Title: Multiple macroevolutionary routes to becoming a biodiversity hotspot

Authors: J. Igea[1]*, A. J. Tanentzap[1]*

Affiliations:

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

*Correspondence to: ji247@cam.ac.uk, ajt65@cam.ac.uk

Abstract:

Why is species diversity so unevenly distributed across different regions on Earth? Regional differences in biodiversity may stem from differences in rates of speciation and dispersal and colonization times, but these hypotheses have rarely been tested simultaneously at a global scale. Here we uncovered the routes that generated hotpots of mammal and bird biodiversity by analyzing the tempo and mode of diversification and dispersal within major biogeographic realms. Hotspots in tropical realms had higher rates of speciation whereas those in temperate realms received more immigrant species from their surrounding regions. We also found that hotspots had higher spatial complexity and energy availability, providing a link between the environment and macroevolutionary history. Our study highlights how assessing differences in macroevolutionary history can help to explain why biodiversity varies so much worldwide.

One Sentence Summary:

Global hotspots of biodiversity arise from faster species generation in the tropics and higher migration in temperate realms.

Main Text:

Biodiversity is extremely unevenly distributed across the globe and understanding why has long fascinated biologists (*1*). For example, there are many exceptions to the tendency for species richness to increase towards the Equator - widely studied as the latitudinal diversity gradient (*1*–*3*). This finer-scale association between biodiversity and geography (*4*) is exemplified by the 35 terrestrial biodiversity hotspots proposed by Myers *et al*. (*5*, *6*) for conservation purposes based on plant endemicity and habitat loss. A third of Myers' hotspots were located in temperate zones and were more diverse than many regions closer to the Equator, demonstrating that high levels of species richness can also be found outside the tropics.

Regional differences in biodiversity may ultimately arise through at least one of three macroevolutionary routes. First, differences in historic rates of *in situ* diversification (i.e., speciation minus extinction) can result in more species accumulating in some areas than others. Second, differences in historic rates of lineage dispersal can result in some areas acting as sources of species that are exported elsewhere and some that are sinks that import species (*7*). Finally, an older age of colonization of a region may promote diversity if there was more time to accumulate species, generating "museums" of biodiversity (*8*), as formalized in the "time-for speciation" hypothesis (*9*, *10*). However, most studies of regional diversity patterns have not compared the relative importance of these different potential routes (*3*).

The three macroevolutionary routes that give rise to regional differences in biodiversity are at least partially paved by the environment (*1*). Some environmental variables may favor cladogenesis, such as past tectonic movements that generate isolation (*11*), whilst others may

favor the establishment of immigrant species, such as historically stable climates that create regional refuges for species during periods of global change (*12*) and favor dispersal from environmentally similar areas, such as because of niche conservatism (*13*). Other environmental variables may favor both cladogenesis and immigration. For example, higher environmental energy might promote speciation by increasing mutation rates and shortening generation times (*14*), and also allow regions to hold more species by expanding their carrying capacity (*15*). Similarly, a higher speciation rate and local carrying capacity are both associated with physiographic heterogeneity (*16*, *17*) and habitat complexity (*18*, *19*).

Here we use terrestrial hotspots of mammal and bird biodiversity to understand how different macroevolutionary routes (*20*) generate extreme spatial differences in species diversity. We delineated hotspots using the number of species in an area divided by the inverse of range of those species. Hotspots based on this measure - known as weighted endemicity (WE, ref. *21*) and not to be confused with counting endemic species - identify the contribution of each area to global biodiversity more accurately than species richness, because widespread species are not counted in every area where they occur and so do not have a disproportionate influence on the metric. Therefore, hotspots based on WE will be more representative of the distribution of biodiversity across multiple regions on Earth than hotspots based on species richness, which are exclusively centered in the tropics. Using diversification rate and historical biogeography inference methods, we then tested which macroevolutionary routes could better explain the existence of mammal and bird hotspots across different regions on Earth. Efforts to reconstruct explicitly the historical rates of migration and diversification of biodiversity hotspots have

largely focused on small clades or specific geographic regions (*20*, *22*, *23*), without a broader global context.

We first used global species maps (*24*, *25*) to delineate hotspots. After overlaying species ranges with a grid of 100×100 km cells, we defined separate mammal and bird hotspots for subsequent analyses as the richest 20% of cells in terms of WE (Figs. S1a, S1b). We chose this threshold to obtain clade-specific hotspots that were roughly equivalent in size to Myers' hotspots (*5, 6*). Despite being partially delineated using habitat threat, we used Myers *et al*. (5) as a basis for comparison because they identified large spatial unevenness in biodiversity. They observed that 20% of the global land area was sufficient to retain many of the most biodiverse biomes on the planet (e.g., Andes, Sundaland, Madagascar, Mediterranean Basin) and >40% of all vertebrate species (*6*). Our resulting WE-based hotspots were largely overlapping with both Myers' hotspots and hotspots of total species richness (see Methods), suggesting that all the measures captured a similar biological pattern.

Using hotspots and neighboring regions within six biogeographic realms, we assessed the spatial variation in the accumulation of ancient and recent lineages. To do this, we first obtained species-specific rates of diversification by estimating the Diversification Rate (DR) metric (*26*) for the most comprehensive phylogenies of mammals and birds (see Methods). DR captures the number of historic diversification events that lead to a given species, weighted by the relative age of those events, but does not explicitly model extinction (*26*, *27*). We then built linear regression models that predicted the richness of species that were either ancient (i.e., older, with DR values in the 1st quartile of the distribution) or recent (younger, with DR values in the 4th quartile of the

distribution). Total species richness in each cell (*27*) was the sole model predictor. By examining the residuals of these linear models, we determined which cells had an excess or a deficit of species in each of the two quartiles within each biogeographic realm (*27*). Our results revealed that hotspots generally had a deficit of ancient lineages and an excess of recent lineages when compared to neighboring regions, except in birds where evidence of the latter was mixed (Fig. 1). We confirmed that these results were robust both to phylogenetic uncertainty and the method used for estimating diversification rates (Fig. S2, Table S1, see Methods).

Next, we assessed whether differences in macroevolutionary routes generated the different patterns of accumulation of ancient and recent species in hotspot and non-hotspot regions across biogeographic realms. We found that the general deficit of ancient lineages and more variable excess of recent lineages in the hotspots compared to nearby regions resulted from contrasting macroevolutionary histories across biogeographic realms. We reached this conclusion by reconstructing assembly dynamics within hotspots and non-hotspots of each realm using historical biogeographic inference (*28*). We found that both mammal and bird species were generated at faster rates in the last 25 million years (Ma) within hotspots compared with nearby regions of largely tropical realms like Australasia, Indo-Malay and the Neotropics. By contrast, cladogenetic rates were similar or lower than surrounding areas in hotspots of the Afrotropics and temperate Palearctic and Nearctic realms (Fig. 2). Therefore, *in situ* cladogenesis could explain the accumulation of biodiversity in most tropical but not temperate hotspots. In temperate but not tropical realms, greater rates of historical dispersal rather than *in situ* cladogenesis could explain the accumulation of mammal and bird diversity within hotspots (Fig. 3). We found no evidence that colonization age alone could explain the differences in

biodiversity as hotspots were generally colonized later than non-hotspot regions in temperate realms, and there was no consistent difference in the age of colonization across tropical realms (Fig. S3a).

Together, our findings suggest that contrasting macroevolutionary routes have shaped the uneven distribution of biodiversity across biogeographic realms. In all primarily tropical realms, except the Afrotropics, hotspots consistently generated and exported species at higher rates than their nearby areas, whereas the disproportionate richness in hotspots of temperate realms could be explained by greater rates of immigration from surrounding regions. The Afrotropics may lay somewhere outside these two routes. Afrotropical hotspots did not generate species more quickly or import them at greater rates in the last 25 Ma. The region became more arid during the late Miocene and early Pliocene as the Sahara Desert was formed (*29*). This change in the regional climate could have generated differences in the extinction dynamics of the Afrotropics hotspots compared to the non-hotspots. However, we could not estimate such extinction dynamics with the available methods. Hotspot diversity may have also been greater for ecological rather than evolutionary reasons, e.g. greater niche space (*15*).

Our results were generally robust to the methodological assumptions. First, species richness is positively correlated with region size (*30*), but we found no evidence that the difference in size between hotspots and non-hotspot regions could alone explain our results. As hotspots of endemicity were defined globally, there were large differences between the sizes of the hotspots and non-hotspot regions within each realm (Table S2). To assess whether these differences could generate the different macroevolutionary patterns that we observed, we repeated our analysis by

randomly sampling combinations of cells with similar size and spatial structure to the hotspot cells in each realm. We consistently found that the real estimates of *in situ* cladogenesis (Fig. S4), dispersal (Fig. S5) and colonization time (Fig. S3b) lay outside the simulated distributions. Therefore, the differences in macroevolutionary patterns that we observed between hotspots and their surrounding areas must have stemmed from differences in species composition and/or environmental features rather than simply due to size. Second, we also found that hotspots were more clustered in space than non-hotspot cells across all realms, particularly in tropical as compared with temperate realms. However, the differences across realms were small and <10% in the most extreme case (Table S3, Fig. S6). The slightly greater clustering of tropical hotspots is therefore unlikely to explain fully the different macroevolutionary routes that we found in tropical and temperate realms (Fig. S6). Third, we confirmed that two alternative ways of delineating biodiversity hotspots were congruent with the results for the WE-based hotspots (Supplementary Text, Figs. S7, S8). These alternate definitions used species richness (Figs. S1c, S1d) and areas where narrow-ranged species occurred (Figs. S1e, S1f), which have been proposed to reflect past opportunities for speciation (*31*).

Finally, we found evidence that unique environments inside the hotspots could have promoted differences in macroevolution when compared to neighboring non-hotspot regions. Linear models with spatial autocorrelation allowed us to compare environmental features potentially associated with differences in rates of *in situ* diversification and dispersal between hotspots and their surrounding areas within each realm. We found that hotspots had a greater mean net primary productivity, terrain ruggedness index and more habitats than their surrounding regions (Fig. 4). These differences were consistent across realms despite the contrasting

macroevolutionary routes between tropical and temperate regions. This finding is not entirely surprising because the same environmental features can be associated with contrasting macroevolutionary routes. Specifically, increased historic opportunities for speciation may have resulted from higher energy availability (*14*) and spatial complexity (*17*, *18*) in the tropics. In temperate regions, the same variables may have elevated carrying capacities and packed more immigrant species into the hotspots, which could have also acted as biodiversity refuges during past climate change (*32*, *33*). The velocity of late Quaternary climate-change was not, however, different between hotspot and non-hotspot areas (Fig. 4), despite previously being shown to correlate negatively with vertebrate endemism (*12*). The extent of the tropics was also larger over much over the past 25 Ma (*3*). Thus, more immigrants could have dispersed into temperate latitudes, i.e. "out-of-tropics" hypothesis (*34*), providing an additional explanation for the different macroevolutionary routes that we observed across realms.

Our study offers an integrative approach to understand why biodiversity varies so much across the globe by using a regional-scale and spatially-explicit reconstruction of historical dispersal and diversification alongside an analysis of ecological gradients. Generally, vertebrate, invertebrate and plant diversity are spatially correlated at regional scales across the planet (*4*), so we expect similar mechanisms to generate biodiversity across the Tree of Life (*35*). Similar analyses carried out in other groups may nevertheless result in clade-specific idiosyncrasies. For instance, the relative roles of *in situ* cladogenesis and dispersal as drivers of regional diversity may be different in taxa with lower vagility than mammals and birds, such as amphibians and insects (*36*, *37*). By simultaneously comparing different macroevolutionary routes and their

macroecological features in two major vertebrate clades, our study now provides a new answer to the old question of why diversity varies so much across the world.

Materials and Methods

Phylogenetic and species distribution data

The maximum clade credibility tree (MCC) for mammals was estimated with 100 random trees from the pseudoposterior provided by Kuhn *et al.* (*38*) using TreeAnnotator (*39*) v.1.8.2. Following Rolland *et al.* (*40*), we recalibrated the dates in this MCC tree and in the initial 100 trees from the pseudoposterior with alternative dates from Meredith *et al.* (*41*) using PATHd8 (*42*). For birds, we used 100 random trees from the updated version of the posterior distribution of trees in Jetz *et al.* (*26*) and we obtained the MCC tree as above.

Species distribution data were obtained from the IUCN Red list for mammals (*24*) (version 5.2) and Birdlife (*25*) (version 6.0) for birds. Marine species were not analyzed. Hotspots were defined using all species with distribution data – 5302 and 11093 species of mammals and birds, respectively. Species names in the phylogenetic trees were standardized with the IUCN (version 5.2) and Birdlife (version 6.0) taxonomies and collated with the distributional data. In total, 4633 and 9622 species of mammals and birds, respectively representing 83.3% and 86.5% of the described species were present in the phylogenetic trees and were used in downstream analyses. We followed the definition of the WWF Simplified Biogeographical Realms (*43*) (version 2.0) for the different biogeographic realms. By definition, the areas in each realm share a common evolutionary and biogeographic history, so comparing them provides a framework for

generalization across realms with distinct biotas (*43*). The Oceanic realm was much smaller than the rest, so we did not include it in our analyses.

Definition of the biodiversity hotspots

We overlaid the distributions of mammals and breeding ranges of birds onto a 100×100 km grid to define hotspots of biodiversity (Figs. S1a, S1b). Weighted endemism (WE) in each grid cell was calculated by weighting each species' occurrence by the inverse of its corresponding range and then summing values across all species in a cell. We then defined the hotspots as the cells with the top 20% of WE values. This definition covered 19.9% of our grid surface, roughly equivalent in size to Myers' hotspots that comprised 17.0% of global land surface (*6*), and resulted in 3826 and 3304 cells for mammals and birds, respectively. There was a 74.4% overlap between the mammal and bird hotspots, and a 69.7 and 71.3% overlap between the Myers' hotspots and the ones we defined for mammals and birds, respectively (Fig. S1). We also found an overlap between our WE-based hotspots and hotspots defined simply with total species richness (SR) of 56.9% and 62.4 % for mammals and birds, respectively). Overlap was expected since WE and SR were positively correlated (Spearman's $\rho = 0.74$ in our dataset).

We also assessed whether alternative ways of delineating biodiversity hotspots changed our results. First, we defined hotspots based on species richness (SR). Hotspots were defined as the cells with the top 20% values of SR and, as expected, were concentrated in the tropics (Figs. S1c, S1d). Two realms, Australasia and the Nearctic, had very few or no hotspot cells or showed a very small overlap with the WE hotspots (Table S4). Therefore, we were only able to carry out further analyses within the Afrotropics, Indo-Malay, Neotropics and the Palearctic. Second, we

delineated alternative hotspots as 'centers of endemism' where narrow-ranged species (NRS) were found. These regions have been proposed to provide many opportunities for past speciation while enabling the survival of narrow-ranged endemics due to stable environments (*31*). Following ref. *31*, we defined narrow-ranged species as species with a range of ≤ 100,000 km^2 (10 cells in our grid), and hotspots as all the cells where these species were found ($n_{mammals}$ = 2723 cells, n_{birds} = 2041 cells). The resulting hotspots substantially overlapped with the WE-based hotspots, particularly for mammals (Figs. S1e, S1f, Table S4). We repeated the DR and historical biogeography analyses described in the main text for the SR and NRS-based hotspots and found that the results were largely congruent with the WE-based hotspots (Supplementary Text, see Figs. S7, S8).

Species diversification rates

The Diversification Rate metric (*26*) (DR) was calculated as the number of nodes that separated each species from the root of the tree weighted by the distance of each node to the present. Thus, DR represents the relative phylogenetic isolation of species and has been used to determine regions that are enriched in actively diversifying lineages and older lineages (*27*, *44*). We divided the species into four quartiles, with older lineages (here termed "ancient") in the first quartile (Q1) and younger, actively diversifying lineages ("recent") in the fourth quartile (Q4). Using simple linear regression models, we predicted Q1 and Q4 species richness from the total species richness in each cell. Positive residuals of the corresponding linear model indicated that a particular cell contained an excess of ancient or recent lineages, respectively, while negative residuals indicated a deficit (*27*, *44*). For each biogeographic realm, we then assessed whether an average hotspot cell had different values for the Q1 and Q4 residuals when compared to an

average non-hotspot cell. To do this, we ran two spatial simultaneous autoregressive (SAR) error models, each with the Q1 and Q4 residuals as a response and hotspot category as predictor using the R package *spdep* (*45*). Finally, we compared the mean values of Q1 and Q4 SAR residuals in hotspots and non-hotspot cells of each biogeographic realm using a Wilcoxon rank sum test.

We also analyzed the effect of phylogenetic uncertainty in our estimates of diversification rates. We obtained DR estimates for 100 random trees from the pseudoposterior distribution of the mammal and bird phylogenies (see above). We then verified that the MCC-based DR estimates (presented in the main text) were positively correlated with the median DR values across the 100 trees (Fig. S2a) and with each individual tree (Fig. S2b). We also verified that the residuals of the DR quartile linear regressions with the MCC tree were correlated with the residual values across the 100 trees (Figs. S2c-S2d).

DR has been shown to correlate to recent speciation events. However, it does not account for extinction and temporal variation in diversification rates within lineages (*26*, *27*). To check if these assumptions could influence our analyses, we estimated species-specific rates of speciation and extinction (i.e., "tip-rates") in the mammal and bird phylogenies using Bayesian Analysis of Macroevolutionary Mixtures (BAMM) (*46*). BAMM models diversification rate heterogeneity across lineages and through time. Although the reliability of BAMM has recently been questioned, a focus of criticism was the influence of unobserved rate shifts within extinct lineages on diversification estimates (*47*). This issue affects most methods but the contribution of shifts in extinct lineages to the overall probability of extinction is marginal under biologically plausible scenarios (*48*). A related concern is that extinction rate estimates obtained from

molecular phylogenies of extant species may sometimes be unreliable (*49*, *50*). Accurate speciation rates and relative diversification rates can however be routinely inferred using a variety of methods (*51*), including BAMM. Simulation studies have shown BAMM infers speciation rates with high accuracy, particularly for large, well-sampled trees such as ours (*48*) and can robustly estimate speciation rates in the absence of paleontological data (*52*). Despite having different assumptions, BAMM and DR estimates of diversification are generally positively correlated (*53*), including in our dataset (Figs. S3e, S3f; Pearson's $r_{mammals} = 0.63$; $r_{birds} = 0.58$; for both, p-value < 0.0001). This result suggests that both measure capture similar information about diversification. We also found that residuals of DR-based and BAMM-based quartile regressions were positively correlated (Figs. S3e, S3f; Pearson's $r_{mammals.Q1} = 0.27$; $r_{mammals.Q4} = 0.68$; $r_{birds.Q1} = 0.65$; $r_{birds.Q4} = 0.77$; for all, p-value < 0.0001). Similarly, comparisons between hotspots and non-hotspot regions across realms using DR residuals (Fig. 1) largely matched those obtained using BAMM residuals (Table S1).

Historical biogeography analysis

We inferred the historical biogeography of mammals and birds using the R package Biogeography with Bayesian [and Likelihood] Evolutionary Analysis in R Scripts (BioGeoBEARS) (*28*, *54*). BioGeoBEARS infers biogeographic history using phylogenetic trees and locality data. Ancestral ranges were estimated by implementing a series of biogeographic models in BioGeoBEARS that allowed for different types of range shifts and that could be compared in a common likelihood framework. Here, we used AICc to compare the fit of six biogeographical models: Dispersal-Extinction-Cladogenesis (DEC, ref. *55*), Dispersal-Vicariance Analysis (DIVA, ref. *56*), and Bayesian Analysis of Biogeography

(BAYAREALIKE, ref. *57*), and versions of the previous three models that allowed for founder-event speciation (+J, ref. *54*), i.e., DEC+J, DIVA+J, BAYAREA-LIKE+J. All the models account for anagenetic dispersal and local extinction and differ in the cladogenetic events that they allow: DEC models subset sympatry and narrow vicariance, DIVA models subset narrow and widespread vicariance, and BAYAREA models subset narrow and widespread sympatry (see ref. *28* for details). To alleviate computational issues related to an excessive number of states (i.e. geographic areas), we performed six independent analyses, one for each biogeographic realm. In each analysis, we used the complete phylogenetic tree and coded the species as present in a maximum of two areas: hotspots in a focal realm, non-hotspot areas in a focal realm, and in any one of the remaining five realms. No constraints to dispersal were set between areas. Presence was assigned when at least 20% of the total species range overlapped a focal region. We then used the parameter rate estimates of the best fitting model (Table S5) for each realm to perform 50 Biogeographic Stochastic Mappings (BSMs) (*58*).

The BSMs first allowed us to count, date and extract the chronology of anagenetic and cladogenetic events in a probabilistic sample of biogeographic histories for hotspots and non-hotspot regions within each realm. We then binned the events into 2-Ma periods and calculated rolling estimates of the rates of regional dispersal out of and into each region following Xing and Ree (*23*). Dispersal rates for each region were calculated as the median number of dispersal events in a time bin across the BSMs divided by the number of lineages present in that region in the previous time bin. For instance, dispersal rates 2-4 Ma from non-hotspots into hotspots (N-H) in the Afrotropics were calculated as the median number of N-H events 2-4 Ma divided by the median number of lineages estimated in the hotspots in the Afrotropics at 4-6 Ma. To calculate

rates of cladogenesis, we first used the BSMs to assign nodes to particular geographic regions. We then obtained the branch estimates of speciation rates for the complete phylogenetic trees with BAMM using the R package BAMMtools (*59*). Finally, we calculated the rate of cladogenesis for a given region (hotspot or non-hotspot) in a time bin as the average of the speciation rates of the branches assigned to that region in that time bin. We discarded estimates that were older than 26 Ma, due to the lack of data (i.e. ≤10 lineages in each region), and younger than 2 Ma, to avoid confounding effects of ongoing speciation events (*26*).

Size effect in historical biogeography analysis

We assessed whether the differences in size between hotspot and non-hotspot regions within realms could explain the differences in macroevolutionary patterns that we observed. We first fitted a cluster point process model to the point pattern of the hotspots in each realm using the R package *spatstat* (*60*). We used these models to simulate 50 sets of "control" hotspots with similar size and spatial structure to the real datasets (Fig. S9). Finally, we repeated the biogeographic inference and the BSMs as detailed above for each of the 50 control hotspots in each realm. We then compared whether the rate estimates obtained with the real datasets were significantly different from the distribution of rate estimates obtained with the simulated hotspots. We found differences between the real and the simulated datasets that suggested that size alone could not explain our results: *i*) rates of *in situ* cladogenesis in the Palearctic and Nearctic hotspots were higher in the real than in the simulated datasets (Fig. S4a); *ii*) rates of *in situ* cladogenesis in the non-hotspot regions in Australasia, Indo-Malay and the Neotropics were smaller in the real than in the simulated datasets (Fig. S4b); and *iii*) rates of dispersal from the

hotspots into non-hotspot regions were consistently higher in the real than in the simulated datasets across all realms (Fig. S5a).

Contiguity of hotspots across biogeographic realms

We determined whether different spatial clustering of hotspot and non-hotspot areas within each realm could influence our results. For each realm, we calculated the median distance of each cell to all neighboring cells of the same class (hotspot or non-hotspot) in a 1000-km radius (Fig. S6). We then compared the differences of the mean distances among hotspots and non-hotspots using a *t*-test.

Environmental differences between hotspots and non-hotspots

We investigated whether there were significant differences between hotspots and surrounding non-hotspot regions for variables that have previously been implicated in generating biodiversity gradients (*12*, *14–19*). The models compared the mean values between hotspots and non-hotspots in the same realm for variables related to: *i*) current energy availability (net primary productivity, NPP); *ii*) spatial heterogeneity (number of habitats and terrain ruggedness index, TRI); and *iii*) historical stability (climate change velocity since the Last Glacial Maximum for mean annual rainfall and mean annual temperature, past tectonic movements). We used simultaneous autoregressive error models that controlled for spatial autocorrelation as implemented in the R package *spdep* (*45*), with hotspot category as a predictor and each environmental variable as a response. The spatial neighborhood matrix was calculated for neighbors within 1,000 km. Variables were centered and scaled to a mean of 0 and s.d. of 1 prior to model fitting so as to generate standardized effects that would be comparable across the different variables. P-values

were Bonferroni-corrected for multiple comparisons. Data for NPP were obtained from the Socioeconomic Data and Applications Center (SEDAC) (*61*). Climate change velocity was calculated following Sandel et al (*12*) and climatic variables for the present and the Last Glacial Maximum (with the MIROC model) were obtained from WorldClim version 1.4 (*62*). We obtained habitat data from the USGS 1 km Global Land Cover Characteristics Data Base Version 2.0, and used the global 30 arc second digital elevation model (GTOPO30, also available from the USGS) to calculate cell-average values of TRI using the *raster* R package (*63*). Tectonic movement was calculated as the standard deviation in the distances between a cell and its neighbors from 65 Ma until the present (*11*), as reconstructed with *gplates* (*64*). We also considered additional variables (mean elevation, elevation range) but they were highly correlated (Spearman's $\rho > 0.70$) to other variables in our dataset and so removed prior to the analyses.

References and Notes

1. D. Schluter, M. W. Pennell, Speciation gradients and the distribution of biodiversity. *Nature*. **546**, 48–55 (2017).
2. E. R. Pianka, Latitudinal gradients in species diversity: a review of concepts. *Am. Nat.* **100** (1966), pp. 33–46.
3. G. G. Mittelbach *et al.*, Evolution and the Latitudinal Diversity Gradient: speciation, extinction and biogeography. *Ecol. Lett.* **10**, 315–331 (2007).
4. G. Kier *et al.*, A global assessment of endemism and species richness across island and mainland regions. *Proc. Natl. Acad. Sci.* **106**, 9322–9327 (2009).
5. N. Myers, R. A. Mittermeier, C. G. Mittermeier, G. A. B. da Fonseca, J. Kent, Biodiversity hotspots for conservation priorities. *Nature*. **403**, 853–858 (2000).
6. R. A. Mittermeier *et al.*, *Hotspots revisited* (CEMEX, Mexico City, 2004).

7. R. E. Ricklefs, Community diversity: relative roles of local and regional processes. *Science.* **235**, 167–171 (1987).

8. K. J. Gaston, T. M. Blackburn, The tropics as a museum of biological diversity: an analysis of the New World avifauna. *Proc. R. Soc. B Biol. Sci.* **263**, 63–68 (1996).

9. P. R. Stephens, J. J. Wiens, Explaining species richness from continents to communities: the time-for-speciation effect in emydid turtles. *Am. Nat.* **161**, 112–128 (2003).

10. J. J. Wiens, The causes of species richness patterns across space, time, and clades and the role of "ecological limits." *Q. Rev. Biol.* **86**, 75–96 (2011).

11. G. F. Ficetola, F. Mazel, W. Thuiller, Global determinants of zoogeographical boundaries. *Nat. Ecol. Evol.* **1**, 0089 (2017).

12. B. Sandel *et al.*, The influence of Late Quaternary climate-change velocity on species endemism. *Science.* **334**, 660–664 (2011).

13. J. J. Wiens, C. H. Graham, Niche conservatism: integrating evolution, ecology, and conservation biology. *Annu. Rev. Ecol. Evol. Syst.* **36**, 519–539 (2005).

14. A. P. Allen, J. F. Gillooly, V. M. Savage, J. H. Brown, Kinetic effects of temperature on rates of genetic divergence and speciation. *Proc. Natl. Acad. Sci.* **103**, 9130–9135 (2006).

15. D. J. Currie, Energy and large-scale patterns of animal- and plant-species richness. *Am. Nat.* **137**, 27–49 (1991).

16. M. J. Steinbauer *et al.*, Topography-driven isolation, speciation and a global increase of endemism with elevation. *Glob. Ecol. Biogeogr.* **25**, 1097–1107 (2016).

17. A. Stein, K. Gerstner, H. Kreft, Environmental heterogeneity as a universal driver of species richness across taxa, biomes and spatial scales. *Ecol. Lett.* **17**, 866–880 (2014).

18. O. Paun *et al.*, Processes driving the adaptive radiation of a tropical tree (*Diospyros*,

Ebenaceae) in New Caledonia, a biodiversity hotspot. *Syst. Biol.* **65**, 212–227 (2016).

19. J. T. Kerr, L. Packer, Habitat heterogeneity as a determinant of mammal species richness in high-energy regions. *Nature*. **385**, 252–254 (1997).

20. C. E. Hughes, Are there many different routes to becoming a global biodiversity hotspot? *Proc. Natl. Acad. Sci. U. S. A.* **114**, 4275–4277 (2017).

21. M. D. Crisp, S. Laffan, H. P. Linder, A. Monro, Endemism in the Australian flora. *J. Biogeogr.* **28**, 183–198 (2001).

22. L. G. Cook, N. B. Hardy, M. D. Crisp, Three explanations for biodiversity hotspots: small range size, geographical overlap and time for species accumulation. An Australian case study. *New Phytol.* **207**, 390–400 (2015).

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

24. The IUCN Red List of Threatened Species. Version 5.2, (available at http://www.iucnredlist.org).

25. BirdLife International and Handbook of the Birds of the World (2016).

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

27. J. D. Kennedy *et al.*, Into and out of the tropics: The generation of the latitudinal gradient among New World passerine birds. *J. Biogeogr.* **41**, 1746–1757 (2014).

28. N. J. Matzke, Probabilistic historical biogeography: new models for founder-event speciation, imperfect detection, and fossils allow improved accuracy and model-testing. *Front. Biogeogr.* **26**, 217–220 (2013).

29. A. Micheels, J. Eronen, V. Mosbrugger, The Late Miocene climate response to a modern

Sahara desert. *Glob. Planet. Change*. **67**, 193–204 (2009).

30. D. Storch, P. Keil, W. Jetz, Universal species–area and endemics–area relationships at continental scales. *Nature*. **488**, 78–81 (2012).

31. W. Jetz, C. Rahbek, R. K. Colwell, The coincidence of rarity and richness and the potential signature of history in centres of endemism. *Ecol. Lett.* **7**, 1180–1191 (2004).

32. F. Médail, K. Diadema, Glacial refugia influence plant diversity patterns in the Mediterranean Basin. *J. Biogeogr.* **36**, 1333–1345 (2009).

33. S. Harrison, R. Noss, Endemism hotspots are linked to stable climatic refugia. *Ann. Bot.* **119**, 207–214 (2017).

34. D. Jablonski, K. Roy, J. W. Valentine, Out of the tropics: evolutionary dynamics of the latitudinal diversity gradient. *Science*. **314**, 102–6 (2006).

35. V. S. F. T. Merckx *et al.*, Evolution of endemism on a young tropical mountain. *Nature*. **524**, 347–350 (2015).

36. Y. Kisel, T. G. Barraclough, Speciation has a spatial scale that depends on kevels of gene flow. *Am. Nat.* **175**, 316–334 (2010).

37. I. Medina, G. M. Cooke, T. J. Ord, Walk, swim or fly? Locomotor mode predicts genetic differentiation in vertebrates. *Ecol. Lett.* **21**, 638–645 (2018).

38. T. S. Kuhn, A. Ø. Mooers, G. H. Thomas, A simple polytomy resolver for dated phylogenies. *Methods Ecol. Evol.* **2**, 427–436 (2011).

39. A. J. Drummond, A. Rambaut, BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol. Biol.* **7**, 214 (2007).

40. J. Rolland, F. L. Condamine, F. Jiguet, H. Morlon, Faster speciation and reduced extinction in the Tropics contribute to the mammalian Latitudinal Diversity Gradient.

PLoS Biol. **12**, e1001775 (2014).

41. R. Meredith, J. Janečka, J. Gatesy, O. Ryder, C. Fisher, Impacts of the cretaceous terrestrial revolution and KPg extinction on mammal diversification. *Science.* **334** (2011).

42. T. Britton, C. L. Anderson, D. Jacquet, S. Lundqvist, K. Bremer, Estimating divergence times in large phylogenetic trees. *Syst. Biol.* **56**, 741–752 (2007).

43. D. M. Olson *et al.*, Terrestrial ecoregions of the world: a new map of life on earth. A new global map of terrestrial ecoregions provides an innovative tool for conserving biodiversity. *Bioscience*. **51**, 933–938 (2001).

44. J. D. Kennedy *et al.*, Does the colonization of new biogeographic regions influence the diversification and accumulation of clade richness among the Corvides (Aves: Passeriformes)? *Evolution.* **71**, 38–50 (2017).

45. R. Bivand, G. Piras, Comparing implementations of estimation methods for spatial econometrics. *J. Stat. Softw.* **63**, 1–36 (2015).

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

47. B. R. Moore, S. Höhna, M. R. May, B. Rannala, J. P. Huelsenbeck, Critically evaluating the theory and performance of Bayesian analysis of macroevolutionary mixtures. *Proc. Natl. Acad. Sci.* **113** (2016), doi:10.1073/pnas.1518659113.

48. D. L. Rabosky, J. S. Mitchell, J. Chang, Is BAMM flawed? Theoretical and practical concerns in the analysis of multi-rate diversification models. *Syst. Biol.* **66**, 477–498 (2017).

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

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

51. H. Morlon, Phylogenetic approaches for studying diversification. *Ecol. Lett.* **17**, 508–525 (2014).

52. J. S. Mitchell, R. S. Etienne, D. L. Rabosky, R. Ree, Inferring diversification rate variation from phylogenies with fossils. *Syst. Biol.* (2018), doi:10.1093/sysbio/syy035.

53. M. G. Harvey *et al.*, Positive association between population genetic differentiation and speciation rates in New World birds. *Proc. Natl. Acad. Sci. U. S. A.* **114**, 6328–6333 (2017).

54. N. J. Matzke, Model selection in historical biogeography reveals that founder-event speciation is a crucial process in island clades. *Syst. Biol.* **63**, 951–970 (2014).

55. R. H. Ree, S. A. Smith, Maximum Likelihood inference of geographic range evolution by dispersal, local extinction, and c ladogenesis. *Syst. Biol.* **57**, 4–14 (2008).

56. I. Sanmartin, H. Enghoff, F. Ronquist, Patterns of animal dispersal, vicariance and diversification in the Holarctic. *Biol. J. Linn. Soc.* **73**, 345–390 (2001).

57. M. J. Landis, N. J. Matzke, B. R. Moore, J. P. Huelsenbeck, Bayesian analysis of Biogeography when the number of areas is large. *Syst. Biol.* **62**, 789–804 (2013).

58. J. Dupin *et al.*, Bayesian estimation of the global biogeographical history of the Solanaceae. *J. Biogeogr.* **44**, 887–899 (2017).

59. D. L. Rabosky *et al.*, BAMMtools: an R package for the analysis of evolutionary dynamics on phylogenetic trees. *Methods Ecol. Evol.* **5**, 701–707 (2014).

60. A. Baddeley, E. Rubak, R. Turner, *Spatial point patterns: methodology and applications with r* (Chapman and Hall/CRC Press, London, 2015; http://www.crcpress.com/Spatial-

Point-Patterns-Methodology-and-Applications-with-R/Baddeley-Rubak-Turner/9781482210200/).

61. M. L. Imhoff *et al.*, Global patterns in human consumption of net primary production. *Nature*. **429**, 870–873 (2004).

62. R. J. Hijmans, S. E. Cameron, J. L. Parra, P. G. Jones, A. Jarvis, Very high resolution interpolated climate surfaces for global land areas. *Int. J. Climatol.* **25**, 1965–1978 (2005).

63. R. J. H. & J. van Etten, raster: Geographic analysis and modeling with raster data (2012), (available at http://cran.r-project.org/package=raster).

64. S. E. Williams, R. D. Müller, T. C. W. Landgrebe, J. M. Whittaker, An open-source software environment for visualizing and refining plate tectonic reconstructions using high-resolution geological and geophysical data sets. *GSA Today*, 4–9 (2012).

Acknowledgments: We thank T. Jucker, V. Soria-Carrasco and J. Garcia-Porta for useful comments that helped improve the manuscript. We thank the Gatsby Charitable Foundation (Grant Number GAT2962), Wellcome Trust (Grant Number 105602/Z/14/Z) and Isaac Newton Trust (Grant Number 17.24r) for funding. J.I and A.J.T designed the study. J.I. performed the analysis. J.I. and A.J.T interpreted the analysis and wrote the manuscript. The authors declare no competing interests. Relevant scripts and input files can be found at https://doi.org/10.6084/m9.figshare.6804500.v1

Fig. 1. Hotspots are poor in ancient lineages and sometimes rich in recent lineages. Residuals from linear models predicting cell-specific richness of a) ancient and b) recent lineages in hotspots (H, shown in red) and non-hotspot regions (N, shown in blue). Positive residuals indicate a regional excess of ancient/recent lineages and negative residuals indicate a deficit. The lower panels show the median of the difference between a random hotspot point (H) and a random non-hotspot point (N) and the 95% confidence interval around that median calculated with a Wilcoxon rank sum test.

Fig. 2. Contrasting rates of *in situ* cladogenesis in hotspots compared to surrounding non-hotspot regions. a) *In situ* cladogenesis rates between 2 to 26 Ma ago within non-hotspots were subtracted from rates within hotspots in each of six biogeographic realms and divided by the overall standard deviation to allow for comparison across realms. Solid lines indicate median differences ± 90% confidence interval. Intervals overlapping the dotted line indicate a lack of statistically significant differences at $\alpha = 0.10$. b) Differences for each 2-Ma time bin.

Fig. 3. Source-sink dynamics of hotspots and their surrounding regions. a) Dispersal rates between 2 to 26 Ma ago from non-hotspots to hotspots (N→H) were subtracted from hotspot to non-hotspots (H→N) rates within each realm. Lines and shaded areas presented as in Figure 2.

Fig. 4. Hotspots are more spatially complex and have more energy than surrounding regions. Standardized differences in the average cell values of environmental variables between hotspots

and non-hotspot cells within each of 6 biogeographic realms based on the coefficients of univariate spatially-autocorrelated linear regressions. Blank areas indicate non-significant differences. NPP = net primary productivity; CCV = climate change velocity; TRI = terrain ruggedness index.

List of Supplementary Materials:

Figs. S1 to S9

Tables S1 to S5

Supplementary Text

Science
AAAS

Supplementary Materials for

Multiple macroevolutionary routes to becoming a biodiversity hotspot

Authors: J. Igea*, A. J. Tanentzap*

Correspondence to: ji247@cam.ac.uk, ajt65@cam.ac.uk

This PDF file includes:

Figs. S1 to S9
Tables S1 to S5
Supplementary Text

Science
AAAS

Supplementary Materials for

Multiple macroevolutionary routes to becoming a biodiversity hotspot

Authors: J. Igea*, A. J. Tanentzap*

Correspondence to: ji247@cam.ac.uk, ajt65@cam.ac.uk

This PDF file includes:

Figs. S1 to S9
Tables S1 to S5
Supplementary Text

Supplementary Text

We first repeated the DR residuals and the historical biogeographic analyses described in the main text using the SR hotspots (Fig. S1c, S1d). We found largely congruent results with the WE-based hotspots: *i*) SR-based hotspots were generally poor in ancient lineages and rich in recent lineages, like WE-based hotspots (Fig. 1, Fig. S7a); *ii*) speciation was generally higher in hotspots than in non-hotspots in tropical realms, while the reverse was true in the Palearctic (Fig. 2, Fig. S8a); and *iii*) hotspots imported species from non-hotspots at faster rates than they exported them in the Palearctic, and the reverse was true in tropical realms (Fig. 3, Fig. S8a).

We similarly repeated the analyses using NRS-based hotspots (Figs. S1e, S1f). Similar to the main analyses based on WE, NRS-based hotspots were poor in ancient lineages and generally rich in recent lineages (Fig. 2, Fig. S7b). As in our main analysis, speciation was higher in non-hotspots than in hotspots in temperate realms. Similarly, in largely tropical realms, most of the parameters had the same sign as in our main analysis (i.e., speciation was generally larger in hotspots than in non-hotspots), but, in several instances, the intervals around the mean differences of speciation in hotspots minus non-hotspots overlapped with zero (Fig. 2, Fig. S8b). Likewise, dispersal results for temperate realms matched the WE-based results but the differences in tropical realms were largely not statistically significant (Fig. 3, Fig. S8b).

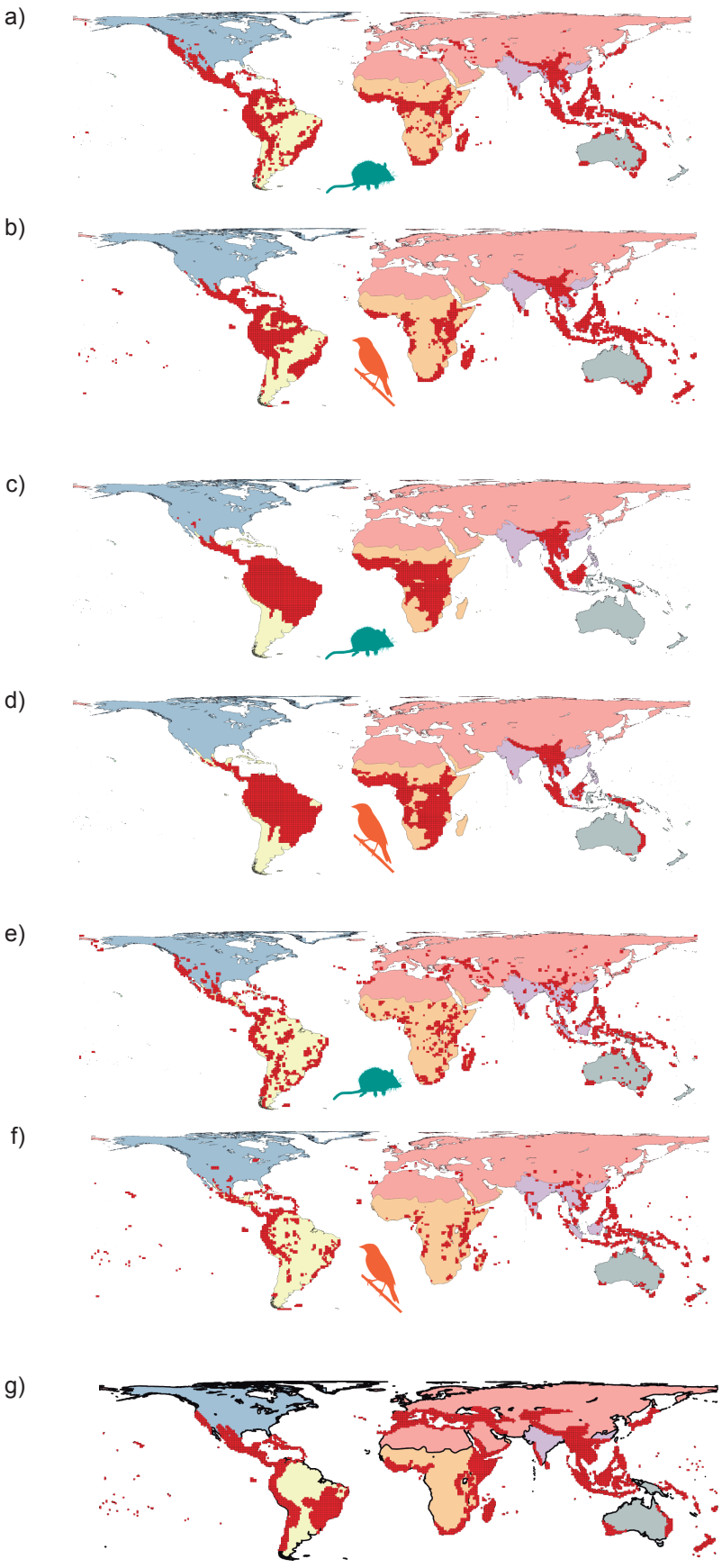

a)
b)
c)
d)
e)
f)
g)

Figure S1 Global maps of the mammal and bird hotspots in this study (shown in red) delineated using weighted endemicity **(**a,b**)**; alternate hotspots using species richness (c,d); alternate hotspots using narrow-ranged species (e,f); and Myers conservation hotspots (g). Afrotropics are shown in orange, Australasia in grey, Indo-Malay in purple, Nearctic in blue, Neotropics in yellow and Palearctic in pink.

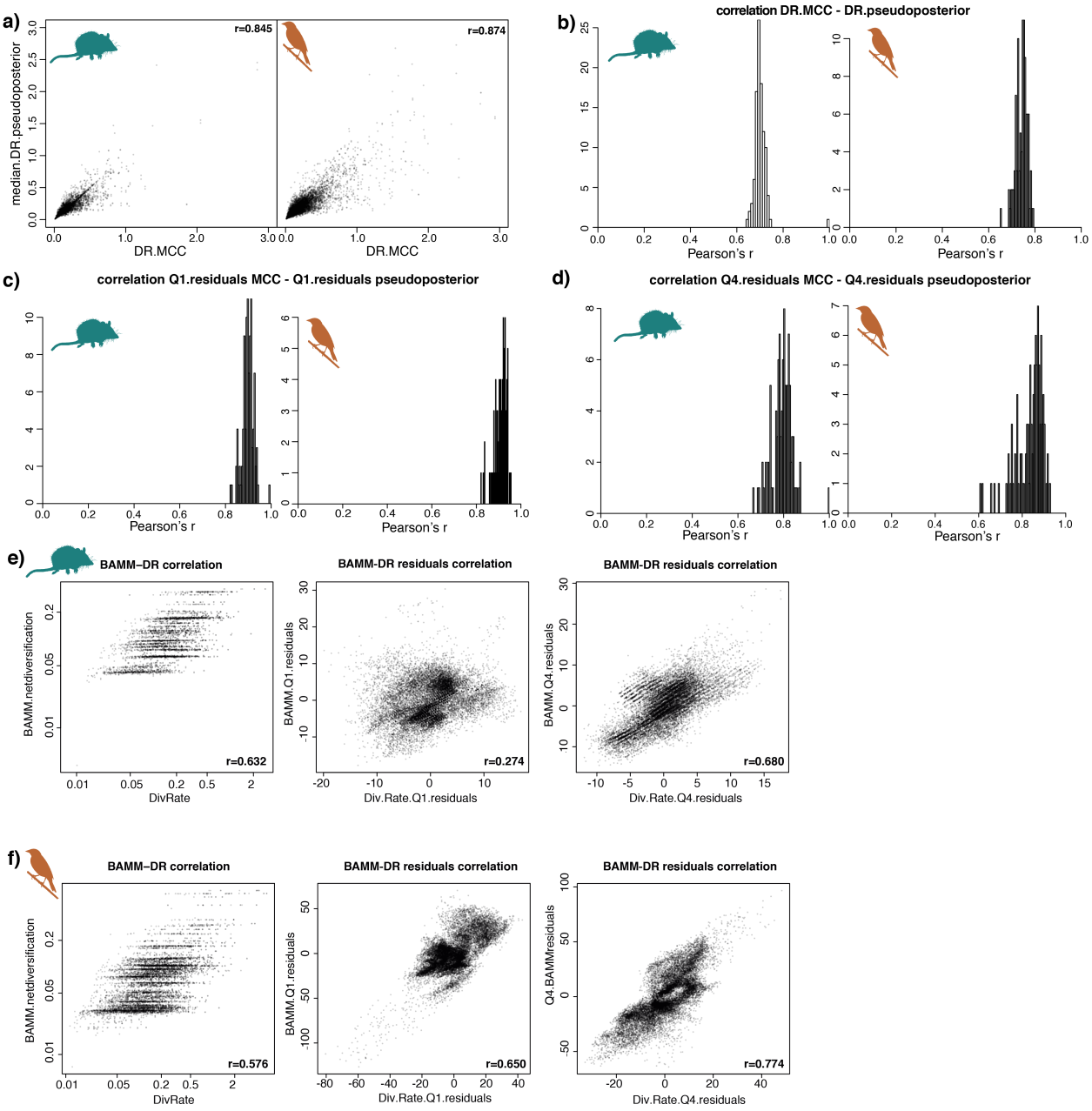


Figure S2. DR estimates are correlated across the pseudoposterior distribution and also correlate with BAMM estimates. a) correlation of DR values from the MCC tree with the median values of DR across the 100 random trees from the pseudoposterior distribution; b) correlations of the DR values from the MCC with each of the 100 random trees from the pseudoposterior; c) correlations of the cell-specific ancient lineages residuals in the MCC and the 100 random trees; d) correlations of the cell-specific recent lineages residuals in the MCC and

the 100 random trees; e) (left to right) correlation of the species-specific DR values in the MCC tree and the species-specific values of net diversification calculated with BAMM; correlation of the cell-specific ancient lineages residuals calculated with DR and with BAMM; correlation of the cell-specific recent lineages residuals calculated with DR and with BAMM in mammals; and f) in birds.

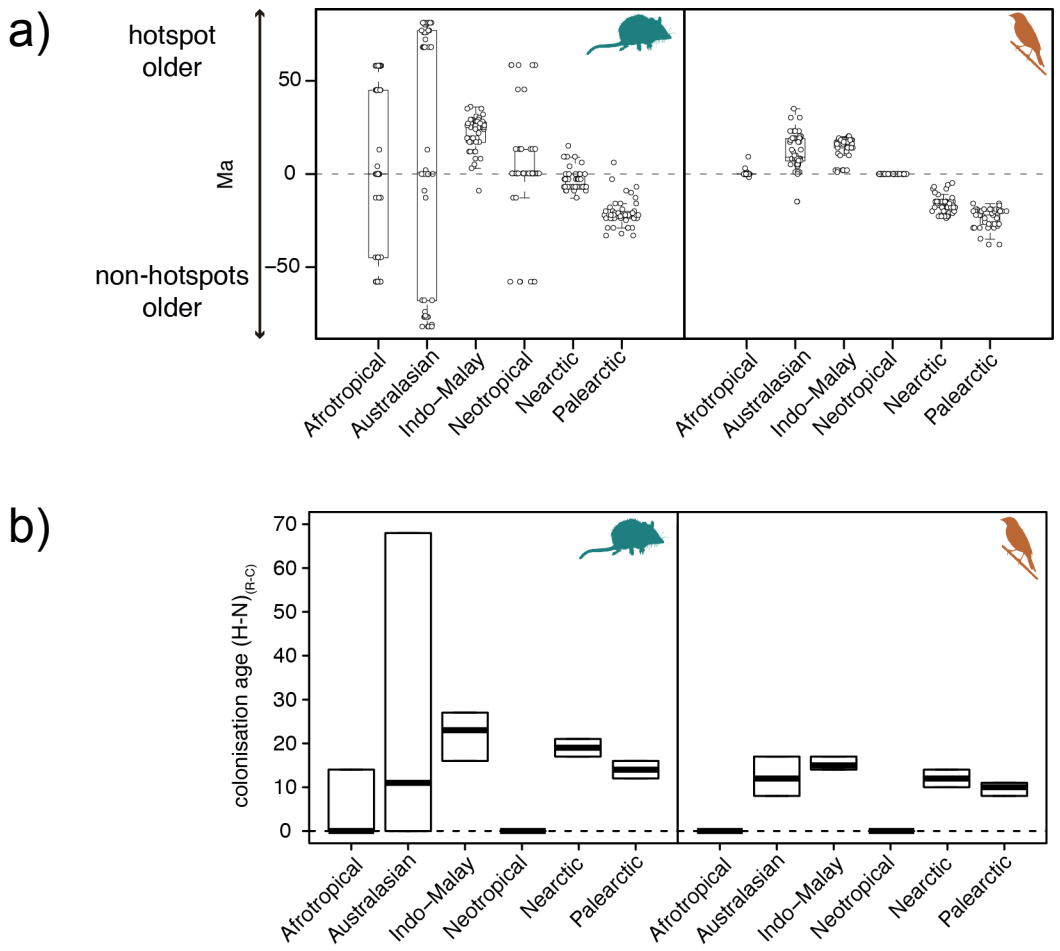


Figure S3. Age of colonization in hotspots and non-hotspots. a) No consistent differences in colonization age between hotspot (H) and non-hotspot (N) regions across realms. Points represent each of the 50 biogeographic stochastic mappings (BSMs) in each realm. Colonization ages in the 50 BSMs were calculated as the first presence of a lineage in a region. Box limits indicate the first and third quartiles and whiskers indicate 1.5 times the interquartile range above the third quartile and below the first quartile. Differences overlapped zero (shown by dashed line) for most tropical realms, whereas non-hotspots were generally older later than non-hotspot regions, contrary to the prediction that greater diversity can arise from an older colonization date. b) Empirically-estimated differences in the age of colonization of hotspots and non-hotspots differ from rates estimated in "control areas" with similar size and spatial structure to the real

hotspots. Differences in the age of colonization between hotspots and non-hotspots (H – N) calculated with the real dataset (R) minus the difference in the simulated "control" (C) hotspots within realms. The difference in the "control" hotspots was calculated for each biogeographic stochastic mapping by subtracting the estimated age of colonization of the non-hotspot region from the estimated age of colonization of the hotspot region (n = 50 replicates × 50 control regions). Solid line shows the median, box limits indicate the first and third quartiles.

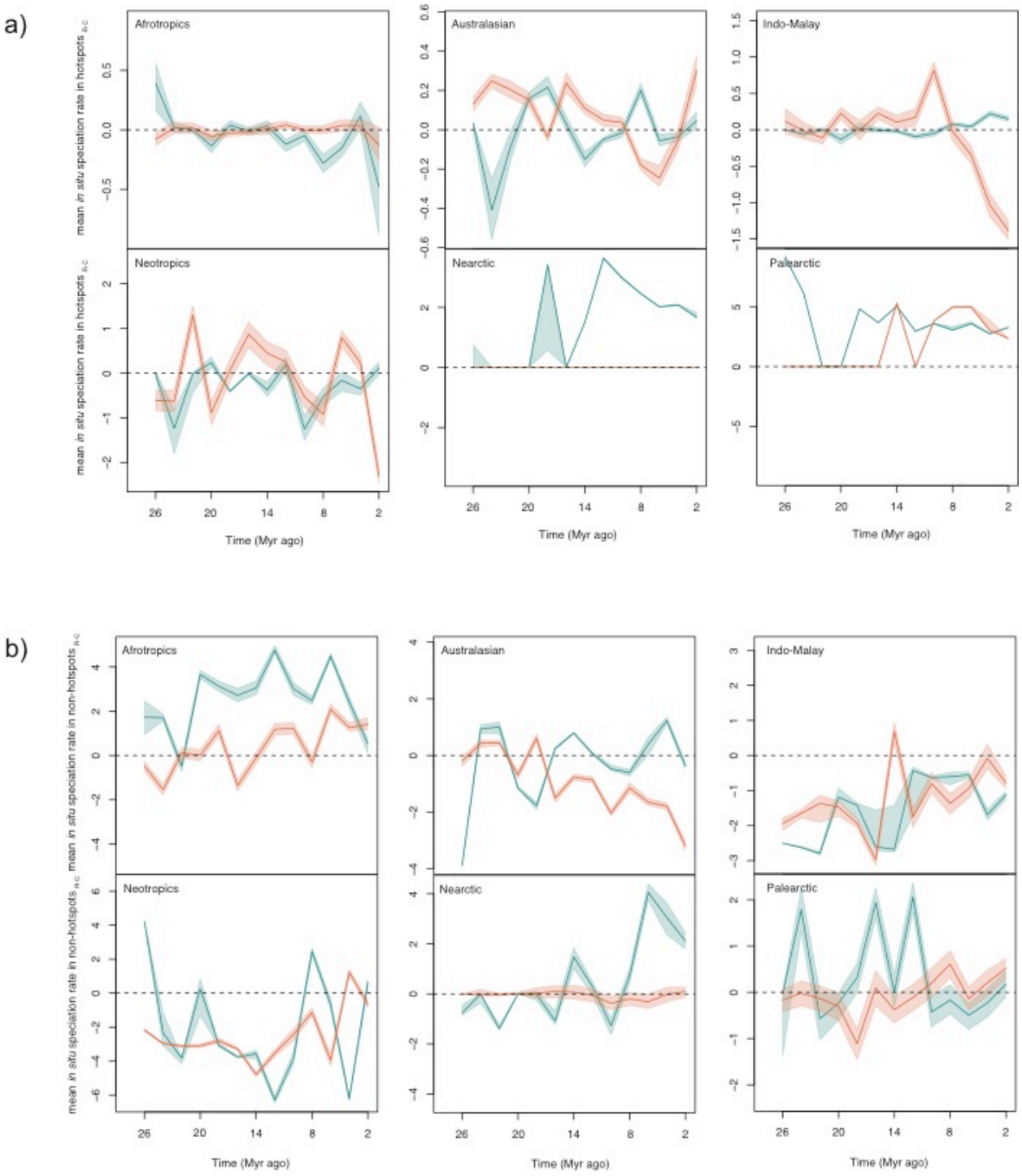


Figure S4. Empirically-estimated *in situ* cladogenetic rates in hotspots and non-hotspots differ from rates estimated in "control areas" with similar size and spatial structure to the

real hotspots. Differences for *in situ* cladogenetic rates in the hotspots (a) and in the non-hotspots (b) calculated with the real (R) dataset minus the simulated "control" (C) hotspots within each realm. Solid lines indicate the median difference in the *in situ* rates at a particular time bin and the shaded areas indicate the 90% confidence interval. Intervals overlapping the dotted line indicate a lack of statistically significant differences at $\alpha = 0.10$. The differences have been standardized by dividing the values by the realm-specific standard deviations.

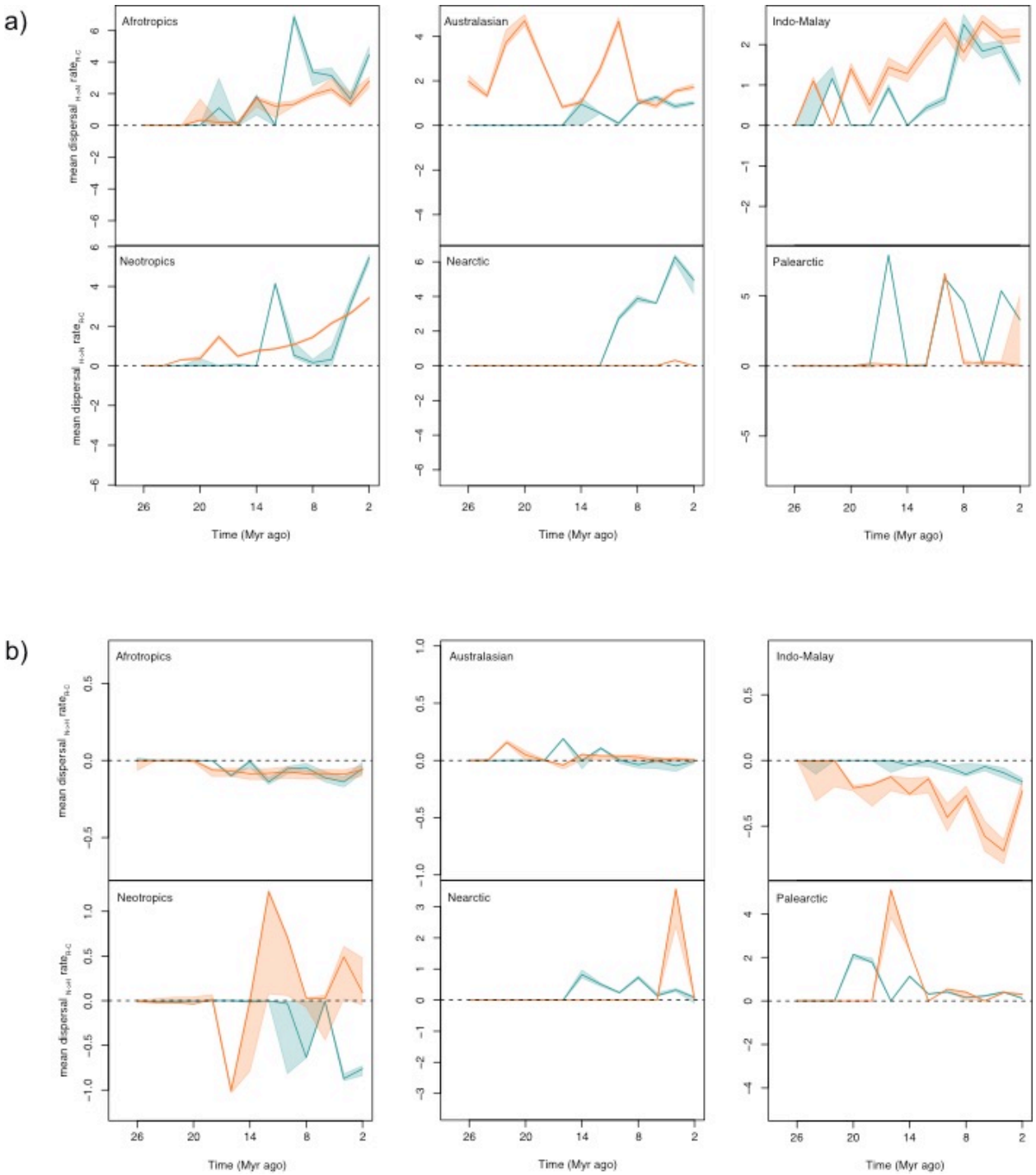


Figure S5. Empirically-estimated dispersal rates from hotspots to non-hotspots (H→N) and from non-hotspots to hotspots (N→H) differ from rates estimated in "control areas" with similar size and spatial structure to the real hotspots. Differences for dispersal rates from

hotspots to non-hotspots (H→N) (a) and from non-hotspots to hotspots (N→H) (b) calculated with the real (R) dataset minus the simulated "control" (C) hotspots within realms. Lines and shaded areas are presented as in Figure S4.

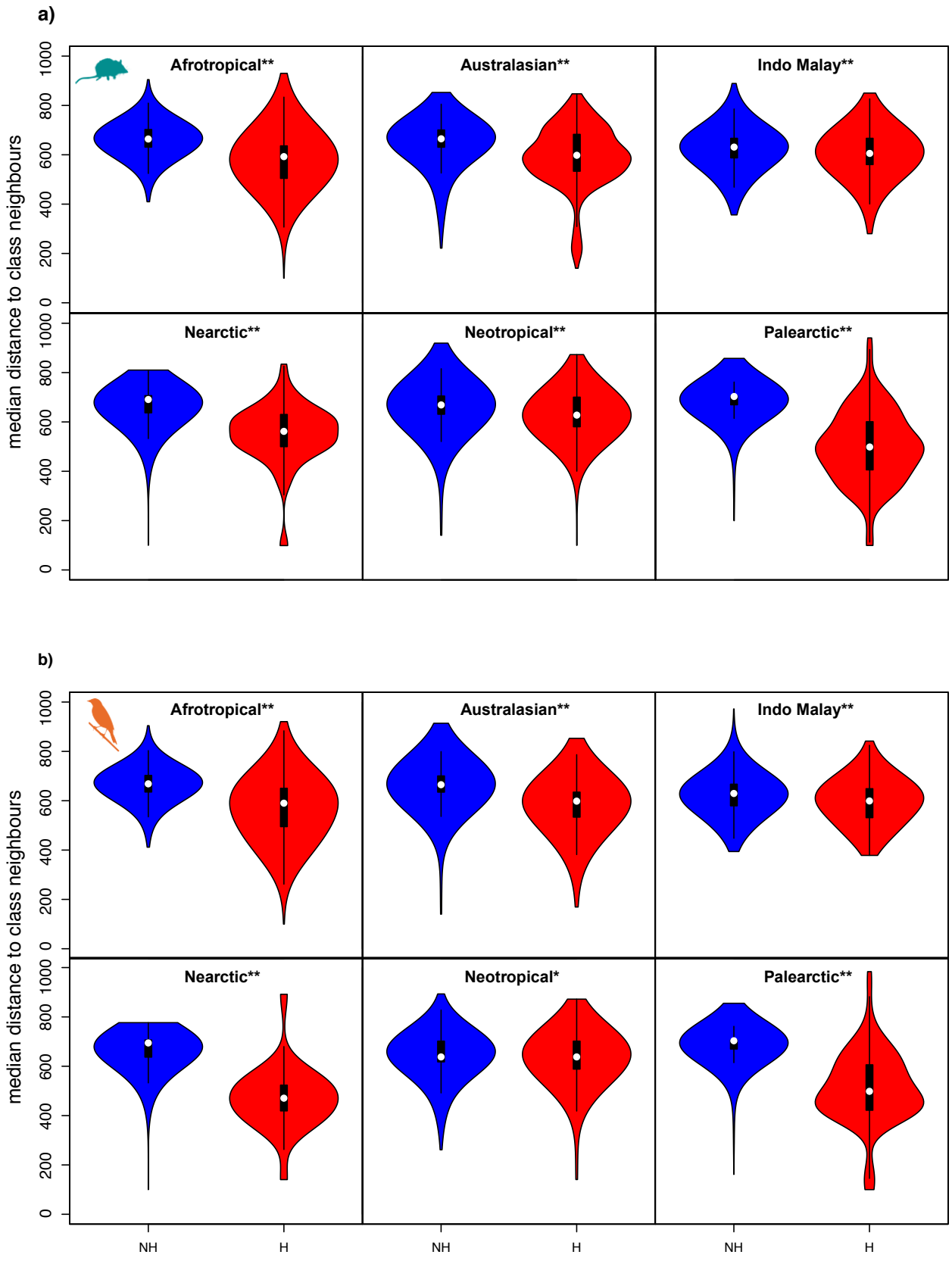

a)
Afrotropical**
Australasian**
Indo Malay**
Nearctic**
Neotropical**
Palearctic**
median distance to class neighbours
0
200
400
600
800
1000
b)
Afrotropical**
Australasian**
Indo Malay**
Nearctic**
Neotropical*
Palearctic**
median distance to class neighbours
NH
H

Figure S6. Similar differences in contiguity of hotspot and non-hotspot cells across biogeographic realms. For each realm, the median distance of each cell to every neighboring cell of the same class in a 1000 km radius is shown for hotspots (red) and non-hotspots (blue) for a) mammals, and b) birds. Asterisks indicate significant differences (* = p-value < 0.05; ** = p-value <0.01) between hotspot and non-hotspot contiguity with a *t*-test.

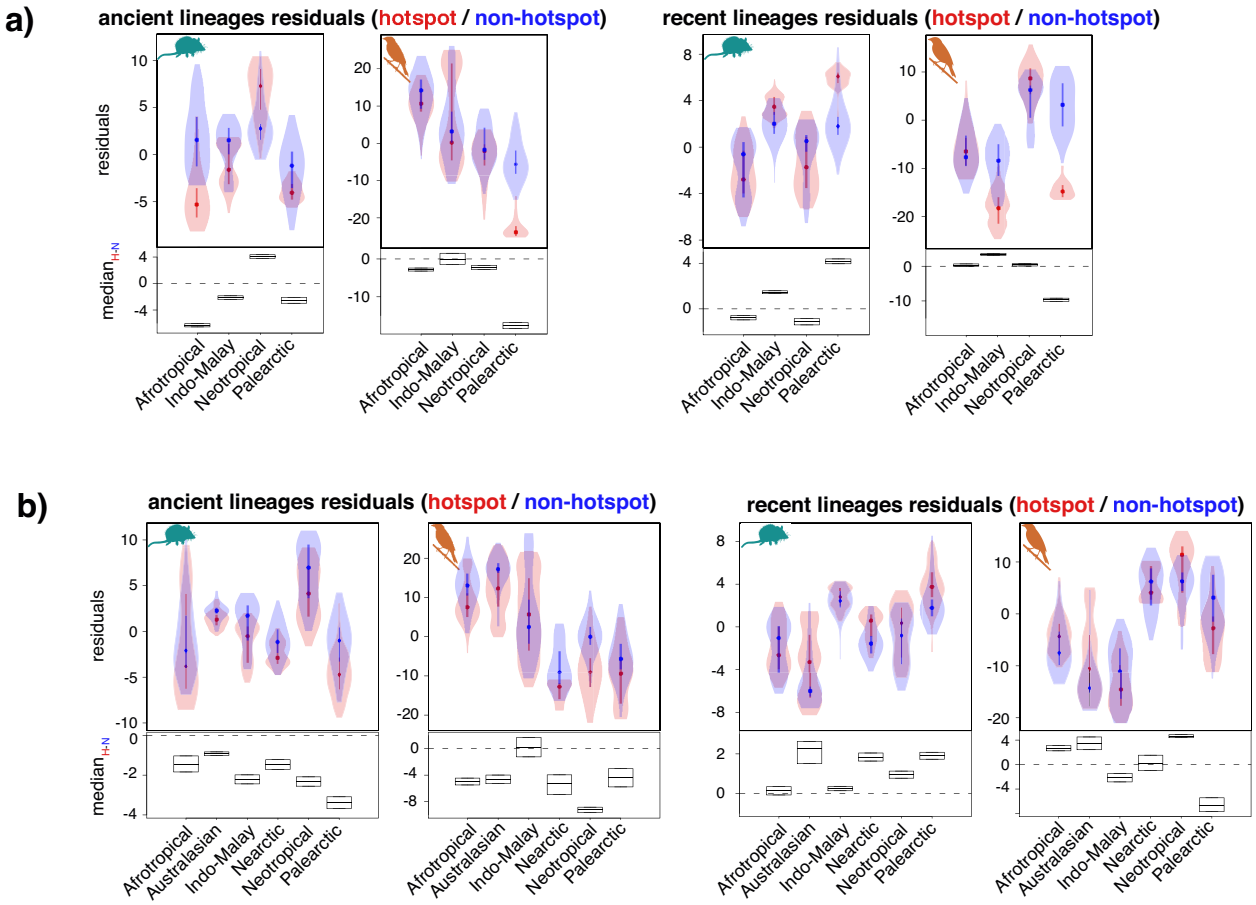
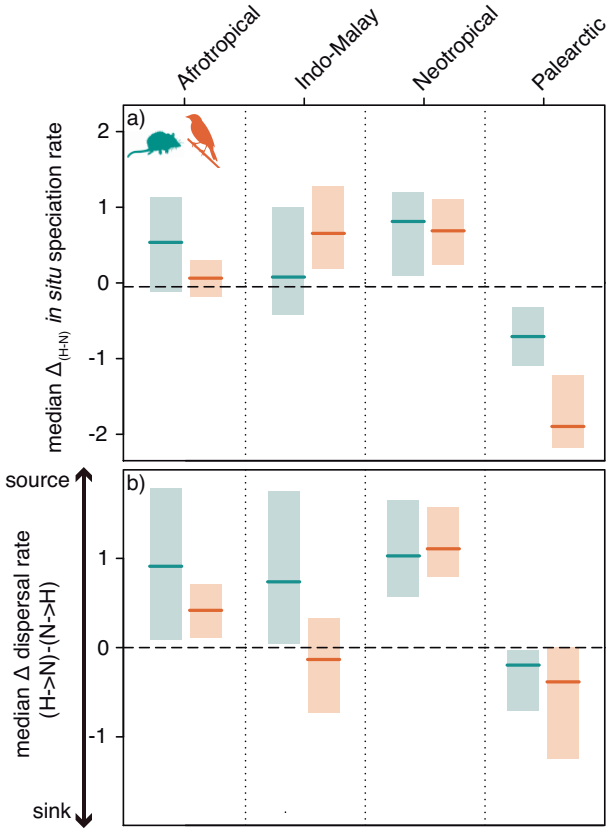


Fig. S7. Species richness-based hotspots (a) and narrow ranged species-based hotspots (b) are poor in ancient lineages and sometimes rich in recent lineages. Residuals from linear models predicting cell-specific richness of ancient and recent lineages in hotspots (H, shown in red) and non-hotspot regions (N, shown in blue). Positive residuals indicate a regional excess of ancient/recent lineages and negative residuals indicate a deficit. Lower panels are presented as in Fig. 1.

a)


Afrotropical
Indo-Malay
Neotropical
Palearctic
a)
median Δ(H+N) in situ speciation rate
2
1
0
-1
-2
b)
source
sink
median Δ dispersal rate
(H->N)-(N->H)
1
0
-1

b)

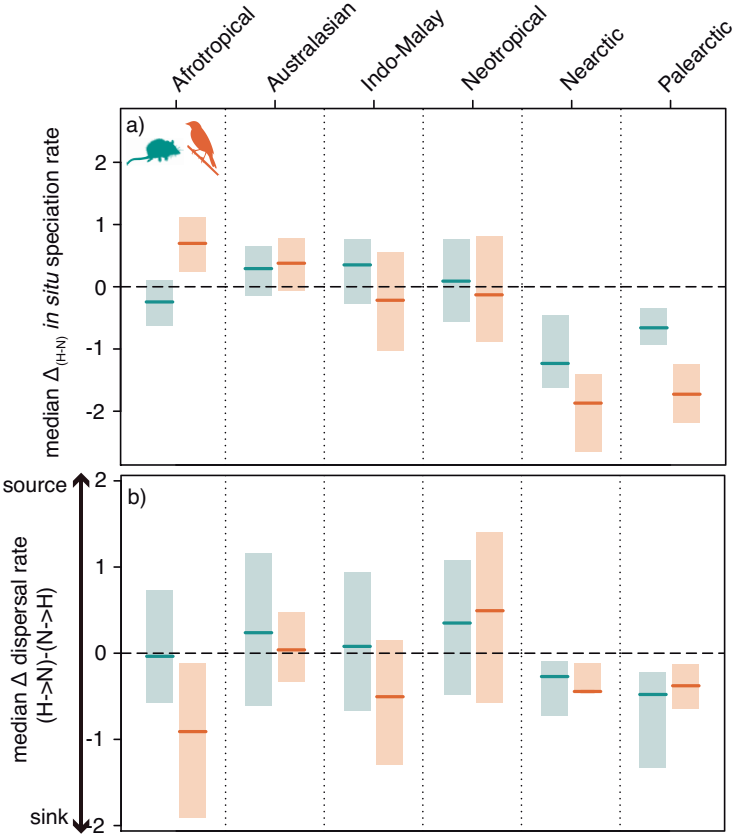

Afrotropical
Australasian
Indo-Malay
Neotropical
Nearctic
Palearctic
a)
median Δ(H+N) in situ speciation rate
2
1
0
-1
-2
b)
source
sink
median Δ dispersal rate
(H->N)-(N->H)
2
1
0
-1
-2

Fig. S8. Contrasting macroevolutionary routes in species richness-based hotspots and non-hotspots (a) and in narrow ranged species-based hotspots and non-hotspots (b) *In situ* cladogenesis rates between 2 to 26 Ma ago within non-hotspots were subtracted from rates within hotspots in each of six biogeographic realms (top panels), and divided by the overall standard deviation to allow for comparison across realms. Dispersal rates between 2 to 26 Ma ago from non-hotspots to hotspots (N→H) were subtracted from hotspot to non-hotspots (H→N) rates within each realm (bottom panels). Lines and shades are presented as in Fig. 2

Figure S9. Example of simulating 'control' hotspots. Map of hotspots (shown in red) in the Afrotropics (observations outlined in grey) and 50 sets of cells with similar size and spatial struct

Table S1. DR and BAMM produce consistent differences between hotspot and non-hotspot regions. For each region, we subtracted the non-hotspot residuals from hotspots and compared the sign between the DR and BAMM approaches. A match indicates that the sign of the comparison is the same in DR and BAMM (e.g., there are less ancient lineages in the hotspots both with DR and BAMM estimates). A mismatch indicates that the sign of the comparison is not the same (e.g., there are less ancient lineages in the hotspots if estimated with DR but more ancient lineages if estimated with BAMM).

	MAMMALS		BIRDS	
	Ancient lineages residuals	Recent lineages residuals	Ancient lineages residuals	Recent lineages residuals
MATCH	4	4	2	4
NO MATCH	2	1*	2*	1*

* indicates nonsensical comparison(s) (i.e., when one of the values in a comparison overlaps zero, so we cannot tell if it is larger or smaller in the hotspots)

Table S2. Total size and proportion of hotspot cells across biogeographic realms.

Realm name	Size (# of cells)	Proportion of hotspots
Afrotropical	2512	0.29
Australasian	1337	0.32
Indo-Malay	1255	0.50
Nearctic	2753	0.08
Neotropical	2304	0.46
Palearctic	6277	0.03

Table S3. Mean of median distances (km) of each cell to every neighboring cell of the same class in a 1000 km radius is shown for hotspots and non-hotspots for mammals and birds.

	mammals		birds	
	Distance to non-hotspot	Distance to hotspot	Distance to non-hotspot	Distance to hotspot
Afrotropical	661.34	584.04	666.37	567.99
Australasian	651.95	591.99	666.74	581.84
Indo-Malay	623.76	605.12	621.29	596.61
Nearctic	663.52	550.97	665.32	470.09
Neotropical	661.37	628.61	643.00	634.51
Palearctic	685.80	496.45	686.73	509.50

5

Table S4. Overlap of weighted endemicity-based (WE) hotspots with species richness-based (SR) and narrow-ranged species-based (NRS) hotspots. Size is the number of grid cells.

	mammals					birds				
	WE hotspots (size)	SR hotspots (size)	overlap (proportion WE ∩ SR)	NR hotspots (size)	overlap (proportion WE ∩ NR)	WE hotspots (size)	SR hotspots (size)	overlap (proportion WE ∩ SR)	NR hotspots (size)	overlap (proportion WE ∩ NR)
Afrotropics	735	1129	0.73	549	0.64	651	1162	0.76	356	0.44
Australasian	433	23	0.05	391	0.77	555	113	0.20	428	0.72
Indo-Malay	629	472	0.68	390	0.58	623	433	0.65	351	0.48
Nearctic	186	23	0.00	177	0.64	27	0	0.00	46	0.48
Neotropical	1064	1531	0.75	844	0.68	1298	1504	0.76	706	0.49
Palearctic	206	92	0.30	358	0.73	98	75	0.61	92	0.38

Table S5. Model fit of BioGeoBEARS in 6 biogeographic realms. Best fitting models according to AICc are shaded.

	Afrotropics		**Australasian**		**Indo-Malay**		**Nearctic**		**Neotropics**		**Palearctic**	
model	mammals	birds	mammals	birds	mammals	birds	mammals	birds	mammals	birds	mammals	birds
DEC	14694	33464	12342	28771	12495	28997	12270	26217	14985	34529	12130	26797
DEC+J	13620	31236	11073	25744	11032	25860	10876	23127	14041	32325	10711	23727
DIVALIKE	15300	34723	12925	30040	12952	29984	12815	27430	15645	35945	12649	27990
DIVALIKE+J	13958	31994	11346	26362	11218	26312	11128	23630	14463	33214	10925	24237
BAYAREALIKE	15399	34012	13744	31789	14538	32939	13954	29596	14894	36114	14003	30516
BAYAREALIKE+J	12751	27798	11256	25874	11237	25958	11013	23217	12737	31418	11004	23895