






# Recent accelerated diversification in rosids occurred outside the tropics

Miao Sun <sup>1,2,3,8</sup>✉, Ryan A. Folk <sup>4,8</sup>✉, Matthew A. Gitzendanner<sup>5,6</sup>, Pamela S. Soltis <sup>1,6,7</sup>, Zhiduan Chen<sup>3</sup>, Douglas E. Soltis <sup>1,5,6,7</sup> & Robert P. Guralnick <sup>1,6</sup>✉

Conflicting relationships have been found between diversification rate and temperature across disparate clades of life. Here, we use a supermatrix comprising nearly 20,000 species of rosids—a clade of ~25% of all angiosperm species—to understand global patterns of diversification and its climatic association. Our approach incorporates historical global temperature, assessment of species' temperature niche, and two broad-scale characterizations of tropical versus non-tropical niche occupancy. We find the diversification rates of most subclades dramatically increased over the last 15 million years (Myr) during cooling associated with global expansion of temperate habitats. Climatic niche is negatively associated with diversification rates, with tropical rosids forming older communities and experiencing speciation rates ~2-fold below rosids in cooler climates. Our results suggest long-term cooling had a disproportionate effect on non-tropical diversification rates, leading to dynamic young communities outside of the tropics, while relative stability in tropical climates led to older, slower-evolving but still species-rich communities.

<sup>1</sup>Florida Museum of Natural History, University of Florida, Gainesville, FL 32611, USA. <sup>2</sup>Department of Bioscience, Aarhus University, Aarhus 8000C, Denmark. <sup>3</sup>State Key Laboratory of Systematic and Evolutionary Botany, Institute of Botany, The Chinese Academy of Sciences, Beijing 100093, China. <sup>4</sup>Department of Biological Sciences, Mississippi State University, Starkville, MS 39762, USA. <sup>5</sup>Department of Biology, University of Florida, Gainesville, FL 32611, USA. <sup>6</sup>Biodiversity Institute, University of Florida, Gainesville, FL 32611, USA. <sup>7</sup>Genetics Institute, University of Florida, Gainesville, FL 32608, USA. <sup>8</sup>These authors contributed equally: Miao Sun, Ryan A. Folk. ✉email: [cactusresponsible@gmail.com](mailto:cactusresponsible@gmail.com); [rfolk@biology.msstate.edu](mailto:rfolk@biology.msstate.edu); [robgur@gmail.com](mailto:robgur@gmail.com)

Attempts to understand the diversification of the angiosperms date to the time of Darwin, yet the drivers of their spectacular species diversity remain to be adequately assessed across much of angiosperm phylogeny as reviewed in Soltis et al.<sup>1</sup>, as well as Sauquet and Magallón<sup>2</sup>. A combination of environmental niche reconstructions and estimates of diversification rates provides one key approach to assess abiotic drivers of angiosperm diversity. Assembly and analysis of such broad-scale data remain challenging<sup>3</sup>. Nevertheless, the identification of climatic drivers of diversification, such as temperature, has seen a surge of recent interest. Numerous studies have demonstrated a correlation between organismal diversification rates and past climate change, particularly with global temperature trends after the Cretaceous–Paleogene (K–Pg) boundary (~65 Million years ago [Myr] to present)<sup>4–8</sup>. The nature of this relationship has been highly variable, with some studies recovering a negative correlation (e.g., accelerated evolution during cool periods<sup>9–11</sup>) and others a positive correlation (rapid evolution during warm periods<sup>6,12</sup>). Likewise, studies of latitudinal gradients in diversification rate have found patterns congruent with both of these relationships (i.e., higher diversification at higher latitudes<sup>13,14</sup>, or higher diversification at lower latitudes<sup>15,16</sup>) or, in some cases, with neither relationship<sup>10,17</sup>. There remains little consensus on the origins of these contrasting biodiversity patterns, with numerous hypotheses proposed across different taxa using different methodological approaches: e.g., the diversification rate hypothesis, the ecological regulation hypothesis, and the tropical conservatism hypothesis<sup>18–20</sup>.

Elucidating the timing and drivers of diversification of the rosid clade of angiosperms is key to better understanding the origin of present-day angiosperm diversity and global vegetation<sup>21</sup>. Rosids, an enormous clade of 90,000–120,000 species (estimated from the Open Tree of Life [OpenTree] and Open Tree Taxonomy [OTT] database<sup>22</sup>), represent more than a quarter of all angiosperms, based on an estimated 400,000 species of angiosperms<sup>23</sup>. The clade originated in the Early to Late Cretaceous (115–93 Myr), followed by rapid diversification of crown groups (sensu APG IV<sup>24</sup>) of fabids (112 to 91 Myr), and malvids (109 to 83 Myr<sup>25,26</sup>) in perhaps as little as 4–5 million years<sup>25</sup>.

Present-day angiosperm-dominated forests are the outcomes of the rapid initial rise of the rosid clade and its subsequent repeated cycles of radiations<sup>25,27</sup>. These forests also profoundly shaped many current terrestrial ecosystems and much of biodiversity<sup>25,28</sup>. Many other lineages of life radiated in the shadow of rosid-dominated angiosperm forests (e.g., ants<sup>29</sup>, amphibians<sup>30</sup>, mammals<sup>31</sup>, liverworts<sup>32</sup>, and most ferns<sup>33</sup>). Rosids are also highly diverse ecologically (e.g., parasite: Rafflesiaceae; mangrove: Rhizophoraceae; aquatic: Podostemaceae) and now occupy most terrestrial habitats (e.g., arctic, alpine, desert, temperate, and tropical forests) with correspondingly large trait diversity (e.g., chemistry, reproductive strategy, and life history<sup>25,34,35</sup>).

Rosids have their highest species richness in the tropics, where they are ecologically dominant<sup>36</sup>, but most rosid subclades extend into subtropical and temperate areas and climate zones<sup>25</sup>, and some lineages, including Brassicaceae, Rosaceae, and Fagales, greatly diversified in temperate habitats<sup>11,37,38</sup>. Despite recent efforts to develop comprehensive rosid phylogenies<sup>25,39–41</sup>, no large-scale assessment of rosid diversification has yet been conducted, likely due to the formidable size of this clade. However, extensive phylogenetic sampling in such large clades is valuable for providing a broad context and clade-level replication for testing macroevolutionary hypotheses<sup>10,36,42</sup>. The enormous diversity of rosids therefore affords the opportunity for evolutionary replication across multiple globally distributed clades,

whereas latitudinal diversity disparities enable testing how climate, and temperature in particular, may have driven diversification.

Here we use a recently compiled 5-locus phylogenetic tree for rosids<sup>41</sup>, with more than twice the taxon sampling used in earlier studies<sup>39,40</sup>, to examine how past climates have driven diversification. Using the 17 rosid orders (clades classified sensu APG IV<sup>24</sup>) as replicated evolutionary tests, we asked whether latitudinal and climatic disparities in rosid diversity are due to (1) differences in diversification rates of contemporary species and communities within and outside of the tropics; (2) historical temperature-dependence of the diversification process; or (3) differences in species ages in tropical versus non-tropical communities. In particular, we consider both geographic and climatic definitions of tropicality (sensu Owens et al.<sup>43</sup>) and use emerging theory linking climatic niche conservatism to increased diversification rates in habitats outside the tropics<sup>19</sup>. We find that rates of speciation are about twofold higher in non-tropical rosid communities, which are also typically younger. This surprising finding likely relates to marked increases in diversification rates of most rosid subclades during cooling associated with the global expansion of temperate habitats in the Miocene. Our approach not only addresses key, long-held diversity gradient hypotheses, but also considers multiple lines of evidence toward a more process-oriented view of species and diversification gradients.

## Results

**Species record accumulation.** We assembled occurrences for all name-validated rosid species via queries using R packages *rgbif* v1.3.0<sup>44</sup> and *ridigbio* v0.3.5<sup>45</sup>, as discussed in the Methods and Supplementary Method 1. Of the 19,740 sampled rosid species, 18,283 species initially had occurrence data in iDigBio and GBIF; we removed species lacking geographic coordinates, geographic outlier records (Methods and Supplementary Method 1), and coordinates where both latitude and longitude were estimated only to the nearest integer degree. After these cleaning steps, 2,691,099 records remained, and each species had an average of 160 occurrence points. We then paired the species with data from the three contemporary niche data sets: mean annual temperature (MAT; 16,986 species passed our filters), geographic tropicality (18,269 species passed), and Köppen–Geiger climatic tropicality (17,635 species passed). See Supplementary Table 1 for detailed summary of each data layer for the 17 rosid orders.

**Diversification and current temperature niche.** Although we fit a variety of pure-birth and birth-death models, we conservatively chose to focus on reporting speciation parameters, given the challenge of estimating extinction rates from extant-only phylogenetic data and that extinction is probably systematically underestimated<sup>46–48</sup>. Hereafter, without qualification we refer to diversification rate (that is, net diversification estimated from birth-death models including both speciation and extinction rates) or to speciation rates estimated from birth-death models, but for trait correlation tests we used present speciation rates (hereafter, tip rates) estimated from DR (diversification rate<sup>17</sup>) and BMM (Bayesian analysis of macroevolutionary mixtures<sup>49</sup>, and STRAPP<sup>50</sup>; see Methods and Supplementary Method 1).

We tested for a correlation between tip rates and MAT using *es-SIM*<sup>51</sup> and STRAPP<sup>50</sup>. We found a significant relationship between current temperature niche and tip rates across all rosids using *es-SIM* (two-tailed  $p = 0.00067$ , test statistic described in Harvey & Rabosky<sup>51</sup>, Table 1, Fig. 1), although the correlation between tip rates and MAT was not significant using STRAPP (Supplementary Table 3). The direction of this relationship was negative (that is, lower temperatures were associated with higher

**Table 1 Summary table of semi-parametric tests for correlation between diversification rate and continuous mean annual temperature (es-SIM).**

Order	es-SIM		
	Rho ( $\rho$ )	p-value	pv.adjust
Brassicales	-0.3945	0.0180	0.2699
Celastrales	0.0185	0.9444	0.9444
Crossosomatales	-0.1255	0.8011	0.9444
Cucurbitales	-0.0471	0.9290	0.9444
Fabales	-0.3290	0.0060	0.0960
Fagales	-0.0767	0.4985	0.9444
Geraniales	0.0655	0.8117	0.9444
Huerteales	-0.6484	0.2413	0.9444
Malpighiales	-0.3256	0.0007	<b>0.0113</b>
Malvales	-0.1251	0.4392	0.9444
Myrtales	-0.2129	0.1113	0.9444
Oxalidales	-0.1205	0.5252	0.9444
Picramniales	-0.3169	0.7711	0.9444
Rosales	-0.2072	0.2186	0.9444
Sapindales	-0.0668	0.5012	0.9444
Vitales	-0.3762	0.1619	0.9444
Zygophyllales	-0.0581	0.7804	0.9444
Total tree	-0.2919	<b>0.0007</b>	/

The significant p-values for adjusted results are in bold. Rho ( $\rho$ ) is the Pearson's correlation coefficient; p-value is for the correlation between mean annual temperature and speciation rate; pv.adjust refers to adjusting the corresponding p-value for family-wise error using the Hochberg<sup>57</sup> method (which can result in identical adjusted p-values as implemented in R). Rosid orders are provided, sensu APG IV<sup>24</sup>.

tip rates and vice versa;  $\rho = -0.29186$ , Table 1). Among the 17 rosid orders (sensu APG IV<sup>24</sup>; Table 1), only Malpighiales showed a significant relationship and only using es-SIM ( $p = 0.01133$ ), but there was a strong overall trend towards a negative correlation ( $\rho$ ) between tip rates and temperature ( $t$ -test of zero mean difference; two-tailed  $p = 0.00039$ ); Celastrales and Geraniales were the only two clades with a positive correlation between climate and tip rates (Table 1). We also tested for phylogenetic conservatism of temperature niche with Pagel's  $\lambda$  test<sup>52</sup> (here we use phylogenetic signal as the criterion of phylogenetic niche conservatism<sup>53</sup>; see Supplementary Method 2). We found very strong evidence for phylogenetic signal in MAT (Likelihood ratio test:  $p \sim 0$  for Pagel's  $\lambda$  test = 0.86209 across the entire rosid tree; Supplementary Table 2).

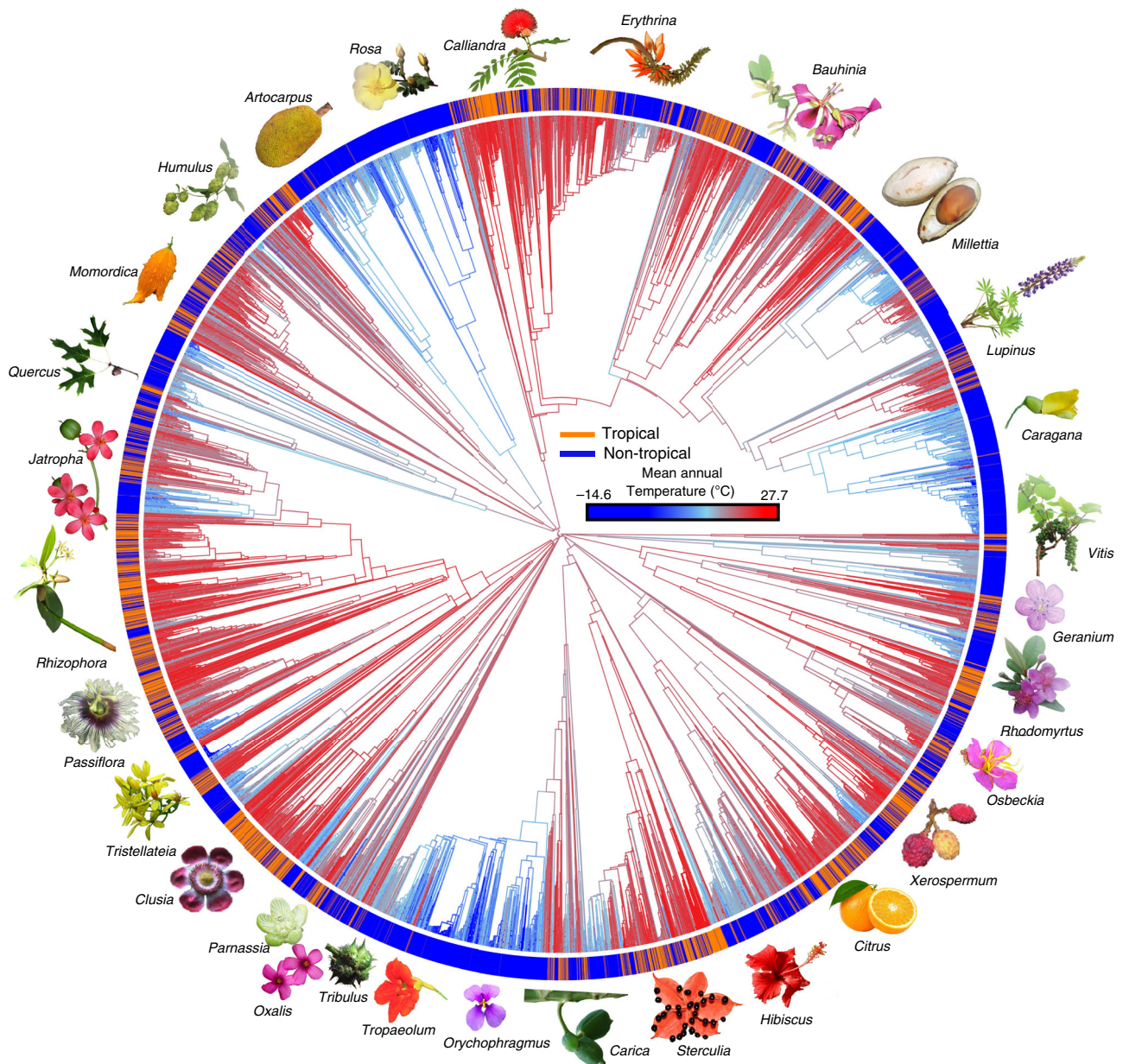
**Diversification and measures of tropicity.** Climatic and geographic definitions resulted in strongly differing delimitations of tropicity. Although we found approximately equal numbers of non-tropical and tropical species with a geographic definition (51.45% and 48.55%, respectively; Supplementary Table 1), climatic criteria resulted in an increased number of non-tropical species (65.07% non-tropical; 34.93% tropical; Supplementary Table 1). The apparent disparities between geographic and climatic tropicity criteria are in subtropical and tropical-montane regions<sup>43</sup>, likely driving the scoring difference. Under the geographic definition, only 27% of the species had at least one point occurrence in both non-tropical and tropical areas. Most species that span both areas have almost all points in one of them, and the majority of species (73%) are found solely in non-tropical or tropical areas. We discuss temperate-tropical biases in more detail below. Under either definition, there was again strong evidence for phylogenetic conservatism of temperature niche using Pagel's  $\lambda$  test (geographic:  $\lambda = 0.93815$ , Likelihood ratio test one-tailed,  $p = 2.61e-306$ ; climatic:  $\lambda = 0.95630$ , Likelihood ratio test one-tailed,  $p = 1.25e-244$ ; Supplementary Table 2).

We tested for a relationship between climatic tropicity and tip rates with two methods: one using semi-parametric tip rates (FiSSE<sup>54</sup>) and the other using tip rates estimated in BAMM (i.e. STRAPP<sup>50</sup>). We found a significant relationship (FiSSE: two-tailed  $p \sim 0$ , test statistic described in Rabosky & Goldberg<sup>54</sup>; Table 2); the average tip rate of species outside the tropics was approximately twofold that of tropical species ( $\lambda_{\text{non-tropical}} = 0.5047$ ;  $\lambda_{\text{tropical}} = 0.2552$ , as estimated in FiSSE; Table 2). Among the 17 orders, climatic tropicity was significantly associated with tip rates for Fabales (in both tests, negatively:  $\lambda_{\text{non-tropical}} > \lambda_{\text{tropical}}$ ): FiSSE ( $p = 0.02997$ ; Table 2) and STRAPP (two-tailed  $p \sim 0$ , test statistic defined in Rabosky & Huang<sup>50</sup>; Supplementary Table 3). Estimated tip rates for non-tropical species ( $\lambda_{\text{non-tropical}}$  as estimated by FiSSE) were consistently and significantly different from those of tropical species ( $\lambda_{\text{tropical}}$ ) across orders (Table 2), and this difference was significant ( $t$ -test for mean difference: two-tailed  $p = 0.0110$ ). However, there was not a significant relationship between the percentage of non-tropical species in an ordinal clade and the disparity between non-tropical and tropical tip rates (linear model of non-tropical proportion vs.  $\lambda_{\text{non-tropical}}/\lambda_{\text{tropical}}$ ;  $F$ -test: two-tailed  $p = 0.8012$ ), indicating that this relationship existed regardless of overall clade temperature niche.

A geographically non-tropical distribution was also significantly associated with different tip rates across the rosids (FiSSE: two-tailed  $p \sim 0$ ; Table 2); geographically non-tropical species had ~1.8-fold the tip rate of tropical species ( $\lambda_{\text{non-tropical}} = 0.5552$ ;  $\lambda_{\text{tropical}} = 0.3089$ ; Table 2). However, this association was weaker than for climatic tropicity, with no order showing a significant association of geographic tropicity with diversification and no significant difference across rosid orders between average tip rates ( $\lambda_{\text{non-tropical}}$  and  $\lambda_{\text{tropical}}$ ;  $t$ -test: two-tailed  $p = 0.0545$ ; Table 2). Similarly, as with climatic tropicity, there was no relationship between the percentage of non-tropical species in the order and the magnitude of the difference between non-tropical and tropical tip rates (linear model of non-tropical proportion vs.  $\lambda_{\text{non-tropical}}/\lambda_{\text{tropical}}$ ;  $F$ -test: two-tailed  $p = 0.4151$ ).

We also assessed the relationship between tropicity and diversification under a fully model-based framework while accounting for the possibility of an unobserved trait driving diversification using a series of HiSSE (hidden-state speciation and extinction<sup>55</sup>) and BiSSE (binary state speciation and extinction<sup>56</sup>) models (Supplementary Method 3). We implemented these tests under a model comparison framework using AIC (Akaike information criterion<sup>57</sup>). Under both definitions of tropicity (climatic and geographic), we find full support with AW (Akaike weight<sup>58</sup>;  $AW \sim 1$ ) for a full HiSSE model, with equal transition rates and relative extinction among tropicity states, but different diversification rates between these two states (Supplementary Table 4). That is, tropicity is significantly associated with diversification, but in concert with an unobserved trait that also drives diversification. Non-tropical diversification is higher under one of these "hidden" traits as expected, but under the other hidden trait, it is slightly (Köppen-Geiger tropicity) or substantially (geographic tropicity) lower. This indicates that, although the overall diversification rate is higher for non-tropical species, the direction of this effect is modulated by an unobserved trait.

Sampling across orders was idiosyncratic with respect to tropicity. For instance, the most prevalently tropical large order, Malpighiales (52.8% climatically tropical, 17.37% species coverage in the tree; cf. Table 1 in Sun et al.<sup>41</sup>), was far better sampled than Rosales, one of the most prominent temperate radiations in the rosids (13.3% tropical; 8.22% species coverage). However, our species richness plots suggest overrepresentation of Rosales in temperate North America and Europe



**Fig. 1 Ancestral reconstruction across the rosid tree for mean annual temperature.** Branches are colored based on mean annual temperature from low ( $-14.6^{\circ}\text{C}$ ; blue) to high ( $27.7^{\circ}\text{C}$ ; red) values. Species tips are colored by the Köppen-Geiger climatic classification<sup>43</sup>; orange represents tropical species, and blue represents non-tropical species, based on climatic criteria; we also represent rosid diversity using icons (flower, fruit, leaves) of clade members to conceptually reflect their morphological character diversity, tropicality, and economic importance; a genus name is provided with each associated icon.

(Supplementary Fig. 1). To test how much both phylogenetic and occurrence record sampling biases might impact predictions of tip rates at the ordinal level, we performed sensitivity analyses (Supplementary Method 4). Based on a series of random taxon dropping treatments representing different disparities of sampling effort among tropical and non-tropical species, we find that overall estimated tip rates are largely insensitive to these sampling differences (Supplementary Method 4 and Supplementary Table 5). Consistent with this finding, although sampling of ordinal trees was indeed moderately negatively correlated with the percent of species that were tropical ( $\rho = -0.2029$ ), this trend was only marginally significant ( $F$ -test: two-tailed  $p = 0.0397$ ). Likewise, the percentage of each order sampled showed no relationship with the disparity between tropical and non-tropical tip rates (i.e.,  $\lambda_{\text{non-tropical}} / \lambda_{\text{tropical}}$ ;  $\rho = -0.0442$ ;  $F$ -test: two-tailed  $p = 0.5779$ ).

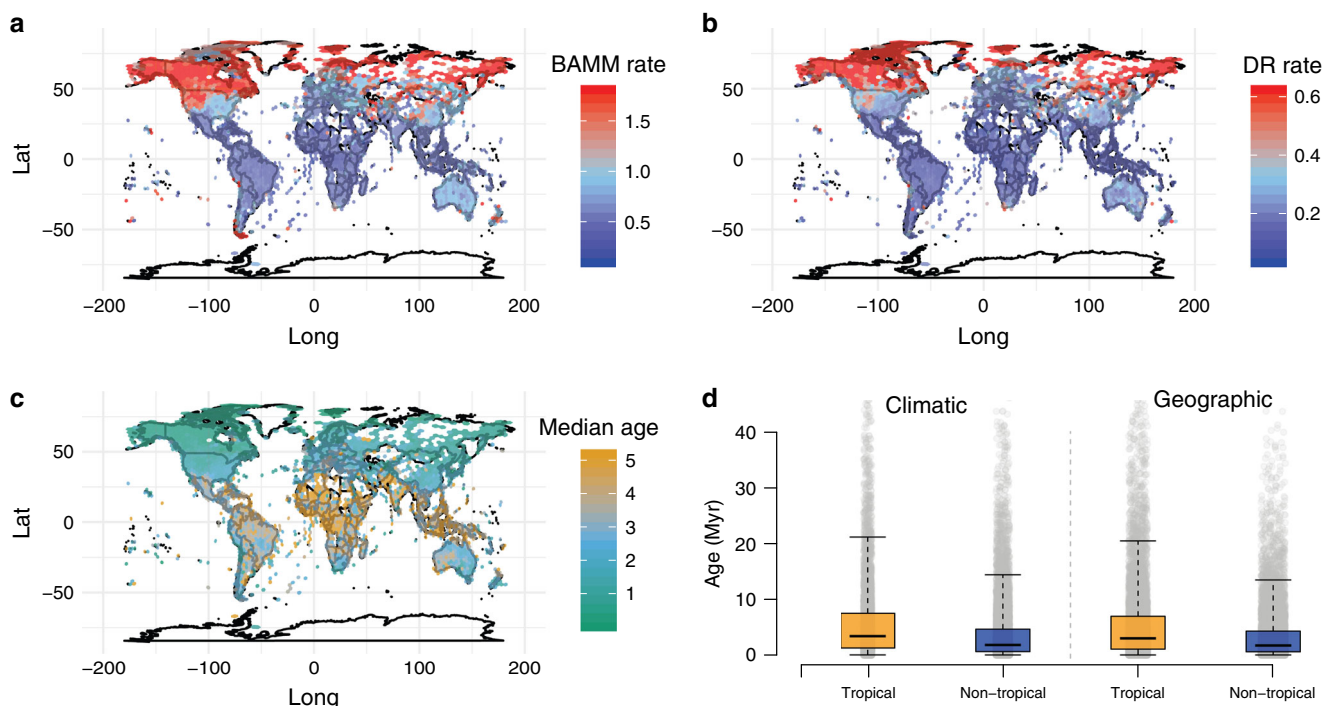
These findings, combined with recovery of largely congruent correlations across clades with different levels of tropicality (above), suggest that the negative diversification-temperature relationship we recovered is robust globally across the phylogeny and independent of sampling effort.

By mapping median tip rates from BAMM and DR analyses on equal-area grids (Fig. 2a, b), we reveal a strong latitudinal pattern in diversification rate specific to the Northern Hemisphere. As with tropicality state-specific estimates, northern non-tropical communities have approximately twofold the median tip rate as communities in tropical areas. Consistent with this pattern, we also show that tropical communities contain substantially older rosid species than communities outside the tropics under both climatic and geographic definitions (Fig. 2c, d;  $t$ -test: both two-tailed; climatic  $p = 3.5518\text{e}-68$ , and geographic  $p = 2.3176\text{e}-82$ ).

**Table 2 Summary of correlation tests between tip rates and the two binary tropicality data sets in FiSSE.**

Order	Köppen–Geiger				Geographic			
	$\lambda_{\text{non-tropical}}$	$\lambda_{\text{tropical}}$	<i>p</i> -value	<i>pv.adjust</i>	$\lambda_{\text{non-tropical}}$	$\lambda_{\text{tropical}}$	<i>p</i> -value	<i>pv.adjust</i>
Brassicales	0.5198	0.1594	0.0985	0.5554	0.5079	0.4410	0.4865	0.7682
Celastrales	0.2208	0.2015	NA	NA	0.2309	0.2046	NA	NA
Crossosomatales	0.0776	0.1073	0.6104	0.6104	0.0893	0.0477	0.2248	0.7682
Cucurbitales	0.3033	0.2353	0.2168	0.5554	0.2919	0.2843	0.4945	0.7682
Fabales	0.6197	0.2946	0.0020	<b>0.0300</b>	0.6962	0.3450	0.0080	0.0959
Fagales	0.6032	0.3385	0.0320	0.4156	0.6276	0.4008	0.0609	0.6703
Geraniales	0.2919	0.0544	0.1578	0.5554	0.2988	0.2366	0.2378	0.7682
Huerteales	0.0278	0.0195	0.1718	0.5554	0.0278	0.0195	0.1768	0.7682
Malpighiales	0.4105	0.2312	0.0230	0.3217	0.4747	0.2582	0.0060	0.0779
Malvales	0.4066	0.2963	0.1898	0.5554	0.4305	0.3639	0.3666	0.7682
Myrtales	0.4766	0.2594	0.0440	0.4835	0.4972	0.3544	0.3437	0.7682
Oxalidales	0.1532	0.0854	0.1698	0.5554	0.1815	0.1503	NA	NA
Picramniales	0.0875	0.0488	0.1389	0.5554	NA	NA	NA	NA
Rosales	0.6407	0.3218	0.0390	0.4675	0.6795	0.4339	0.1149	0.7682
Sapindales	0.2576	0.2202	0.1159	0.5554	0.2815	0.2280	0.1089	0.7682
Vitales	0.7599	0.1055	NA	NA	0.9223	0.1014	NA	NA
Zygophyllales	0.0628	0.0537	0.2777	0.5554	0.0570	0.0663	0.7682	0.7682
Total tree	0.5047	0.2552	<b>0</b>	/	0.5552	0.3089	<b>0</b>	/

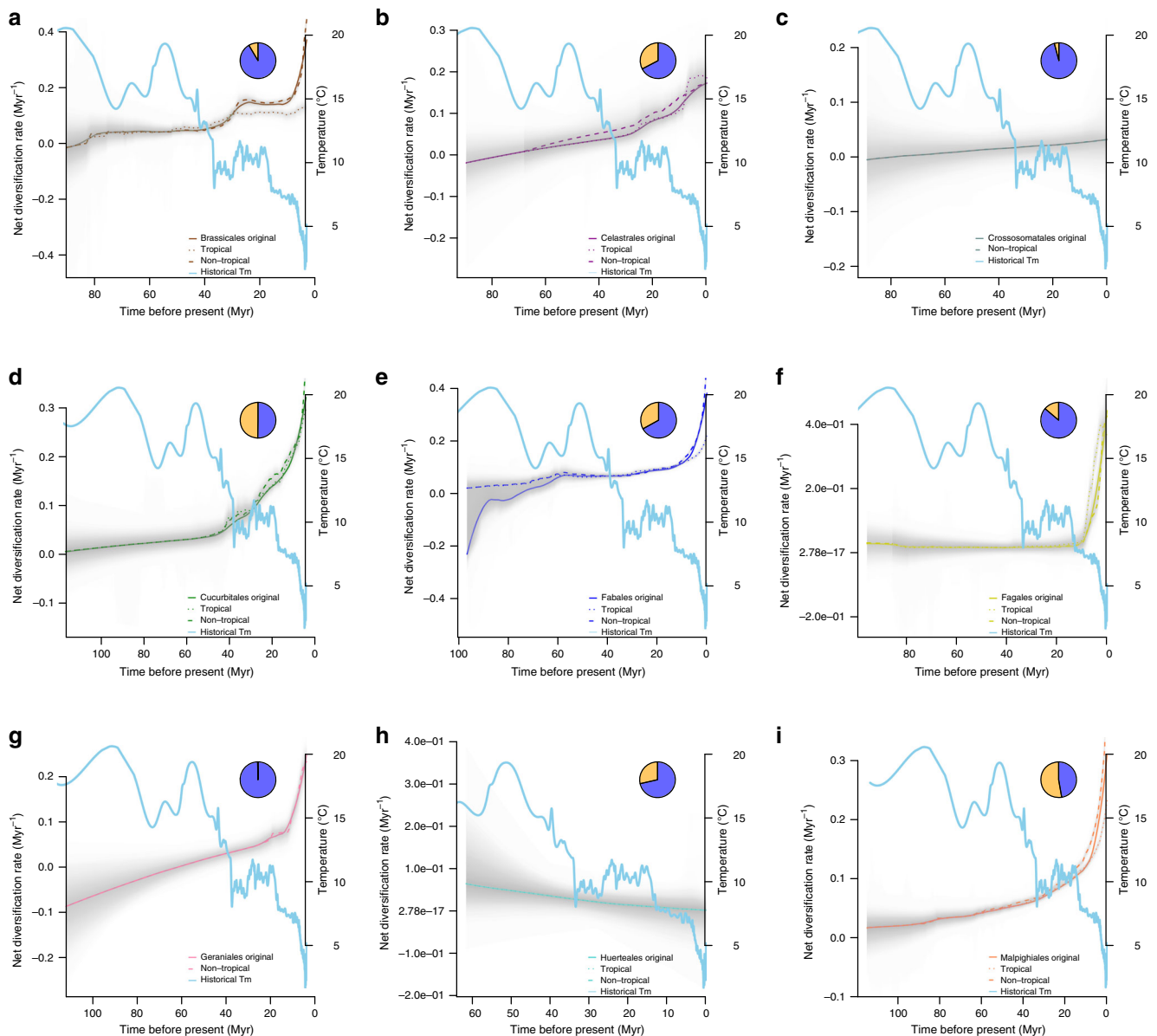
The significant *p*-values for adjusted results are in bold; *p*-value is for the correlation between the Köppen–Geiger climatic/Geographic tropicality and tip rate; *pv.adjust* refers to adjusting the corresponding *p*-value for family-wise error using the Hochberg<sup>37</sup> method. For the binary data, computation of some *p*-values in small subclades failed due to monomorphic traits or mismatched transition rates, which are marked as “NA”. Rosid orders are provided, sensu APG IV<sup>24</sup>.



**Fig. 2 Global distribution patterns of median tip rates and median species age.** Tip rates estimated by BAMM (a) and DR (b) (red = faster, and blue = slower); species tip age (c) is defined as the age of the parent node for any given tip and reported as the grid cell median (orange = older; green = younger); d boxplot of median species age for tropical (orange) and non-tropical (blue) species under both climatic and geographic definitions. t-test: both two-tailed; climatic  $p = 3.5518e-68$  and geographic  $p = 2.3176e-82$ ; median—solid line in the box, box—interquartile range [25% and 75%], whiskers—5% and 95% intervals;  $n = 6160, 11,475, 8870,$  and  $9399,$  respectively; see gray dots for data distribution. All values in a–c are plotted in equal-area grid cells ( $\sim 322 \text{ km}^2$ ); source data are provided in Source Data file.

**Diversification and global historical temperature.** We used the Sun et al.<sup>41</sup> rosid tree with global paleo-temperature data<sup>59</sup> and RPANDA<sup>60</sup> to test alternative diversification models, with more details discussed in Methods below and in Supplementary

Method 5. Model selection demonstrated that temperature-dependent models were favored without exception (Supplementary Table 6). In all 17 orders, the optimal model had constant speciation and extinction with respect to temperature. This result



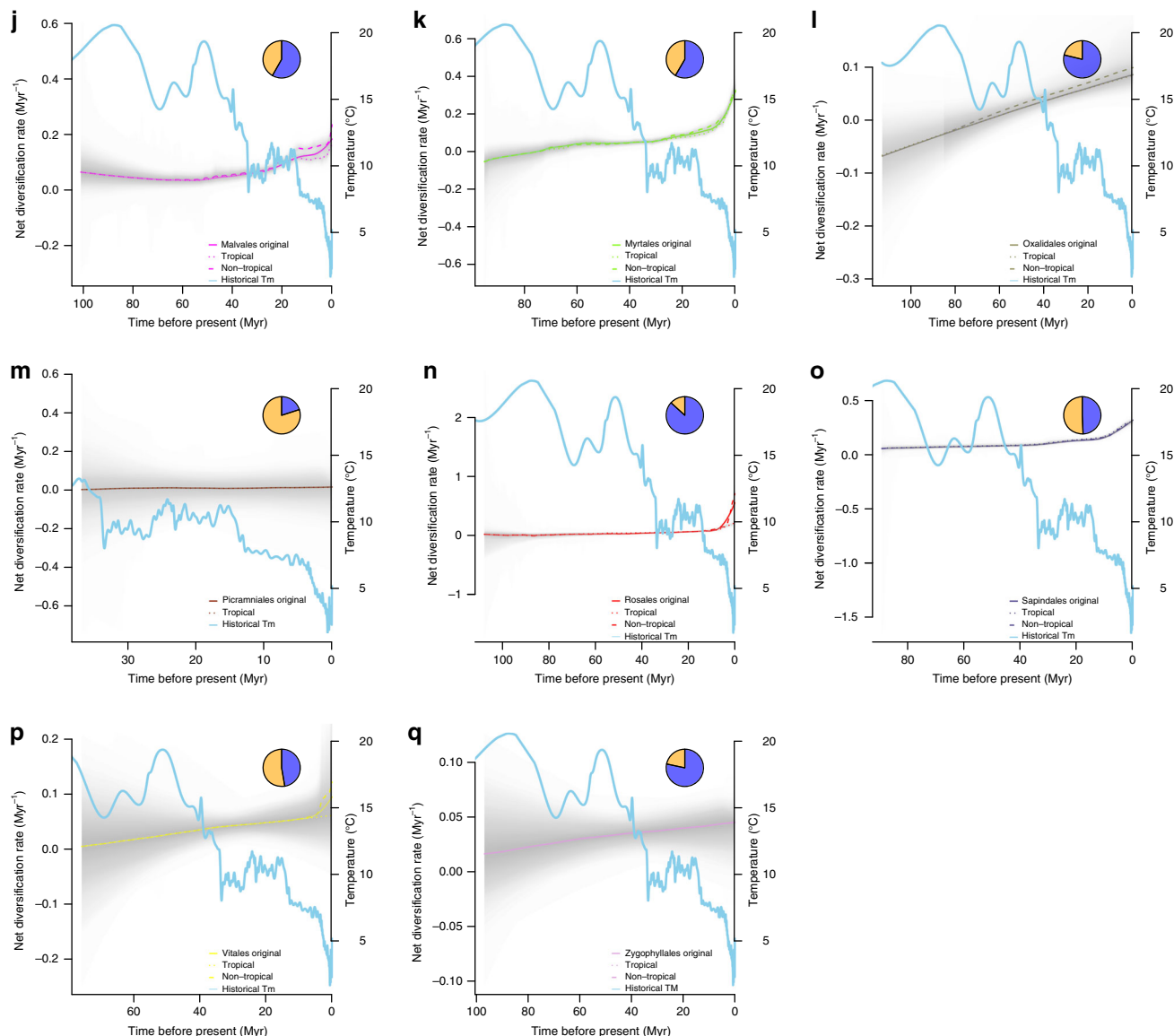
**Fig. 3** Continued

indicates a proportional relationship of diversification with temperature; the coefficients of the models indicated the expected negative relationship. The AW for this model was  $\sim 1$  with the exception of the two smallest orders (Huerteales, 30 species with 28.57% tropical species, AW = 0.768; Picramniales, 57 species with 100% tropical species, AW = 0.812; also see Supplementary Tables 1 and 6).

We also found strong evidence for a negative correlation between estimated global paleo-temperature data and tree-wide net diversification rates (here, not tip rates) using BAMM<sup>49</sup>. In most orders, the association between temperature and net diversification rate was negative and linear (model choice based on AIC against an exponential model; median  $R^2 = 0.712$ ; Supplementary Fig. 2). The exception was Huerteales, the smallest rosids order (30 species), which exhibited a positive correlation (Supplementary Fig. 2h). Across subclasses, more than half of the diversification shifts (97/182 in total, summed over all 17 orders) were detected from 10 to 15 Myr to present, during which the global temperature was cooling, whereas <5 shifts were detected at the Eocene Thermal Maximum ( $\sim 55$  Myr) when the global

temperature was more than 5–8 °C above average. Moreover, among most rosids orders, rate accelerations occurred primarily from  $\sim 10$  to 15 Myr to the present, corresponding to the end of the mid-Miocene Climatic Optimum (Fig. 3; Supplementary Fig. 3). Such increased rates were also evident in some clades (Brassicales, Celastrales, Cucurbitales, and Sapindales; Fig. 3a, b, d, o) closer to an earlier cooling period at the Eocene-Oligocene Glacial Maximum ( $\sim 34$  Myr). Five orders showed no signs of recent bursts of diversification and no shifts in diversification rate detected by BAMM (Crossosomatales, Huerteales, Oxalidales, Picramniales, and Zygophyllales; Fig. 3c, h, l, m, q). This finding may represent a statistical power issue, as four of these five orders are the smallest orders of rosids (Supplementary Table 1).

The response of individual rosids orders to global historical climate change was idiosyncratic, but in general most clades experienced substantial diversification rate shifts and increases from  $\sim 10$  to 15 Myr to the present (Fig. 3; Supplementary Fig. 3). For instance, Fagales maintained a relatively constant net diversification rate close to zero from its origination ( $\sim 95.6$  Myr) to  $\sim 10$  Myr and then exhibited a marked increase to the



**Fig. 3 Trends in net rosid diversification rate-through-time plots of tropical and non-tropical lineages against global temperature.** Net diversification rate curves were generated using BAMM for each of the 17 rosid orders (clades) (a–q); each curve is a median line colored by order; the solid line is the original net diversification rate curve, the dotted line represents the rate of tropical lineages, and the dashed line represents the rate of non-tropical lineages (no curves displayed for any lineage with sample number < 3). The gray segment is the 95% confidence interval; the sky blue line is the trend of global temperature change since the Late Cretaceous (~113 Myr), plotted as a 5-point mean derived from Cramer et al.<sup>59</sup>; the pie-chart represents the percentage of tropical (orange) and non-tropical (blue) species within each order summarized from the Köppen–Geiger climatic classification<sup>43</sup>. Each timescale (x axis) is rescaled to the clade age.

present (Fig. 3f). The diversification rate curve of Rosales was relatively flat after 107 Myr; then there was a rate drop at ~90 Myr, followed by an abrupt rate increase from ~15 Myr (around the Middle Miocene Climatic Optimum) to the present (Fig. 3n). Only a few very small orders exhibited an approximately constant (Picramniales; all tropical species; Fig. 3m) or decreasing (Huerteales; Fig. 3h) trend in diversification towards the present.

### Discussion

Species richness of rosid communities is higher in tropical regions than elsewhere on the planet<sup>36</sup> (Supplementary Fig. 1), and yet we found that tip rates are about twofold higher in non-tropical rosid species (Fig. 2a, b; Table 2). This pattern is only marginally weaker under geographic definitions of tropicality. Individual

rosid orders sometimes showed much greater disparities, such as Vitales (7.2-fold higher for climatically non-tropical species; 9.1-fold higher for geographically non-tropical species). Part of the explanation for this startling pattern may be attributed to our finding of a strong negative relationship between diversification (or tip rates) and both historical and present-day temperature. This association is present both globally across all rosids and in rosid subclades, and is robust to several analytical methods, sensitivity tests, and several data sets representing temperature–species associations. Similar patterns of increased diversification with global cooling have been found in mammals<sup>9</sup>, birds<sup>7</sup>, and some other flowering plant lineages<sup>10,11,14,38</sup>. Both falling global temperature and aridification are compatible with the timing of diversification reconstructed here for rosids and most rosid subclades (~10–15 Myr to the present), a timeframe that also corresponds to the origin of many modern-day temperate plant

communities<sup>61</sup>. Consistent with this observation, we found support for hidden-state models of tropicality (Supplementary Table 4), indicating a role for unobserved drivers of diversification that may work in concert with temperature and should be the focus of future study.

Despite concordant overall patterns in the diversification–climate relationship across most rosid orders, rate-through-time plots demonstrate a diversity of responses to global temperature (Fig. 3). These differences among subclades may reflect phylogenetic constraints imposed by ancestral temperature niche and other factors<sup>62</sup>. Nevertheless, temperature is negatively associated with diversification not only in largely temperate clades such as Rosales and Fabales, but also in large clades such as Malpighiales, in which more species diversity is tropical (52.76% are tropical under more conservative Köppen–Geiger definitions<sup>43</sup>; 67.21% under geographic definition; see Supplementary Table 1); even in clades that are primarily tropical, tip rates tend to be highest in non-tropical species (Fig. 2a, b). Although the estimated direction of the temperature–diversification relationship was consistent across analyses, relationships were not significant for many orders. This finding is likely due to a loss of power resulting both from corrections for multiple comparisons and from subtrees with limited numbers of tips<sup>50,51,54,63–65</sup>, and perhaps most importantly, the limited information that can be extracted when combining extant-only phylogenetic data with a single univariate trait<sup>65</sup>.

We found compelling evidence of a negative relationship between temperature and diversification in the rosids. These results are part of an emerging pattern in large-scale diversification studies<sup>10,13,14,66</sup> (but see refs. <sup>15,16</sup>) suggesting that non-tropical species exhibit high diversification rates despite the lower species richness of communities outside of the tropics. Our results are consistent with recent work that examined speciation rates across geographic space in angiosperms and showed lower angiosperm speciation in the tropics<sup>14</sup>. Our work significantly extends this effort by explicitly considering multiple definitions of tropicality, as well as the relationship between temperature and diversification over geological time and across multiple subclades. Our diversification results challenge some explanations of high species diversity in the tropics invoking higher tropical diversification rates<sup>21,67</sup> (also reviewed in Donoghue<sup>68</sup>). The negative correlation between species richness and diversification may instead suggest little turnover of ancient, species-rich communities in the tropics (the tropical conservatism hypothesis<sup>18,66</sup>) in contrast to the dynamism of high-latitude biomes in the recent past (the last 15 million years<sup>10,13,69–71</sup>). We note, however, that the tropics themselves are heterogeneous and that seasonally dry tropical forests are poorly understood but may form exceptions to the diversity gradient patterns found to date.

We determined clade ages across gridded rosid communities and found that the tropics generally contain older rosid communities (Fig. 2c, d) compared to younger communities in the high latitudes of the Northern Hemisphere. Those same communities also possess species with the highest diversification rates (Fig. 2) of anywhere else on the globe. Our finding of diversification rate differences between non-tropical areas of the Northern and Southern Hemispheres has been observed before<sup>72</sup> and is likely due to the greater historical climatic stability of the biomes of the Southern Hemisphere. That same dynamism suggests that extinction rates in especially temperate and polar regions may also be higher<sup>73</sup>, as both global cooling and perturbations during the Pleistocene likely impacted those regions more strongly. Because extinction rates are challenging to estimate from extant-only phylogenetic data<sup>46–48</sup>, we conservatively chose to focus on reporting speciation parameters (i.e., tip rates; but see net diversification reported in Fig. 3 and Supplementary Fig. 2).

Ecological opportunity due to Earth’s dynamic climatic past has been frequently invoked to explain high diversification rates in species currently inhabiting temperate<sup>14,69–71,74</sup> and arid<sup>62,75</sup> environments. Significantly, these studies do not typically invoke adaptive radiations. Rather, these diversifications appear to rely on a passive mechanism via bursts of speciation for lineages already well-adapted to these conditions, and in some cases acquisition of cold-tolerant traits quite clearly predates diversification<sup>10,69</sup>. Among the most important traits in temperate and cold environments, and a key, previously documented constraint on plant communities and flowering plant evolution is freezing tolerance (reviewed in Zanne et al.<sup>76</sup>). Freezing tolerance is underlain by a restricted set of traits<sup>76,77</sup> that has evolved many times via exaptation from existing vegetative variation<sup>78</sup>. Despite its many origins, freezing tolerance has been gained in fewer than half of all flowering plant families from their probable tropical ancestor<sup>76,79,80</sup>. Our results are consistent with supposition of pre-existing cold-tolerant traits across most rosid lineages, all of which are far older than the onset of cold climates in the Pliocene–Pleistocene. This was followed by a largely passive diversification in response to the strong spatial expansion of extra-tropical habitats beginning in the Miocene and mediated by what appears to be strong niche conservatism of cold-tolerant or intolerant traits<sup>68</sup>. This finding is, in essence, a refinement of the tropical conservatism hypothesis (first proposed by Wiens & Donoghue<sup>18</sup>). Further work on developing climatic layers over time that can trace the expansion of strong seasonality, and especially areas with seasonal freezing temperatures is a key next step. These layers, when coupled with more fine-grained, modeled past and present species distribution information would provide a strong basis for more explicit testing of this more mechanism-focused hypothesis.

Although our taxon sampling was extensive (~20,000 species), we recognize that our sampling of the enormous rosid clade (90,000–120,000 species) remains limited, introducing possible biases. However, these biases are inherent in any large-scale biodiversity analyses<sup>36,41</sup>. For instance, differences in sampling rate between tropical and non-tropical species could be important for estimating niche-specific speciation rates. Significantly, we found surprisingly little impact of sampling rates on the strong and repeated trends we observed across well-sampled clades, a robustness that is likely due to the relatively weak association of incomplete sampling with phylogeny<sup>81</sup> and with niche tropicality in this data set. Methods for modeling incomplete clade sampling do not replace the role of molecular phylogenetic data<sup>82,83</sup>. Lastly, fully incorporating estimates of phylogenetic uncertainty in a data set of nearly 20,000 species is currently computationally intractable with the models used here. However, in our previous work<sup>41,83</sup>, we have extensively studied the properties of this phylogenetic data set with respect to phylogenetic and diversification analyses. In particular, we performed a series of experiments on the properties of RPANDA and BAMM on the current rosid phylogenetic estimate and found diversification estimates to be robust on two supermatrix trees differing strongly in sampling and phylogenetic support<sup>83</sup> (see also similar sampling experiments in Igea & Tanentzap<sup>14</sup>). We believe our exploration of phylogenetic uncertainty is representative of current approaches given the phylogenetic scale considered.

Despite the challenges of computational tractability and sampling for megatrees, our work here still provides a baseline set of hypotheses to test further, with yet more robust sampling and improving tools, methods, and resources. Future studies on large phylogenetic, niche, and trait data sets such as those we have started to build in rosids will be critical for further identifying and refining drivers of the high diversity of rosids in many contemporary terrestrial habitats and potentially predicting how



these key members of present-day plant communities may respond during future, rapid climate warming.

## Methods

**5-Locus rosid supermatrix.** We used a tree generated from a supermatrix of 19,740 species of rosids<sup>41</sup> for diversification analysis. This is the most comprehensive existing phylogeny for the rosid clade. Briefly, to construct this tree, Sun et al.<sup>41</sup> mined GenBank for commonly used genetic regions using the PHYLogeny Assembly With Databases pipeline<sup>84</sup> (PHLAWD, version 3.4a, <https://github.com/blackrim/phlawd>). The regions used represent all three plant genomes, including the chloroplast genes *atpB*, *matK*, and *rbcl*, the mitochondrial gene *matR*, and the nuclear ribosomal ITS (i.e., ITS-1, 5.8S, and ITS-2 regions). We validated the species names following The Plant List (<http://www.theplantlist.org/>) and removed invalid names and any taxon names with “sp.”, “subsp.”, “var.”, “x”, “cf.” and “aff.”, or other non-species designations; we curated the 5-locus data set iteratively by screening individual loci and concatenated matrices for rogue taxon behavior through manual inspection of initial trees; detailed cleaning steps are fully discussed in Sun et al.<sup>41</sup>. After cleaning, we concatenated the loci and inferred the best ML tree with RAxML<sup>85</sup> using the extended Majority Rule Criterion (autoMRE) as a bootstrap stopping rule<sup>86</sup> (reached at 352 replicates). The final 5-locus data set contained 19,740 ingroup species (135 families and 17 orders) and 554 outgroup species; the latter were excluded from downstream diversification analyses.

Divergence time estimation was conducted in treePL v.1.0<sup>87</sup> using 59 calibration points covering 15 of the 17 rosid orders<sup>24,41</sup> with a root constraint following Wang et al.<sup>25</sup>. For determining community age (Fig. 2c, d), we defined the age of a tip species as the age of its closest internal tree node and reported community-level ages as the grid cell median using the R package *dggridR* v. 2.0.3<sup>88</sup>. We then used both climatic and geographic definitions of tropicality in order to test for an age difference between tropical and non-tropical species (Fig. 2d).

We assessed the quality of the final tree by comparing the topology and recovered clade ages to a series of previous benchmark studies<sup>26,39</sup> and compared sampling statistics to the OTT taxonomy<sup>22</sup>. More details of tree quality control are discussed in Sun et al.<sup>41,83</sup>. Overall, we found broad congruence with previous results (reviewed in Sun et al.<sup>39,41</sup>), but with increased resolution in many areas of the phylogeny enabled by denser taxon sampling. Sampling remains incomplete (also see Supplementary Table 7), but this phylogeny represents a strong improvement for the clade with more than twice the coverage of previous efforts (reviewed in Sun et al.<sup>41</sup>).

Given both our motivation to use the rosid clade for evolutionary replication of diversification patterns and its prohibitive size for many diversification methods, we split the phylogenetic tree into 17 subclasses corresponding to all rosid orders recognized in APG IV<sup>24</sup>. Likewise, for computational feasibility, all analyses were run in parallel across 17 order subtrees, unless otherwise stated.

**Assembly of species distribution data.** We queried all rosid species sampled in our tree from iDigBio and GBIF using R packages *rgbif* v1.3.0<sup>44</sup> and *ridigbio* v0.3.5<sup>45</sup> on June 4th, 2019. No further taxonomic name resolution was conducted for these rosid names, because all rosid species names in our tree were validated and reconciled by The Plant List and OpenTree databases in Sun et al.<sup>41</sup>, and we queried rosid species occurrence records with accepted rosid names. During our initial data acquisition, we extracted only data records that were georeferenced and excluded any coordinates with zero and/or integer latitude and longitude. We then used Python scripts from Folk et al.<sup>10</sup> and custom R scripts to remove geographic outliers beyond three standard deviations of Euclidean distance from the geographic centroid of species occurrences and to filter missing environmental data. For more details on processing occurrence records, see Supplementary Method 1.

**Temperature data layers.** We used four data sets to assess the relationship between temperature and rosid diversification rates. We first assessed the relationship of species-specific diversification rates (i.e., tip rates; see below) with contemporary temperature niche by calculating species mean temperature. For each species, we associated all assembled occurrence points with average monthly annual mean temperature data from 1970 to 2000 using WorldClim (i.e., bio1). We then used a geological historical oxygen isotope ( $\delta^{18}\text{O}$ ) curve from the Late Cretaceous to the present (~113 Myr to present<sup>59</sup>), converting these values to 5-point temperature means following Folk et al.<sup>10</sup>, to assess the impact of global historical temperature trends on diversification rates as detailed below.

Many previous studies have invoked high diversification rate as an explanation for high biodiversity in the tropics<sup>89–91</sup>, although this has proved controversial (e.g., reviewed in Rabosky<sup>92</sup>). We tested this hypothesis by developing two further binary data sets to represent a tropical/non-tropical contrast. We used both geographic and climatic definitions to classify rosid species as occurring either inside or outside of tropical environments. For the geographic definitions, we calculated the mean latitude of occurrence records for each species and classified a species as tropical if this mean value fell between the Tropics of Cancer and Capricorn (23.43677°N and 23.43677°S), and non-tropical (i.e., temperate + polar) if it fell outside this interval. The climatic tropicality data set used the standard Köppen–Geiger climatic tropics definition as calculated by Owens et al.<sup>43</sup>, which defines as tropical those regions with year-round monthly mean temperatures >18 °C. We associated species

occurrences with the Owens et al.<sup>43</sup> data set and classified each species based on the more frequent value (e.g., species with >50% of occurrences outside tropical sites were classified as non-tropical). We tested for the presence of phylogenetic signal in the three contemporary niche data sets using the lambda transform and Likelihood Ratio Test (Pagel’s  $\lambda$ ) functions in R packages *phytools* v0.7.00<sup>52</sup> and *geiger* v2.0.6.2<sup>93</sup>. This test assumes a phylogenetic signal definition of niche conservatism, one among several that have been proposed<sup>53,94</sup>.

**Diversification rate inference and statistical analysis.** We first used BAMM v2.5.0<sup>49</sup> to explore the diversification profiles of the rosids, and we inferred species-specific diversification rates (i.e., tip rates), diversification rate shifts, and diversification rate-through-time matrices and plots. Then we applied these diversification metrics, together with other diversification methods and the statistical tests introduced below, to test, assess, and cross-check the correlation pattern we observed.

For the initial BAMM method, we confirmed convergence of the MCMC chains and effective sample sizes >200 for both the number of shifts and log likelihoods, after discarding 10% burn-in (Supplementary Table 7). We accounted for incomplete sampling in BAMM<sup>49</sup> and RPANDA<sup>60</sup> based on species richness summarized from the OTT database<sup>41</sup> (Supplementary Table 7).

We used tip rates as estimated in BAMM and DR<sup>17,83</sup> (Supplementary Method 3) as reliable estimators to investigate speciation rate dynamics and trait-dependent (e.g., temperature-dependent) speciation as reviewed in Title & Rabosky<sup>95</sup>. We implemented a series of explicitly phylogenetic parametric and semi-parametric methods to test the relationship of tip rates to mean annual temperature and to the tropicality of contemporary rosid species (Supplementary Method 2). For parametric analyses, we used STRAPP<sup>50</sup> as implemented in BAMMtools<sup>96</sup>, testing each of the 17 rosid clades recognized as orders. To cross-check the correlation pattern observed by STRAPP in BAMM, semi-parametric analyses were employed. We used tip rates estimated by DR in the es-SIM tests<sup>51</sup> for the continuous mean annual temperature data set and FISSE tests<sup>54</sup> for the categorical tropicality data sets. Because it was computationally feasible to use semi-parametric methods for larger trees, we tested both the 17 rosid ordinal clades and the total tree; STRAPP was only used on the ordinal subtrees. All tests across rosid orders were controlled for family-wise error by the Hochberg<sup>97</sup> method.

As an additional parametric approach to test for associations between tropicality and diversification rate, and to test for potential unobserved diversification drivers, we fit two binary state speciation and extinction (BiSSE<sup>56</sup>) and two hidden-state speciation and extinction (HiSSE<sup>55</sup>) models. These approaches estimate state-specific rates of diversification and state transition, assessing whether diversification rates are significantly different for the alternative states and whether unobserved traits influence diversification rates under a model selection framework. These analyses were executed in the R package HiSSE v1.9.6<sup>55</sup> (for further implementation details, see Supplementary Method 3 and Supplementary Table 4).

In addition, we explored the spatial diversification patterns of rosids using tip rates estimated from both BAMM and DR via the R package *dggridR* v. 2.0.3<sup>88</sup> to view global patterns (Fig. 2a, b).

**Diversification rates and historical temperature.** To explore the patterns of global paleo-temperature on rosid diversification, we used RPANDA v1.4<sup>60</sup> to fit a series of time- and temperature-dependent likelihood diversification models<sup>3,60</sup>. The latter models used the global historical temperature data set described above. Speciation and extinction dependence were modeled as all possible combinations of constant, linear, and exponential relationships, as well as pure-birth models for all three relationships. All models used are described in Supplementary Method 5.

We also used metrics from BAMM to test the relationship of diversification rate with historical global temperature found in RPANDA. We extracted rate-through-time curves in 100 time-unit slices and fit linear and exponential regression models with temperature, using AIC for model choice (Supplementary Fig. 2). Furthermore, we used diversification rate-through-time matrices and plots generated by BAMM v2.5.0<sup>49</sup> to assess the relationship between geological historical temperature and net diversification rate curves for tropical and non-tropical lineages, respectively (Fig. 3). These analyses, unlike similar analyses in RPANDA (see “Results” section), do not account for temporal autocorrelation of diversification rates. However, given that RPANDA results show that the temperature–diversification relationship we observed cannot be explained by time-dependency alone, we have included regression results to present an independent method with easily interpreted correlation statistics.

We accounted for incomplete phylogenetic sampling in BAMM, STRAPP, RPANDA, and HiSSE/BiSSE with global sampling fractions based on species richness summarized from the OTT database<sup>41</sup> (Supplementary Table 7). Semi-parametric DR and es-SIM methods do not parameterize incomplete sampling so these approaches were used directly (but see the similarity of tip rates across methods, Fig. 2a, b). The best model was selected by the smallest AIC value (or  $\Delta\text{AIC}$ ) and largest AW for HiSSE and RPANDA.

**Sensitivity tests.** We assessed the sensitivity of our results to potentially uneven occurrence data accumulation between tropical and non-tropical species, the

former of which are expected to be under-sampled<sup>98</sup>, by a series of random taxon dropping experiments (Supplementary Method 5). For each ordinal subtree, 10%, 30%, or 50% of non-tropical species were dropped while all tropical species were retained. We then re-ran STRAPP analyses following the methods described above and assessed the impact on estimated tip rates and downstream analyses.

**Reporting summary.** Further information on research design is available in the Nature Research Reporting Summary linked to this article.

### Data availability

Phylogenetic trees, species distribution data, WorldClim temperature data, and BMM diversification rate-through-time matrices are available at [https://github.com/Cactusolo/rosid\\_NCOMMS-19-37964-T](https://github.com/Cactusolo/rosid_NCOMMS-19-37964-T) (<https://doi.org/10.5281/zenodo.3843441>)<sup>99</sup>; the source data underlying Fig. 2, pie-chart of Fig. 3, and Supplementary Fig. 3 are provided as a Source Data file; all other relevant intermediate data are available from the authors upon request.

### Code availability

All code used in this study is available at [https://github.com/Cactusolo/rosid\\_NCOMMS-19-37964-T](https://github.com/Cactusolo/rosid_NCOMMS-19-37964-T) (<https://doi.org/10.5281/zenodo.3843441>)<sup>99</sup>.

Received: 15 November 2019; Accepted: 5 June 2020;

Published online: 03 July 2020

### References

- Soltis, P. S., Folk, R. A. & Soltis, D. E. Darwin review: angiosperm phylogeny and evolutionary radiations. *Proc. R. Soc. Lond. B* **286**, 20190099 (2019).
- Sauquet, H. & Magallón, S. Key questions and challenges in angiosperm macroevolution. *N. Phytol.* **219**, 1170–1187 (2018).
- Morlon, H. Phylogenetic approaches for studying diversification. *Ecol. Lett.* **17**, 508–525 (2014).
- Allen, A. P., Gillooly, J. F., Savage, V. M. & Brown, J. H. Kinetic effects of temperature on rates of genetic divergence and speciation. *Proc. Natl Acad. Sci. USA* **103**, 9130–9135 (2006).
- Wang, W. et al. Menispermaceae and the diversification of tropical rainforests near the Cretaceous–Paleogene boundary. *N. Phytol.* **195**, 470–478 (2012).
- Cantalapiedra, J. L. et al. Dietary innovations spurred the diversification of ruminants during the Cenozoic. *Proc. Biol. Sci.* **281**, 20132746 (2014).
- Claramunt, S. & Cracraft, J. A new time tree reveals Earth history's imprint on the evolution of modern birds. *Sci. Adv.* **1**, e1501005 (2015).
- Pires, M. M., Rankin, B. D., Silvestro, D. & Quesada, T. B. Diversification dynamics of mammalian clades during the K–Pg mass extinction. *Biol. Lett.* **14**, 20180458 (2018).
- Stadler, T. Mammalian phylogeny reveals recent diversification rate shifts. *Proc. Natl Acad. Sci. USA* **108**, 6187–6192 (2011).
- Folk, R. A. et al. Rates of niche and phenotype evolution lag behind diversification in a temperate radiation. *Proc. Natl Acad. Sci. USA* **116**, 10874–10882 (2019).
- Huang, X. C., German, D. A. & Koch, M. A. Temporal patterns of diversification in Brassicaceae demonstrate decoupling of rate shifts and mesopolyploidization events. *Ann. Bot.* **125**, 29–47 (2020).
- Vieites, D. R., Min, M.-S. & Wake, D. B. Rapid diversification and dispersal during periods of global warming by plethodontid salamanders. *Proc. Natl Acad. Sci. USA* **104**, 19903–19907 (2007).
- Rabosky, D. L. et al. An inverse latitudinal gradient in speciation rate for marine fishes. *Nature* **559**, 392–395 (2018).
- Igea, J. & Tanentzap, A. J. Angiosperm speciation cools down in the tropics. *Ecol. Lett.* **23**, 692–700 (2020).
- Cardillo, M., Orme, C. D. L. & Owens, I. P. F. Testing for latitudinal bias in diversification rates: An example using new world birds. *Ecology* **86**, 2278–2287 (2005).
- Ricklefs, R. E. Global variation in the diversification rate of passerine birds. *Ecology* **87**, 2468–2478 (2006).
- Jetz, W., Thomas, G. H., Joy, J. B., Hartmann, K. & Mooers, A. O. The global diversity of birds in space and time. *Nature* **491**, 444–448 (2012).
- Wiens, J. J. & Donoghue, M. J. Historical biogeography, ecology and species richness. *Trends Ecol. Evol.* **19**, 639–644 (2004).
- Schluter, D. & Pennell, M. W. Speciation gradients and the distribution of biodiversity. *Nature* **546**, 48–55 (2017).
- Economo, E. P., Narula, N., Friedman, N. R., Weiser, M. D. & Guénard, B. Macroecology and macroevolution of the latitudinal diversity gradient in ants. *Nat. Commun.* **9**, 1778 (2018).
- Ricklefs, R. E. Estimating diversification rates from phylogenetic information. *Trends Ecol. Evol.* **22**, 601–610 (2007).
- Hinchliff, C. E. et al. Synthesis of phylogeny and taxonomy into a comprehensive tree of life. *Proc. Natl Acad. Sci. USA* **112**, 12764–12769 (2015).
- Govaert, R. How many species of seed plants are there? *Taxon* **50**, 1085–1090 (2001).
- Byng, J. W., Christenhusz, M., Fay, M. F. & Chase, M. An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG IV. *Bot. Linn. Soc.* **181**, 1–20 (2016).
- Wang, H. C. et al. Rosid radiation and the rapid rise of angiosperm-dominated forests. *Proc. Natl Acad. Sci. USA* **106**, 3853–3858 (2009).
- Magallón, S., Gómez-Acevedo, S., Sánchez-Reyes, L. L. & Hernández-Hernández, T. A meta-calibrated timetree documents the early rise of flowering plant phylogenetic diversity. *N. Phytol.* **207**, 437–453 (2015).
- Soltis, P. S., Soltis, D. E., Chase, M. W., Endress, P. K. & Crane, P. R. In *The Tree of Life* (eds Cracraft, J. & Donoghue, M.) 154–167 (Oxford & New York, 2004).
- Boyce, C. K., Lee, J. E., Field, T. S., Brodribb, T. J. & Zwieniecki, M. A. Angiosperms helped put the rain in the rainforests: The impact of plant physiological evolution on tropical biodiversity. *Ann. Missouri Bot. Gard.* **97**, 527–540 (2010).
- Moreau, C. S. & Bell, C. D. Testing the museum versus cradle tropical biological diversity hypothesis: phylogeny, diversification, and ancestral biogeographic range evolution of the ants. *Evolution* **67**, 2240–2257 (2013).
- Roelants, K. et al. Global patterns of diversification in the history of modern amphibians. *Proc. Natl Acad. Sci. USA* **104**, 887–892 (2007).
- Bininda-Emonds, O. R. et al. The delayed rise of present-day mammals. *Nature* **446**, 507–512 (2007).
- Feldberg, K. et al. Epiphytic leafy liverworts diversified in angiosperm-dominated forests. *Sci. Rep.* **4**, 5974 (2014).
- Schneider, H., Schuettpelz, E., Pryer, K. M. & Cranfill, R. Ferns diversified in the shadow of angiosperms. *Nature* **428**, 553–557 (2004).
- Magallón, S., Crane, P. R. & Herendeen, P. S. Phylogenetic pattern, diversity, and diversification of eudicots. *Ann. Mo. Bot. Gard.* **86**, 297–372 (1999).
- Stevens, P. F. Angiosperm phylogeny website. Version 14. <http://www.mobot.org/MOBOT/research/APweb/>. Accessed 1 January 2018 (2017).
- Folk, R. A. et al. Challenges of comprehensive taxon sampling in comparative biology: Wrestling with rosids. *Am. J. Bot.* **105**, 433–445 (2018).
- Töpel, M., Antonelli, A., Yesson, C. & Eriksen, B. Past climate change and plant evolution in western North America: a case study in Rosaceae. *PLoS ONE* **7**, e50358 (2012).
- Xiang, X. G. et al. Large-scale phylogenetic analyses reveal fagalean diversification promoted by the interplay of diaspores and environments in the Paleogene. *Perspect. Plant. Ecol.* **16**, 101–110 (2014).
- Sun, M. et al. Phylogeny of the Rosidae: a dense taxon sampling analysis. *J. Syst. Evol.* **54**, 363–391 (2016).
- Zhao, L. et al. Phylogenomic analyses of large-scale nuclear genes provide new insights into the evolutionary relationships within the rosids. *Mol. Phylogenet. Evol.* **105**, 166–176 (2016).
- Sun, M. et al. Exploring the phylogeny of rosids with a five-locus supermatrix from GenBank. Preprint at <https://www.biorxiv.org/content/10.1101/694950v1> (2019).
- Beaulieu, J. M. & O'Meara, B. C. Can we build it? Yes we can, but should we use it? Assessing the quality and value of a very large phylogeny of campanulid angiosperms. *Am. J. Bot.* **105**, 417–432 (2018).
- Owens, H. L. et al. The latitudinal diversity gradient in New World swallowtail butterflies is caused by contrasting patterns of out-of-and into-the-tropics dispersal. *Glob. Ecol. Biogeogr.* **26**, 1447–1458 (2017).
- Chamberlain, S. et al. rgbif: Interface to the Global Biodiversity Information Facility API. R package version 1.3.0. <https://CRAN.R-project.org/package=rgbif> (2019).
- Michonneau, F. & Collins, M. ridigbio: Interface to the iDigBio data API. R package version 0.3.5. <https://CRAN.R-project.org/package=ridigbio> (2017).
- Rabosky, D. L. Extinction rates should not be estimated from molecular phylogenies. *Evolution* **64**, 1816–1824 (2010).
- Pyron, R. A. & Burbrink, F. T. Phylogenetic estimates of speciation and extinction rates for testing ecological and evolutionary hypotheses. *Trends Ecol. Evol.* **28**, 729–736 (2013).
- Revell, J. L. Comparing the rates of speciation and extinction between phylogenetic trees. *Ecol. Evol.* **8**, 5303–5312 (2018).
- Rabosky, D. L. Automatic detection of key innovations, rate shifts, and diversity-dependence on phylogenetic trees. *PLoS ONE* **9**, e89543 (2014).
- Rabosky, D. L. & Huang, H. A robust semi-parametric test for detecting trait-dependent diversification. *Syst. Biol.* **65**, 181–193 (2016).
- Harvey, M. G. & Rabosky, D. L. Continuous traits and speciation rates: alternatives to state-dependent diversification models. *Methods Ecol. Evol.* **9**, 984–993 (2018).

52. Revell, J. L. Phytools: an R package for phylogenetic comparative biology (and other things). *Methods Ecol. Evol.* **3**, 217–223 (2012).
53. Wiens, J. J. et al. Niche conservatism as an emerging principle in ecology and conservation biology. *Ecol. Lett.* **13**, 1310–1324 (2010).
54. Rabosky, D. L. & Goldberg, E. E. FISSE: a simple nonparametric test for the effects of a binary character on lineage diversification rates. *Evolution* **71**, 1432–1442 (2017).
55. Beaulieu, J. M. & O'Meara, B. C. Detecting hidden diversification shifts in models of trait-dependent speciation and extinction. *Syst. Biol.* **65**, 583–601 (2016).
56. Maddison, W. P., Midford, P. E. & Otto, S. P. Estimating a binary character's effect on speciation and extinction. *Syst. Biol.* **56**, 701–710 (2007).
57. Akaike, H. A new look at the statistical model identification. *IEEE Trans. Autom. Contr.* **19**, 716–723 (1974).
58. Wagenmakers, E. J. & Farrell, S. AIC model selection using Akaike weights. *Psychon. Bull. Rev.* **11**, 192–196 (2004).
59. Cramer, B. S., Toggweiler, J. R., Wright, J. D., Katz, M. E. & Miller, K. G. Ocean overturning since the Late Cretaceous: Inferences from a new benthic foraminiferal isotope compilation. *Paleoceanography* **24**, PA4216 (2009).
60. Morlon, H. et al. RPANDA: an R package for macroevolutionary analyses on phylogenetic trees. *Methods Ecol. Evol.* **7**, 589–597 (2016).
61. Pound, M. J., Haywood, A. M., Salzmann, U. & Riding, J. B. Global vegetation dynamics and latitudinal temperature gradients during the Mid to Late Miocene (15.97–5.33 Ma). *Earth Sci. Rev.* **112**, 1–22 (2012).
62. Wang, Q. G. et al. Niche conservatism and elevated diversification shape species diversity in drylands: evidence from Zygophyllaceae. *Proc. R. Soc. Lond. B* **285**, 20181742 (2018).
63. Rabosky, D. L. & Goldberg, E. E. Model inadequacy and mistaken inferences of trait-dependent speciation. *Syst. Biol.* **64**, 340–355 (2015).
64. Gamisch, A. Notes on the statistical power of the Binary State Speciation and Extinction (BiSSE) model. *Evol. Bioinformatics* **12**, 165–174 (2016).
65. Grundler, M. C. & Rabosky, D. L. Macroevolutionary analysis of discrete traits with rate heterogeneity. Preprint at <https://www.biorxiv.org/content/10.1101/2020.01.07.897777v1.full> (2020).
66. Weir, J. T. & Schluter, D. The latitudinal gradient in recent speciation and extinction rates of birds and mammals. *Science* **315**, 1574–1576 (2007).
67. Buzas, M. A., Collins, L. S. & Culver, S. J. Latitudinal difference in biodiversity caused by higher tropical rate of increase. *Proc. Natl Acad. Sci. USA* **99**, 7841–7843 (2002).
68. Donoghue, M. J. A phylogenetic perspective on the distribution of plant diversity. *Proc. Natl Acad. Sci. USA* **105**, 11549–11555 (2008).
69. Nürk, N. M., Uribe-Convers, S., Gehrke, B., Tank, D. C. & Blattner, F. R. Oligocene niche shift, Miocene diversification–Cold tolerance and accelerated speciation rates in the St. John's Worts (*Hypericum*, Hypericaceae). *BMC Evol. Biol.* **15**, 80 (2015).
70. King, Y. & Ree, R. H. Uplift-driven diversification in the Hengduan mountains, a temperate biodiversity hotspot. *Proc. Natl Acad. Sci. USA* **114**, E3444–E3451 (2017).
71. Lu, L. M. et al. Evolutionary history of the angiosperm flora of China. *Nature* **554**, 234–238 (2018).
72. Dunn, R. R. et al. Climatic drivers of hemispheric asymmetry in global patterns of ant species richness. *Ecol. Lett.* **12**, 324–333 (2009).
73. Jablonski, D. Extinction and the spatial dynamics of biodiversity. *Proc. Natl Acad. Sci. USA* **105**(Suppl 1), 11528–11535 (2008).
74. Schluter, D. Speciation, ecological opportunity, and latitude (American Society of Naturalists Address). *Am. Nat.* **187**, 1–18 (2016).
75. Arakaki, M. et al. Contemporaneous and recent radiations of the world's major succulent plant lineages. *Proc. Natl Acad. Sci. USA* **108**, 8379–8384 (2011).
76. Zanne, A. E. et al. Functional biogeography of angiosperms: life at the extremes. *N. Phytol.* **218**, 1697–1709 (2018).
77. Diaz, S. et al. The global spectrum of plant form and function. *Nature* **529**, 167–171 (2016).
78. Zanne, A. E. et al. Three keys to the radiation of angiosperms into freezing environments. *Nature* **506**, 89–92 (2014).
79. Ricklefs, R. E. & Renner, S. S. Species richness within families of flowering plants. *Evolution* **48**, 1619–1636 (1994).
80. Preston, J. C. & Sandve, S. R. Adaptation to seasonality and the winter freeze. *Front. Plant Sci.* **4**, 167 (2013).
81. Pender, J. E. et al. How sensitive are climatic niche inferences to distribution data sampling? A comparison of biota of North America Program (BONAP) and Global Biodiversity Information Facility (GBIF) datasets. *Ecol. Informatics* **54**, 100991 (2019).
82. Rabosky, D. L. No substitute for real data: a cautionary note on the use of phylogenies from birth–death polytomy resolvers for downstream comparative analyses. *Evolution* **69**, 3207–3216 (2015).
83. Sun, M. et al. Estimating rates and patterns of diversification with incomplete sampling: A case study in the rosids. *Am. J. Bot.* **107**, 1–15 (2020).
84. Smith, S. A., Beaulieu, J. M. & Donoghue, M. J. Mega-phylogeny approach for comparative biology: An alternative to supertree and supermatrix approaches. *BMC Evol. Biol.* **9**, 37 (2009).
85. Stamatakis, A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30**, 1312–1313 (2014).
86. Pattengale, N. D. et al. (eds) *ISBRA* Vol. 6053, 128–139 (Springer, 2010).
87. Smith, S. A. & O'Meara, B. C. treePL: divergence time estimation using penalized likelihood for large phylogenies. *Bioinformatics* **28**, 689–690 (2012).
88. Barnes, R. dgggridR: discrete global grids for R. R package version 0.1.12. <https://github.com/r-barnes/dgggridR/> (2017).
89. Jablonski, D., Roy, K. & Valentine, J. W. Out of the tropics: evolutionary dynamics of the latitudinal diversity gradient. *Science* **306**, 102–106 (2006).
90. Kozak, K. H. & Wiens, J. J. Accelerated rates of climatic-niche evolution underlie rapid species diversification. *Ecol. Lett.* **13**, 1378–1389 (2010).
91. Mullen, S. P., Savage, W. K., Wahlberg, N. & Willmott, K. R. Rapid diversification and not clade age explains high diversity in neotropical Adelpha butterflies. *Proc. R. Soc. Lond. B* **278**, 1777–1785 (2011).
92. Rabosky, D. L. Ecological limits on clade diversification in higher taxa. *Am. Nat.* **173**, 662–674 (2009).
93. Pennell, M. W. et al. Geiger v2.0: An expanded suite of methods for fitting macroevolutionary models to phylogenetic trees. *Bioinformatics* **30**, 2216–2218 (2014).
94. Losos, J. B. Phylogenetic niche conservatism, phylogenetic signal and the relationship between phylogenetic relatedness and ecological similarity among species. *Ecol. Lett.* **11**, 995–1003 (2008).
95. P. O. & Rabosky, D. L. Tip rates, phylogenies and diversification: what are we estimating, and how good are the estimates? *Methods Ecol. Evol.* **10**, 821–834 (2019).
96. 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).
97. Hochberg, Y. A sharper Bonferroni procedure for multiple tests of significance. *Biometrika* **75**, 800–803 (1988).
98. Meyer, C., Kreft, H., Guralnick, R. & Jetz, W. Global priorities for an effective information basis of biodiversity distributions. *Nat. Commun.* **6**, 8221 (2015).
99. Sun, M. & Folk, R. A. Code and data for rosid\_NCOMMS-19-37964. Zenodo. <https://doi.org/10.5281/zenodo.3843441> (2020).

## Acknowledgements

This work was supported by the U.S. National Science Foundation (DEB-1208809 to D.E.S.; DEB-1442280 to P.S.S. and D.E.S.; DBI-1458640 to D.E.S. and P.S.S.; DBI-1547229 to P.S.S.; DBI-1523667 to R.A.F.; DEB-1916632 to R.A.F., R.P.G., P.S.S. and D.E.S.), U. S. Department of Energy (DE-SC0018247 to R.A.F., R.P.G., P.S.S. and D.E.S.), the Strategic Priority Research Program of the Chinese Academy of Sciences (both XDA19050103 and XDB31000000 to Z.D.C.), and National Natural Science Foundation of China (Grant No. 31590822 to Z.D.C.). We also thank Dr. Stephen A. Smith for comments on the first draft and Dr. Daijiang Li for diagnosing the analytical capability of function “fit\_env” in RPANDA. The HiPerGator cluster at the University of Florida provided extensive computational resources.

## Author contributions

M.S., R.A.F. and R.P.G. designed the study; M.S. and R.A.F. performed the analyses and interpreted the results; M.S., R.A.F., D.E.S. and R.P.G. drafted the manuscript; R.A.F., R.P.G., D.E.S., P.S.S., Z.D.C. and M.A.G. revised the manuscript. All authors contributed to and approved the final manuscript.

## Competing interests

The authors declare no competing interests.

## Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41467-020-17116-5>.

Correspondence and requests for materials should be addressed to M.S., R.A.F. or R.P.G.

Peer review information *Nature Communications* thanks Evan Economo, Renske Onstein and the other, anonymous, reviewer for their contribution to the peer review of this work. Peer reviewer reports are available.

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020