region can hold, and that these limits vary with latitude[42].
205 For example, the larger diversity in the tropics could be explained by expanded niche
206 spaces and greater niche packing[43].
207
208 Recent findings have suggested that rates of species origination are highest in species-
209 poor areas[17,44]. Our results are consistent with these findings and therefore suggest
210 that latitudinal variation in speciation rates is not a major engine of global diversity
211 patterns in one of the largest eukaryotic radiations. Using time-heterogeneous
212 estimates of speciation, we were able to exclude faster rates of evolution, such as

213 arising from more favourable environmental conditions or more intense biotic
214 interactions, as responsible for shaping the angiosperm LDG. Future work must now
215 resolve whether recent rapid speciation nearer the poles is widespread across the ToL
216 and its underlying causes.
217
218 **Methods**
219
220 *Latitudinal and phylogenetic datasets*
221 We obtained latitudinal data for 236,894 angiosperms that included the median,
222 maximum and minimum latitude for each species[28]. We then intersected these data
223 with speciation rates estimated with BAMM v.2.5.0[27] for 29,703 angiosperms[26]. The
224 input phylogeny was obtained from a time-calibrated phylogenetic tree for 31,389
225 land plant species[45,46]. BAMM detects heterogeneity in speciation and extinction rates
226 through time and across lineages. Non-random incomplete taxon sampling was
227 incorporated into the estimation process by calculating the number of species sampled
228 in each family.
229
230 The species names in the GBIF and BAMM datasets were standardised with The
231 Plant List (TPL) using the *Taxonstand*[47] and *Taxonlookup*[48] packages before collation,
232 resulting in 28,057 species in our full dataset. Tropical and temperate species were
233 defined as those where the absolute median latitude was below and above 23.5°,
234 respectively. Species were also grouped into 10° bins according to their median
235 latitude to define latitudinal bands (starting from a band centered in the Equator from
236 -5° to 5°, then from 5° to 15, etc). We repeated the analyses using only densely
237 sampled species (i.e., with five or more data points in the GBIF dataset). Strictly

238 tropical and temperate species were defined as species with median, maximum and
239 minimum latitude occurring inside and outside the tropics, respectively. We discarded
240 widespread species (i.e., neither strictly tropical nor strictly temperate) for analyses
241 comparing strict tropical and temperate taxa (retaining 20,766 species). Finally, we
242 discarded data with median latitudes in the highest latitudinal bands (>50° and <-50°),
243 which contained very few species with extreme estimates of λ (retaining 25,189
244 species).
245
246 *BAMM-based analyses*
247 We correlated BAMM speciation rates and latitudinal data using Structured Rate
248 Permutations on Phylogenies (STRAPP)[29]. STRAPP assesses the significance of the
249 empirical association of macroevolutionary rates and phenotypic traits by comparing
250 it to a null distribution generated by permuting the speciation rates across the
251 phylogeny while maintaining the position of the rate shifts in the phylogenetic tree.
252 We performed all STRAPP two-tailed tests using the *traitdependentBAMM* function
253 of the BAMMtools package[49] with 1000 replicates and logging the rates. We assessed
254 whether *i*) tropical and temperate species had different speciation rates using a Mann
255 Whitney U test; *ii*) speciation rates for each of the latitudinal bands were different
256 using the Kruskal-Wallis rank sum statistic; and *iii*) species absolute median latitude
257 was correlated with its speciation rate using Spearman's rank correlation coefficient.
258
259 We analysed the potential effect of the disproportionate amount of temperate species
260 in our full dataset using BAMM. Subsampled datasets were consistent with the LDG
261 and either *i*) maintained the proportions of tropical (57.5%) and temperate (42.5%)
262 taxa observed in the GBIF dataset ('unbiased'); or *ii*) assumed that the species absent

263 from the GBIF dataset from ref. 28 were all tropical ('extreme tropical') (n = 109,461
264 species; obtained from subtracting the number of species in The Plant List according
265 to *TaxonLookup*[48] minus the species in the GBIF dataset). Both subsampled datasets
266 were generated to assess the effect of the excess of temperate species in our full
267 dataset, but the extreme tropical dataset represented the most extreme scenario of the
268 effect of undescribed biodiversity in our analyses. The size of these datasets was fixed
269 at 10,000 species so that the tropical species present in each subsample, which were
270 obtained by random draws of 5,750 and 7,091 species for the unbiased and extreme
271 tropical respectively from a total pool of 10,749 species, were not excessively
272 redundant. We selected 10 random replicates of the 10,000-species subsampled
273 datasets and reran the BAMM analyses as detailed in ref. 26. We then assessed
274 whether the relationship of the full (n=28,057 species) and both the unbiased and
275 extreme tropical λ estimates was different for tropical and temperate species. To do
276 this, we predicted the full λ estimates using the subsampled λ estimates, 'tropicality'
277 (i.e. a binary variable indicating if a species was tropical based on our aforementioned
278 definition) and the interaction of the λ estimates and tropicality in each of the 10
279 replicates for the unbiased and extreme tropical datasets. We incorporated the effect
280 of the phylogeny using the *phylolm* package[50]. We found that the effect size of the
281 interaction of tropicality and the λ estimates was not different from zero across all the
282 10 replicates for the unbiased (mean = 0.0009, standard deviation = 0.0015) and the
283 extreme tropical (mean = 0.0004, standard deviation = 0.0016) datasets. This result
284 shows that there is no difference in the correlation between λ of the full and
285 subsampled datasets between temperate and tropical lineages.
286

287 We also confirmed that our results could not be explained by latitudinal variation in
288 the clade sampling fractions of our BAMM analysis. We found a positive relationship
289 (slope = 0.059, p-value < 0.0001) between the species median latitude and its
290 sampling fraction (i.e. the percentage of species in the same family present in our
291 28,057 species phylogenetic tree). This result meant that a 10° increase in median
292 latitude corresponded with an average of 0.59% increase in sampling fraction.
293 Furthermore, we found that there was a negative relationship between sampling
294 fraction and the log-transformed estimated speciation rate (slope = -0.055, p-value <
295 0.0001). Again, this effect size was small. For example, an increase in sampling
296 fraction from 20% to 30% corresponded with an average decrease of 0.33
297 lineages/myr in the estimated speciation rate.
298
299 Our full dataset contained λ estimates for 28,057 species, which represent 8% of
300 described angiosperm species (total n = 346,365 calculated with *Taxonlookup*). To
301 assess the reliability of λ estimates given the large number of missing taxa, we
302 randomly generated 10 subsamples ('small' datasets) with 8% of the species in our
303 full dataset (n = 2,273). As above, we ran BAMM for each of the 10 subsamples and
304 predicted the full λ estimates using the small λ estimates. Similar to the unbiased and
305 extreme tropical datasets, we found that the effect size of the interaction of tropicality
306 and λ was not different from zero across all the 10 replicates (mean = -0.0033,
307 standard deviation = 0.0189).
308
309 Although the reliability of BAMM estimates has recently been questioned[51],
310 simulations have shown that accurate estimates of speciation can be obtained even in
311 the absence of fossil data[39,52]. Tip-based estimates of speciation like the ones we use

312 here should also be robust to the effect of extinction and are increasingly used when
313 analysing the effect of traits on lineage diversification[20]. Extinction estimates are
314 thought to be more unreliable[19,53,54] so we do not discuss them here.
315
316 *FiSSE-based analyses*
317 FiSSE analyses the correlation of a binary trait and diversification and has shown to
318 have a lower type I error than similar trait-dependent diversification methods[31]. It
319 calculates the speciation rate associated with each trait state and is based on the
320 diversification rate metric (DR[55]), which was originally proposed to estimate net
321 diversification but has since been shown to approximate speciation rates better[20]. The
322 distributions of DR for each trait value (i.e. tropical and temperate) had long right-
323 tails and the means were much larger than the medians, so the latter measure better
324 described the central tendency of the DR distributions. Therefore, following ref. 56
325 we modified the FiSSE function provided in ref. 31 to compare the median
326 diversification rate for each trait value instead of the mean.
327
328 We also assessed whether the lack of significant differences between tropical and
329 temperate speciation rates that we found with the 100 unbiased datasets could be
330 explained by their sample size (n=10,000 species). We generated 100 replicates of the
331 full dataset of 2,000; 4,000; 6,000; 8,000; 10,000; and 12,000 species and reran FiSSE
332 for each of them. We then ran an exponential regression with the size of a dataset as a
333 predictor and the logged p-value of the association of speciation and tropicality as a
334 response.
335
336 *Clade-based analyses*

337 We estimated clade-level measures of speciation across the angiosperm phylogenetic
338 tree. We used a set of 4 million-year-wide time intervals from 0 to 24 myr to define
339 the ages of the clades. Clades were delimited as the largest non-overlapping
340 monophyletic groups of four or more species where the sampling was larger than
341 30%. This criterion ensures that the probability of retrieving the correct crown age of
342 the clade is above 70%[30]. At least 50% of the species in the clades also had to have
343 latitudinal data. We used RPANDA (R: Phylogenetic ANalyses of DiversificAtion[32])
344 to fit six different models of diversification: *i*) constant speciation (λ) and extinction
345 (μ) fixed at 0; *ii*) constant λ and μ; *iii*) exponential λ and μ fixed at 0; *iv*) exponential λ
346 and constant μ; *v*) constant λ and exponential μ; and *vi*) exponential λ and μ.
347 Incomplete sampling within clades was incorporated into the estimates by calculating
348 the proportion of sampled species within the genera in each clade. Speciation rates
349 were estimated using model averaging with AIC weights. The correlation of the
350 clade-level estimates of speciation and the proportion of temperate and tropical
351 species in each clade was assessed using phylogenetic least squares regression as
352 implemented in the R package *nlme*[57]. The lambda parameter was optimised using a
353 maximum likelihood approach. We weighted the regressions by the proportion of
354 GBIF data points in each clade to account for the effect of different sampling levels of
355 distribution data. We then confirmed that potential latitudinal differences in sampling
356 proportions could not explain our results. To do that, we analysed if the sampling
357 proportions of largely tropical and largely temperate clades (defined as clades where
358 >50% of the species were tropical and temperate, respectively) across the 6 time
359 intervals were different using Pearson's chi-squared test. We found that temperate
360 clades had significantly denser average sampling only in the 16-20 (sampling

361 proportion $_{tropical}$ = 0.65, sampling proportion $_{temperate}$ = 0.56) and 20-24 (sampling
362 proportion $_{tropical}$ = 0.69, sampling proportion $_{temperate}$ = 0.57) myr intervals.
363
364 *Latitudinal variation of seed size*
365 Smaller seeded species have been shown to speciate faster than large seeded species[26].
366 We assessed whether seed size differed between temperate and tropical lineages as it
367 has previously been shown to do so[35], and so latitudinal variation in seed size could
368 explain the differences in speciation between tropical and temperate species. We
369 obtained seed size measurements for 13,127 species in our full dataset from a
370 previous study[26]. The resulting dataset was disproportionately temperate (76.04% of
371 the species were temperate). Average seed sizes were larger in tropical than in
372 temperate species (mean seed size$_{tropical}$ = 0.149 g; mean seed size$_{temperate}$ = 0.065 g; t =
373 -298.7; df = 4189.2; p-value < 0.001). This difference remained when phylogeny was
374 considered (phylANOVA: t = 35.293; p value = 0.001; significance assessed with
375 1,000 random simulations).
376
377 **Data availability**
378 Scripts to reproduce the analyses will be deposited upon acceptance in figshare
379
380
381 **Author contributions**
382 J.I. conceived the study and performed the analysis; J.I. and A.J.T interpreted the
383 analysis and wrote the manuscript.
384
385

Acknowledgments

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. The authors declare no competing interests.

Declaration of interests

The authors declare no competing interests.

Supplementary Figures and Tables

dataset name	$\lambda_{temperate}$	$\lambda_{tropical}$	p-value	dataset size
more.GBIF.data	1.242	0.802	0.014	26,568
no.widespread	1.337	0.802	0.013	20,766
no.high.latitude	1.207	0.813	0.030	25,819

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.

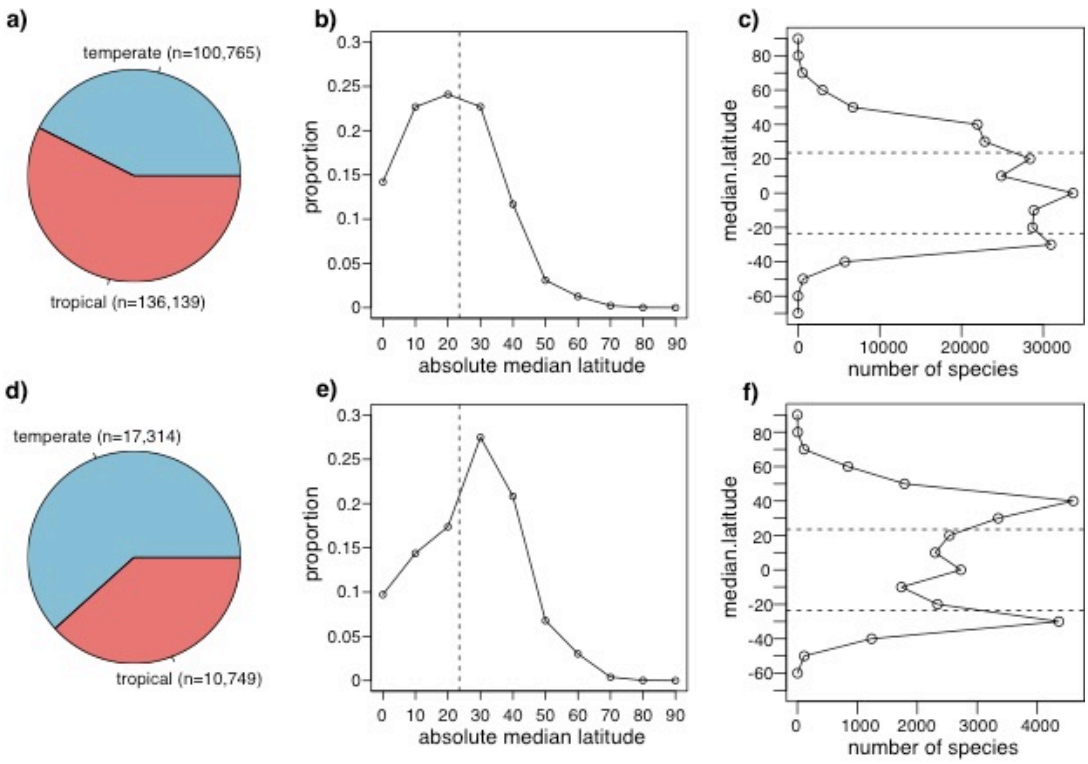


Figure S1. The full dataset had a disproportionate number of temperate species. **a)** Proportion of tropical and temperate species in the GBIF dataset (n = 236,894); **b)** proportion of species in each latitudinal band in the GBIF dataset; **c)** number of species in each latitudinal band in the GBIF dataset; **d)** proportion of tropical and temperate species in the full dataset (n = 28,057); **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.

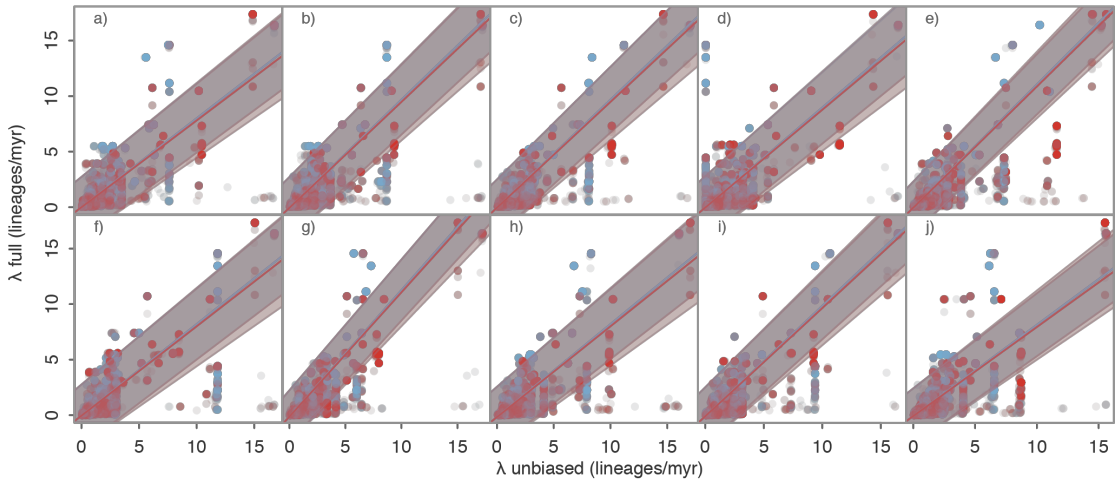


Figure S2. BAMM λ estimates were positively correlated in the full and 10 unbiased datasets with no effect of tropicality on this correlation. **a) - j)** Solid lines are phylogenetic linear regressions predicting λ in the full tree (n = 28,057 species) with λ in the unbiased tree (n = 10,000 species) in temperate (shown in blue) and tropical (shown in red) species. Shaded areas indicate the 95% confidence intervals.

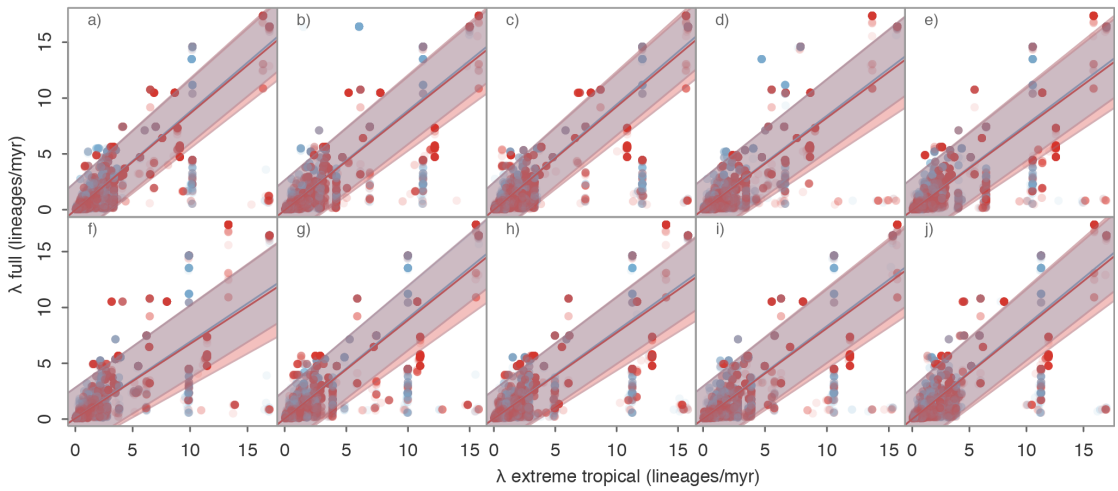


Figure S3. BAMM λ estimates were positively correlated in the full and 10 extreme tropical datasets with no effect of tropicality on this correlation. **a) - j)** Solid lines are phylogenetic linear regressions predicting λ in the full tree (n = 28,057 species) with λ

in the extreme tropical tree (n = 10,000 species) in temperate (shown in blue) and tropical (shown in red) species. Shaded areas indicate the 95% confidence intervals.

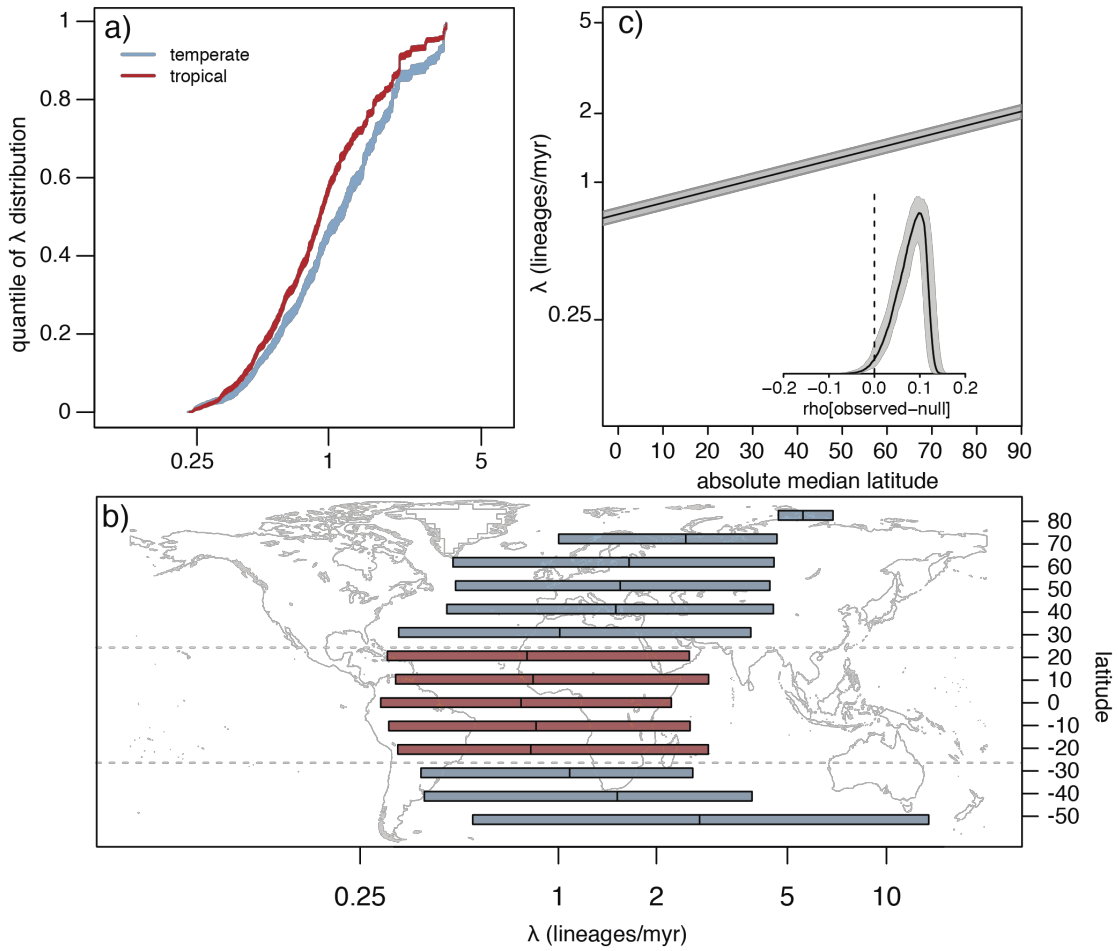


Figure S4. Temperate species have higher speciation rates (λ) than tropical species for 100 subsampled "unbiased" datasets with 42.5% of temperate and 57.5% of tropical species. **a)** Rank ordered distribution of λ inferred with BAMM for tropical (red) and temperate (blue) species. **b)** λ grouped by latitudinal band of each species. Boxes 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. 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.

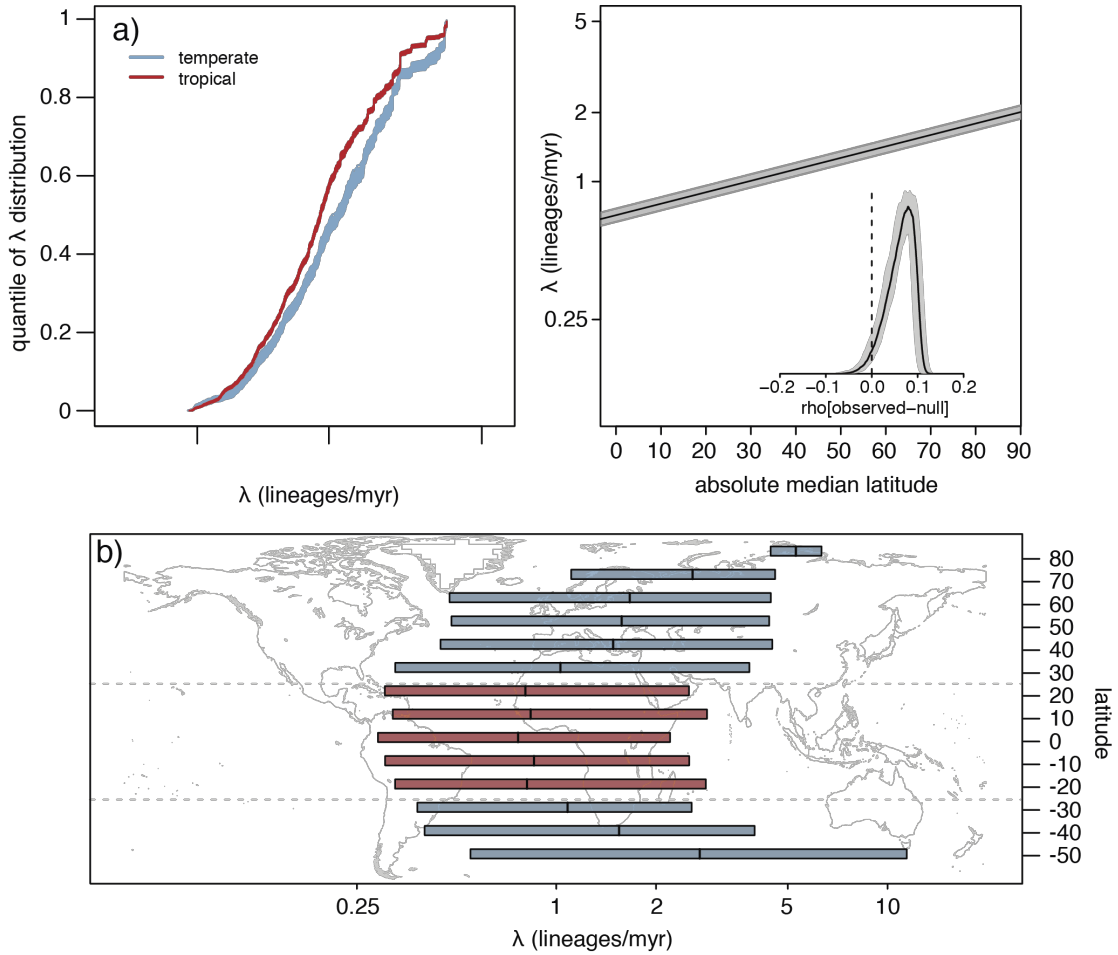


Figure S5. Temperate species have higher speciation rates (λ) than tropical species for 100 subsampled "extreme tropical" datasets with 29.1% temperate and 70.9% tropical species. **a)** Rank ordered distribution of λ inferred with BAMM for tropical (red) and temperate (blue) species. **b)** λ grouped 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. Lines and inset shown as in Fig. S4

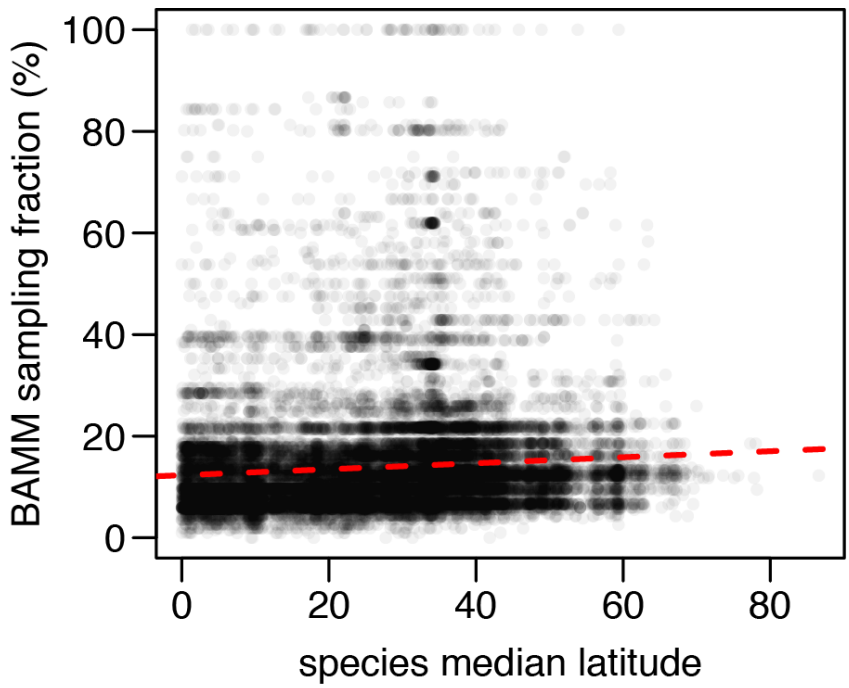


Figure S6. Relationship between the species absolute median latitude and the family-level sampling fraction used in the BAMM analyses. The red line is the slope of the linear regression (slope = 0.059, p-value < 0.0001).

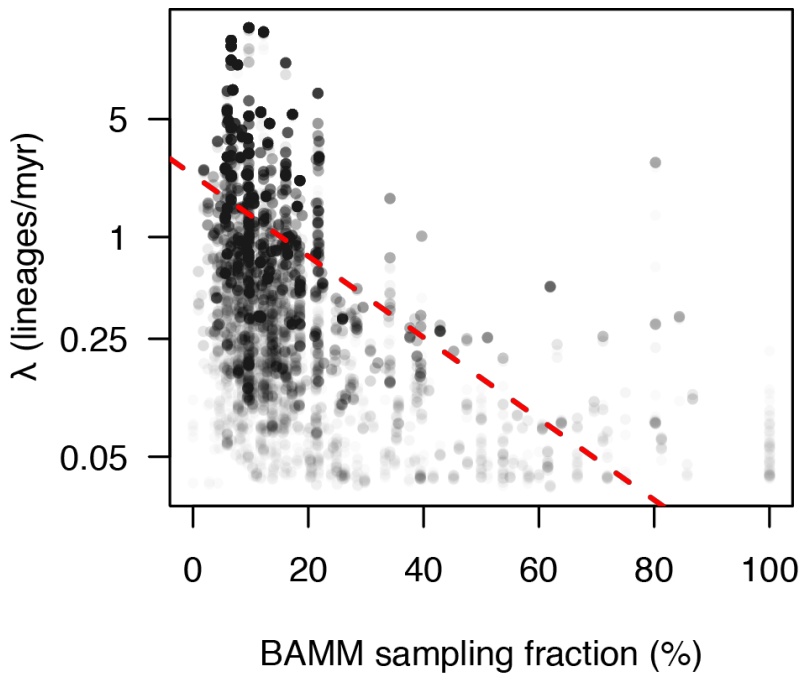


Figure S7. More densely sampled clades have smaller estimates of speciation rate (λ). The dotted line is the slope of the linear regression of the log(λ) and the BAMM sampling fraction (slope = -0.055 , p-value < 0.0001).

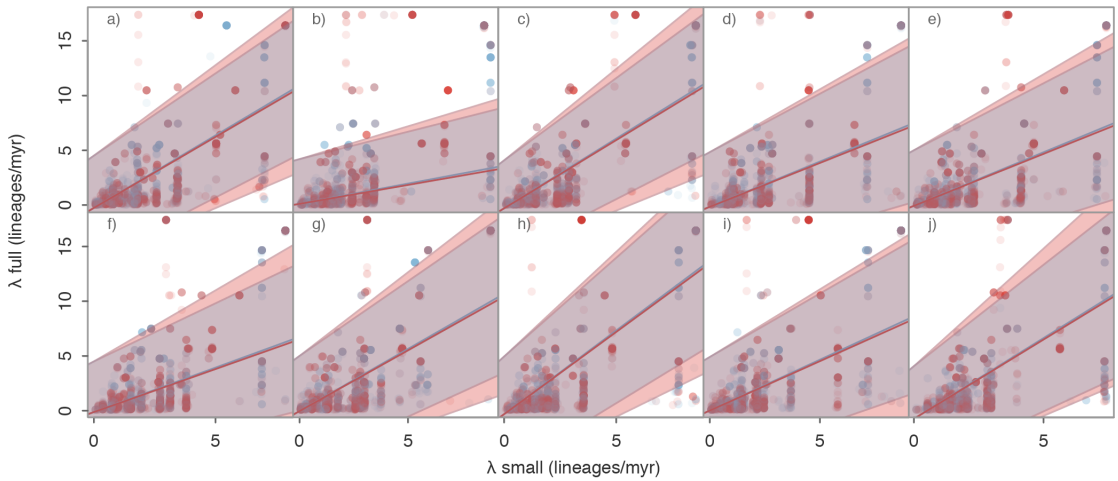
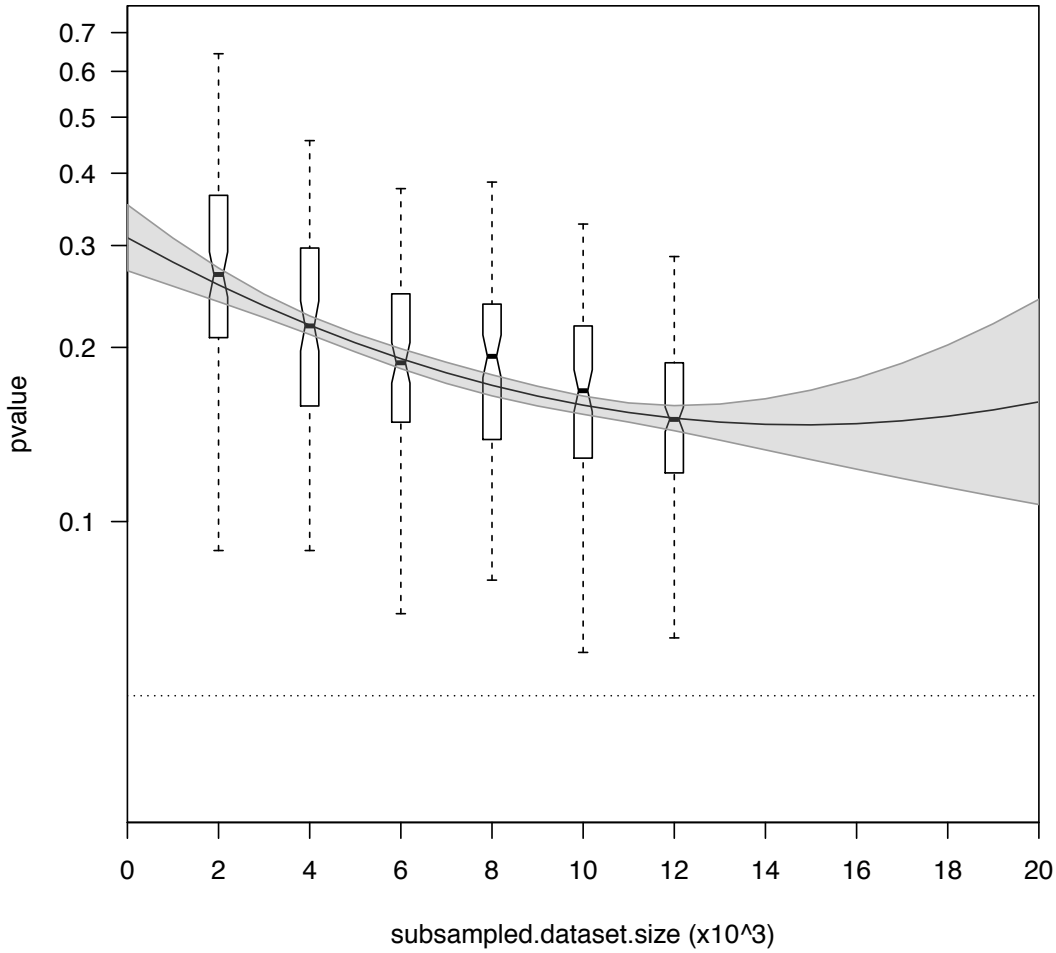


Figure S8. BAMM λ estimates are positively correlated in the full and 10 small datasets and tropicality has no effect on this correlation. **a) - j)** Solid lines are phylogenetic linear regressions predicting λ in the full tree (n = 28,057 species) with λ in the small tree (n = 2,273 species) in temperate (shown in blue) and tropical (shown in red) species. Shaded areas indicate the 95% confidence intervals.



483

484 **Figure S9.** FiSSE power analysis. Each boxplot shows the p-values for the difference
485 of $DR_{temperate}$ and $DR_{tropical}$ for 100 subsamples of varying size taken from the full
486 dataset (n = 28,057 species) whilst retaining proportions of tropical and temperate
487 species observed in the GBIF dataset. Notches in boxplots indicate 95% confidence
488 intervals around the median, denoted by thick vertical lines. Boxes indicate the
489 interquartile range (IQR) and whiskers extend to the most extreme data points that are
490 not further away than ±1.5×IQR. The solid line is an exponential model predicting the
491 logged mean p-value ± 95% confidence interval (shaded area) from the size of the
492 dataset. The dotted line indicates p-value = 0.05.

493

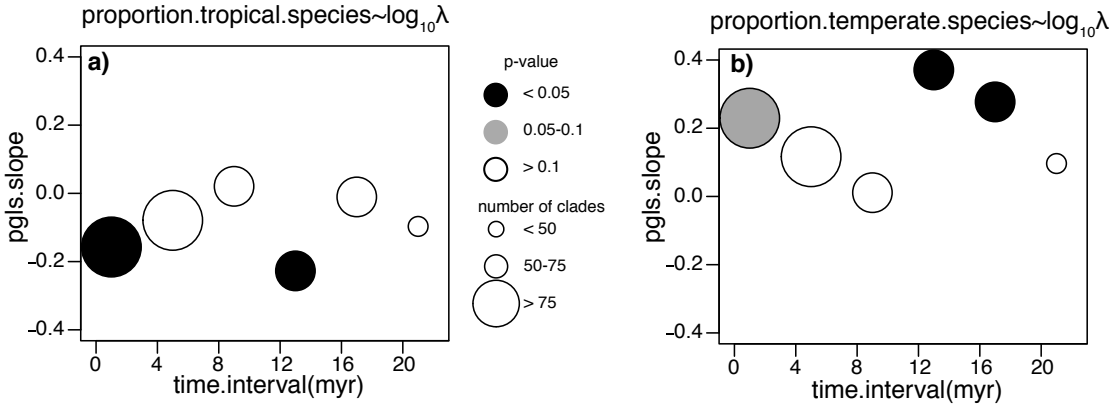


Figure S10. Temperate clades have higher speciation rates in the clade-based analysis. Correlation of a) the proportion of tropical and b) temperate species in each clade with the corresponding speciation rate (λ) estimated with RPANDA. The correlation coefficient is the phylogenetic generalised least 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.

514 **References**

515

516 1. von Humboldt, A. *Ansichten der Natur mit wissenschaftlichen*
517 *Erlauterungen*. (J.G. Cotta, 1808).
518 2. Wallace, A. R. *Tropical nature and other essays*. (Macmillan, 1878).
519 3. Pianka, E. R. Latitudinal gradients in species diversity: a review of
520 concepts. *Am. Nat.* **100**, 33–46 (1966).
521 4. Hawkins, B. A. Ecology's oldest pattern? *Trends Ecol. Evol.* **16**, 470
522 (2001).
523 5. Mittelbach, G. G. *et al.* Evolution and the latitudinal diversity gradient:
524 speciation, extinction and biogeography. *Ecol. Lett.* **10**, 315–331 (2007).
525 6. Dowle, E. J., Morgan-Richards, M. & Trewick, S. A. Molecular evolution
526 and the latitudinal biodiversity gradient. *Heredity* **110**, 501–510 (2013).
527 7. Rohde, K. Latitudinal gradients in species diversity: the search for the
528 primary cause. *Oikos* 514–527 (1992).
529 8. Brown, J. H. Why are there so many species in the tropics? *J. Biogeogr.*
530 **41**, 8–22 (2014).
531 9. Rolland, J., Condamine, F. L., Jiguet, F. & Morlon, H. Faster speciation
532 and reduced extinction in the tropics contribute to the mammalian
533 latitudinal diversity gradient. *PLoS Biol.* **12**, e1001775 (2014).
534 10. Pyron, R. A. & Wiens, J. J. Large-scale phylogenetic analyses reveal the
535 causes of high tropical amphibian diversity. *Proc. R. Soc. Lond. B Biol.*
536 *Sci.* **280**, 20131622 (2013).
537 11. Condamine, F. L., Sperling, F. A., Wahlberg, N., Rasplus, J.-Y. & Kergoat,
538 G. J. What causes latitudinal gradients in species diversity? Evolutionary

539 processes and ecological constraints on swallowtail biodiversity. *Ecol.*
540 *Lett.* **15**, 267–277 (2012).
541 12. Rabosky, D. L. & Huang, H. Minimal effects of latitude on present-day
542 speciation rates in New World birds. *Proc R Soc B* **282**, 20142889 (2015).
543 13. Economo, E. P., Narula, N., Friedman, N. R., Weiser, M. D. & Guénard, B.
544 Macroecology and macroevolution of the latitudinal diversity gradient in
545 ants. *Nat. Commun.* **9**, 1778 (2018).
546 14. Weir, J. T. & Schluter, D. The latitudinal gradient in recent speciation and
547 extinction rates of birds and mammals. *Science* **315**, 1574–1576 (2007).
548 15. Schluter, D. Speciation, ecological opportunity, and latitude: (American
549 Society of Naturalists address). *Am. Nat.* **187**, 1–18 (2016).
550 16. Schluter, D. & Pennell, M. W. Speciation gradients and the distribution of
551 biodiversity. *Nature* **546**, 48 (2017).
552 17. Rabosky, D. L. *et al.* An inverse latitudinal gradient in speciation rate for
553 marine fishes. *Nature* **559**, 392–395 (2018).
554 18. Nee, S., Holmes, E. C., May, R. M. & Harvey, P. H. Extinction rates can be
555 estimated from molecular phylogenies. *Philos. Trans. R. Soc. Lond. B.*
556 *Biol. Sci.* **344**, 77–82 (1994).
557 19. Marshall, C. R. Five palaeobiological laws needed to understand the
558 evolution of the living biota. *Nat. Ecol. Evol.* **1**, 0165 (2017).
559 20. Title, P. O. & Rabosky, D. L. Tip rates, phylogenies and diversification:
560 What are we estimating, and how good are the estimates? *Methods Ecol.*
561 *Evol.* **0**,
562 21. Francis, A. P. & Currie, D. J. A globally consistent richness-climate
563 relationship for angiosperms. *Am. Nat.* **161**, 523–536 (2003).

564 22. Jansson, R. & Davies, T. J. Global variation in diversification rates of
565 flowering plants: energy vs. climate change. *Ecol. Lett.* **11**, 173–183
566 (2008).
567 23. Svenning, J.-C., Borchsenius, F., Bjorholm, S. & Balslev, H. High tropical
568 net diversification drives the New World latitudinal gradient in palm
569 (Arecaceae) species richness. *J. Biogeogr.* **35**, 394–406 (2008).
570 24. Jansson, R., Rodríguez-Castañeda, G. & Harding, L. E. What can multiple
571 phylogenies say about the latitudinal diversity gradient? A new look at the
572 tropical conservatism, out of the tropics, and diversification rate
573 hypotheses. *Evolution* **67**, 1741–1755 (2013).
574 25. Antonelli, A. *et al.* An engine for global plant diversity: highest evolutionary
575 turnover and emigration in the American tropics. *Front. Genet.* **6**, (2015).
576 26. Igea, J., Miller, E. F., Papadopulos, A. S. & Tanentzap, A. J. Seed size
577 and its rate of evolution correlate with species diversification across
578 angiosperms. *PLoS Biol.* **15**, e2002792 (2017).
579 27. Rabosky, D. L. Automatic detection of key innovations, rate shifts, and
580 diversity-dependence on phylogenetic trees. *PloS One* **9**, e89543 (2014).
581 28. Mounce, R., Smith, P. & Brockington, S. Ex situ conservation of plant
582 diversity in the world's botanic gardens. *Nat. Plants* **3**, 795 (2017).
583 29. Rabosky, D. L. & Huang, H. A robust semi-parametric test for detecting
584 trait-dependent diversification. *Syst. Biol.* **65**, 181–193 (2016).
585 30. Rabosky, D. L. Challenges in the estimation of extinction from molecular
586 phylogenies: A response to Beaulieu and O'Meara. *Evolution* **70**, 218–228
587 (2016).

588 31. Rabosky, D. L. & Goldberg, E. E. FiSSE: A simple nonparametric test for
589 the effects of a binary character on lineage diversification rates. *Evol. Int.*
590 *J. Org. Evol.* **71**, 1432–1442 (2017).
591 32. Morlon, H. *et al.* RPANDA: an R package for macroevolutionary analyses
592 on phylogenetic trees. *Methods Ecol. Evol.* **7**, 589–597 (2016).
593 33. Simpson, G. *The major features of evolution*. (Columbia University Press,
594 1953).
595 34. Cutter, A. D. & Gray, J. C. Ephemeral ecological speciation and the
596 latitudinal biodiversity gradient. *Evolution* **70**, 2171–2185 (2016).
597 35. Moles, A. T. *et al.* Global patterns in seed size. *Glob. Ecol. Biogeogr.* **16**,
598 109–116 (2007).
599 36. Revell, L. J. phytools: An R package for phylogenetic comparative biology
600 (and other things). *Methods Ecol. Evol.* **3**, 217–223 (2012).
601 37. Weir, J. T. & Price, T. D. Limits to speciation inferred from times to
602 secondary sympatry and ages of hybridizing species along a latitudinal
603 gradient. *Am. Nat.* **177**, 462–469 (2011).
604 38. Igea, J. & Tanentzap, A. J. Multiple macroevolutionary routes to becoming
605 a biodiversity hotspot. *Sci. Adv.* **5**, eaau8067 (2019).
606 39. Mitchell, J. S., Etienne, R. S. & Rabosky, D. L. Inferring diversification rate
607 variation from phylogenies with fossils. *Syst. Biol.* syy035 (2018).
608 40. Wiens, J. J. & Donoghue, M. J. Historical biogeography, ecology and
609 species richness. *Trends Ecol. Evol.* **19**, 639–644 (2004).
610 41. Fine, P. V. A. & Ree, R. H. Evidence for a time-Integrated species-area
611 effect on the latitudinal gradient in tree diversity. *Am. Nat.* **168**, 796–804
612 (2006).

42. Rabosky, D. L. Ecological limits and diversification rate: alternative paradigms to explain the variation in species richness among clades and regions. *Ecol. Lett.* **12**, 735–743 (2009).
43. Pellissier, V., Barnagaud, J.-Y., Kissling, W. D., Şekercioğlu, Ç. & Svenning, J.-C. Niche packing and expansion account for species richness–productivity relationships in global bird assemblages. *Glob. Ecol. Biogeogr.* **27**, 604–615 (2018).
44. Quintero, I. & Jetz, W. Global elevational diversity and diversification of birds. *Nature* **555**, 246–250 (2018).
45. Zanne, A. E. *et al.* Three keys to the radiation of angiosperms into freezing environments. **506**, 89–92 (2014).
46. Qian, H. & Jin, Y. An updated megaphylogeny of plants, a tool for generating plant phylogenies and an analysis of phylogenetic community structure. *J. Plant Ecol.* **9**, 233–239 (2016).
47. Cayuela, L., Cerda, Í. G. la, Albuquerque, F. S. & Golicher, D. J. taxonstand: An r package for species names standardisation in vegetation databases. *Methods Ecol. Evol.* **3**, 1078–1083 (2012).
48. Pennell, M. W., FitzJohn, R. G. & Cornwell, W. K. A simple approach for maximizing the overlap of phylogenetic and comparative data. *Methods Ecol. Evol.* **7**, 751–758 (2016).
49. Rabosky, D. L. *et al.* BAMMtools: an R package for the analysis of evolutionary dynamics on phylogenetic trees. *Methods Ecol. Evol.* **5**, 701–707 (2014).
50. Tung Ho, L. si & Ané, C. A linear-time algorithm for gaussian and non-gaussian trait evolution models. *Syst. Biol.* **63**, 397–408 (2014).

51. Moore, B. R., Höhna, S., May, M. R., Rannala, B. & Huelsenbeck, J. P. Critically evaluating the theory and performance of Bayesian analysis of macroevolutionary mixtures. *Proc. Natl. Acad. Sci.* **113**, 9569–9574 (2016).
52. Rabosky, D. L., Mitchell, J. S. & Chang, J. Is BAMM flawed? Theoretical and practical concerns in the analysis of multi-rate diversification models. *Syst. Biol.* **66**, 477–498 (2017).
53. Rabosky, D. L. Extinction rates should not be estimated from molecular phylogenies. *Evolution* **64**, 1816–1824 (2010).
54. Mitchell, J. S., Etienne, R. S. & Rabosky, D. L. Inferring diversification rate variation from phylogenies with fossils. *Syst. Biol.* **68**, 1–18 (2019).
55. Jetz, W., Thomas, G. H., Joy, J. B., Hartmann, K. & Mooers, A. O. The global diversity of birds in space and time. *Nature* **491**, 444 (2012).
56. Faurby, S. & Antonelli, A. Evolutionary and ecological success is decoupled in mammals. *J. Biogeogr.* **45**, 2227–2237 (2018).
57. Pinheiro, J., Bates, D., DebRoy, S. & Sarkar, D. R Core Team (2014) nlme: linear and nonlinear mixed effects models. R package version 3.1-117. *available at https://CRAN.R-project.org/package=nlme.* (2014).