

Supplementary methods

Distribution data was used to estimate the variable *species richness* for each bioregion. The addition of phylogenetic data enabled us to estimate the evolutionary, or direct, predictors: dispersal (*dispersal rate*), evolutionary rates (*speciation or diversification rates*), and evolutionary time (*maximum clade age*) (electronic supplementary material, figure S2). Evolutionary rates and evolutionary time were calculated for selected clades, i.e. clades for which all component species are found in the focal bioregion. Among the indirect variables, the two ecological variables were calculated as the mean temperature (*temperature*) and total productivity (*productivity*) of each bioregion. The third indirect variable, climatic stability (*area through time*) is the integration of the focal bioregion area over the past 55 My (electronic supplementary material, figure S2).

Distribution data

We obtained species range maps from the IUCN Red List of Threatened Species for amphibians and terrestrial mammals (1; see also 2, 3) and from BirdLife International for birds (4). Marine species were removed. For terrestrial 'reptiles' (crocodylians, squamate reptiles, and turtles), high quality distribution data derived from recent IUCN Red List workshops (e.g. 5) are available comprehensively for the Western Hemisphere. We facilitated most (nine) of these workshops with regional experts and assembled distributional data on 3,536 species into consistent formats (6) for the analyses. For Eastern Hemisphere 'reptiles' (crocodylians, squamate reptiles, turtles and one tuatara species) we used the distribution data available for species assessed for the IUCN Red List (2,439 species) and added the remaining 2,786 species to these from "Wildfinder" (7, 8), where species list by bioregion are available.

Phylogenetic data

We used smoothed timetrees with interpolated species (species without any genetic data but with at least one member of their genus in the phylogeny) of mammals (5,364 species of which 30% were interpolated) and birds (9,879 species of which 31% were interpolated) (9), and amphibians (7,340 species of which 59% were interpolated) (10). Interpolated species were added to the squamate reptile (9,378 species of which 55% were interpolated) and turtle (143 species of which 22% were interpolated) timetrees, which were subsequently smoothed, following the same protocol (birth-death polytomy resolution method) used for the other three tetrapod groups (9, 10). The smoothing correction was necessary to remove the artificial bias on diversification caused by polytomies and also because the program we used to estimate speciation and diversification rates (BAMM; 11) requires fully resolved timetrees with non-zero branch lengths. Those timetrees and the crocodylian timetree (23 species), for which no species could be

added, are part of a single timetree of life (9) compiled from 2,274 studies representing 50,632 species. The tetrapod timetree (32,128 species) is composed of six clades with interpolated species (amphibians, birds, crocodylians, mammals, squamate reptiles, and turtles) and one tuatara species. We did not consider 'reptiles' as a clade because this group is not monophyletic, but rather considered crocodylians, squamate reptiles, turtles, and the tuatara species separately. Not all species could be interpolated because not all genera have been sequenced. Thus, the correspondence between our timetrees and the IUCN maps led to the use of a slightly smaller number of tetrapods (27,412 species representing 84.2 % of described species (1, 12)), amphibians (6,196 species, 94.9 % of described species (1)), birds (9,210 species, 88.3 % of described species (1)), mammals (4,980 species, 89.9 % of described species (1)), and squamate reptiles (6,951 species, 69.0 % of described species (12)). Geographic and phylogenetic data for crocodylians and turtles were available for only 75 species. Therefore, we included those crocodylian and turtle species, and the tuatara species, in tetrapod-level analyses but did not conduct separate analyses on them.

Our analyses concerning the relative influence of evolution, ecology, and climatic stability on species richness were conducted on a global scale. Analyses were performed in R 3.3.0 (13). The "maptools" (14), "raster" (15), and "rgdal" (16) packages were used to derive assemblage species lists for each bioregion described in Olson et al. (17) (electronic supplementary material, figure S1) for all tetrapods combined and each of the four main groups (birds, mammals, amphibians, and squamate reptiles) separately (electronic supplementary material, table S7).

Evolutionary hypotheses

To estimate clade age for each bioregion, we first determined the clades for which all the species occurred in each bioregion based on the species lists established by bioregion (electronic supplementary material, table S7), and recorded the *maximum clade age* (electronic supplementary material, table S8) representing the time during which species have successfully evolved and diversified in a given bioregion (electronic supplementary material, figure S2). To detect potential outliers, we compared the mean and maximum clade age for each bioregion and taxon studied. To take into account the uncertainty in phylogenetic reconstructions, we tested the lower and upper bound of the 95% confidence interval (9) of the node time used as the *maximum clade age*, termed *age lower bound* (model 3a), *age upper bound* (model 3b) and *random lower/upper age* (model 3c) corresponding to the lower or upper bound randomly sampled. We also tested *cumulative clade age*, i.e. the sum of the clade ages, which represents the total evolutionary time in a given bioregion (model 3d).

Clade age might influence richness indirectly, with older clades expected in richer areas. We used simulations to test the possibility of a spurious correlation between *maximum clade age* and *richness*. A timetree of 32,128 tips

(size of the tetrapod tree) was simulated under a variable birth-death model using five different diversification rates from the range of tetrapod diversification rates. We sampled X clades, where X was sampled from the range of tetrapod number of clades. *Richness* and *maximum clade age* were calculated for the X clades. We repeated this step 100 times, as if there were 100 bioregions and calculated the correlation between *maximum clade age* and *richness* among the 100 regions. These steps were repeated 1000 times to obtain a distribution of the correlation coefficients. We compared our estimates to the distribution (electronic supplementary material, figure S6). We found a correlation between *maximum clade age* and *richness*, but the observed values are higher than expected by chance.

Extinction rate estimates are not reliable in phylogenies that have diversification rate heterogeneity (18, 19). For this reason, we primarily used the mean of the clade speciation rates by bioregion as the variable *speciation rate* (electronic supplementary material, table S8 and figure S2). As for the estimation of *maximum clade age*, the clades used to calculate the rate variables are the clades for which all component species are found in the focal bioregion. We extracted the speciation (or diversification) rate of each clade (see below) corresponding to the integration of rate with respect to time along each branch in the clade. The clade rates were averaged to obtain the variable *speciation rate* for each bioregion. We also tested the influence of mean diversification rate (speciation minus extinction) as well as median speciation and median diversification rates in explaining species richness variation. To estimate evolutionary rates (speciation and net diversification) we used the BAMMtools package (11) on the smoothed trees with interpolated species of amphibians, birds, crocodylians, mammals, squamate reptiles, and turtles (9). The tuatara species was not captured in any of the clades and therefore we did not estimate evolutionary rates for this taxon. The aim of the program BAMM (Bayesian analysis of macroevolutionary mixtures) is to model speciation and extinction dynamics over time and between lineages using reversible jump Markov chain Monte Carlo (MCMC). The function 'setBAMMpriors' was used to generate a prior block that matched the "scale" (e.g., depth of the tree) of our data. Both the speciation and the extinction rate were allowed to vary through time and across lineages, and 50,000,000 generations of MCMC simulation were performed. A global sampling fraction (number of species in the timetree over the number of described species) was specified by setting the 'globalSamplingFraction' parameter for each timetree (0.99, 0.99, 0.92, 0.97, 0.94, and 0.42 for amphibians, birds, crocodylians, mammals, squamate reptiles, and turtles respectively). We applied a burn in of 0.5 and checked the convergence by calculating the effective sample size of the log-likelihood and of the number of shift events present in each sample that should be over 200. Although the accuracy BAMM has been questioned (20), the author of BAMM replied by pointing out that the problem was generated by inappropriate use of a hidden ('developer-only') setting (21). Under the default option in BAMM, the posterior is well-behaved with respect to the prior, even in variable rate phylogenies (21). We have used the default

option in this study. In a previous study (9), we showed that rate was constant at the eukaryote scale. However because smaller taxonomic groups, such as birds and mammals, showed variable rates through time, speciation (and diversification) rates could be predictors of species richness pattern. Moreover, beside time, speciation (and diversification) rates were significantly correlated with species richness in birds, mammals, and squamates, showing that we were able to pick up a signal across bioregions.

Dispersal rates were estimated with the program GeoSSE (Geographic State Speciation and Extinction) (22) (electronic supplementary material, figure S2). This program assumes 3 states, A (species endemic to region A), B (species endemic to region B), and AB (species widespread in both regions). It allows the estimation of dispersal (d_A and d_B , range expansion from A to AB and B to AB respectively), speciation (s_A , s_B , and s_{AB}), and extinction (x_A and x_B) rates for two different geographic areas A and B, with s_{AB} being the between-region speciation rate (i.e. speciation rate of species distributed on both areas). For each bioregion (compared to the remaining bioregions) and taxonomic group, nineteen models (electronic supplementary material, table S9) were compared using the Akaike Information Criterion (AIC). We checked support for the selected model against all other models nested within it using the likelihood ratio test ($p < 0.05$). We first computed a pilot run to obtain reasonable step sizes, then a Bayesian Markov chain Monte Carlo (MCMC) analysis (function 'mcmc') with the set of parameters selected by AIC and 10,000 steps chains of which 2,500 steps were discarded. We checked the convergence by calculating the effective sample size of the log-likelihood and of the number of shift events present in each sample that should be over 200. The estimate d_B representing the dispersal rate toward area A, the bioregion of interest, was used as the *dispersal rate* variable. This program requires at least one species endemic to region A or B, and at least one species to occur in both regions (region AB). This was not the case for a few bioregions, therefore the variable *dispersal rate* could not be estimated for 2, 7, 10, 4, and 5 bioregions for tetrapods, amphibians, birds, mammals, and squamates respectively.

Some concerns have been recently raised concerning the spurious trait dependency of rate that can be inferred by the binary state-speciation and extinction models (23, 24). Davis et al. (23) pointed out that the program BiSSE (25) led to an incorrect result in the case of low sampling or highly unequal representation of one of the two states. Because GeoSSE and BiSSE use the same mathematical framework (22), they might be affected by the same bias. In this study we used large phylogenies (7,990 species in average), although one region was always over-represented with an average ratio of 0.024 (species endemic to the area B / species endemic to the region A). To test the accuracy of the speciation and dispersal rate estimations we simulated (function 'tree.geosse') trees of 5,000 and 9,000 tips, corresponding to the range of tree sizes of the four main tetrapod groups with specific speciation and dispersal rates. In the first set of simulations the parameters s_A , s_B , and s_{AB} were set at 1, x_A and x_B were set at 0.5, d_A

was set at 60, and d_b was set at 1 in order to produce trees with a region ratio (A/B) close to the average region ratio observed in our data. In the second set of simulations the parameter s_b was set at 1.5, s_A and s_{AB} were set at 1, x_A and x_B were set at 0.5, d_A was set at 60, and d_b was set at 1 in order to produce trees with a similar tip ratio as in the first set of simulations but different speciation rates in region A and B. Ten trees were simulated for each set of simulations and tree size and the method described above was followed to estimate s_A , s_B , s_{AB} , x_A , x_B , d_A , and d_b . According to our results, the estimated values, and particularly d_b (the *dispersal rate* variable), were similar to the simulated values (electronic supplementary material, table S10), validating our use of GeosSE. We also compared the speciation estimates obtained with the program BAMM and the program GeosSE and obtained similar results for amphibians ($r^2=0.77$; $p\text{-value}=4.13e^{-09}$), birds ($r^2=0.83$; $p\text{-value}=1.95e^{-09}$), mammals ($r^2=0.74$; $p\text{-value}=2.54e^{-09}$), and squamates ($r^2=0.80$; $p\text{-value}=1.1e^{-10}$).

Ecological hypotheses

We evaluated the two major components of the ‘energy hypothesis’, *productivity*, and *temperature*. For *productivity*, we estimated the total productivity of a region with the cumulative vegetation production index (DHI1) calculated globally at a 1 km resolution from MODIS fPAR data (<http://silvis.forest.wisc.edu/dhi>) using a recent derivation of the Dynamic Habitat Index (26) (electronic supplementary material, figure S2). For *temperature*, we estimated the mean temperature of a region with 2.5 arc-minute raster data for annual mean temperature (BIO1) derived from WorldClim (27). Using the R packages “raster” (15) and “rgdal” (16), raster data for productivity and mean temperature were aggregated at $1/4^\circ$ resolution and the mean temperature and the total productivity of each bioregion were used in our analyses (electronic supplementary material, table S8 and figure S2).

Climatic stability

Climatic stability of bioregions was estimated as the integration of bioregion area over the past 55 My, also termed *TimeArea* (28, 29) (electronic supplementary material, figure S2). This variable represents the area of a bioregion (with specific climatic conditions) through time. If a bioregion has been impacted by climatic fluctuations, its area through time will vary to reflect the range contractions and expansions of the particular climatic conditions characterizing the focal bioregion. We used this variable as our estimate of climatic stability, which we describe here as *area through time*. The last 55 million years appeared to be an appropriate interval of time because by 55 Ma the biosphere has recovered from the massive extinction that occurred 65 Ma, with a very different fauna and flora than during the

Cretaceous (28). Moreover ancestors of the current dominant vertebrate and plant lineages composed the Eocene biota (30).

REFERENCES

1. IUCN. 2016 IUCN Redlist of Threatened Species. Available at: <http://www.iucnredlist.org>. Last accessed 08 January 2016.
2. Stuart SN, Chanson JS, Cox NA, Young BE, Rodrigues AS, Fischman DL, Waller RW. 2004 Status and trends of amphibian declines and extinctions worldwide. *Science* 306, 1783–86.
3. Schipper J, Chanson JS, Chiozza F, Cox NA, Hoffmann M, Katariya V, Lamoreux J, Rodrigues AS, Stuart SN, Temple HJ, Baillie J. 2008 The status of the world's land and marine mammals: diversity, threat, and knowledge. *Science* 322, 225–230.
4. BirdLife International, NatureServe. 2016 Bird species distribution maps of the world. BirdLife International/NatureServe, Cambridge, UK/Arlington, USA.
5. Young B. 2012 Red listing central American squamates. *Herpetological Review* 43, 368–70.
6. IUCN. 2013 Documentation standards and consistency checks for IUCN Red List assessments and species accounts. Version 2. Adopted by the IUCN Red List Committee and IUCN SSC Steering Committee. Downloadable from: http://www.iucnredlist.org/documents/RL_Standards_Consistency.pdf.
7. Lamoreux JF, Morrison JC, Ricketts TH, Olson DM, Dinerstein E, McKnight MW, Shugart HH. 2006 Global tests of biodiversity concordance and the importance of endemism. *Nature* 440, 212–4.
8. World Wildlife Fund. 2006 WildFinder: Online database of species distributions. Available at: www.worldwildlife.org/WildFinder (accessed 18 January 2016).
9. Hedges SB, Marin J, Suleski M, Paymer M, Kumar S. 2015 Tree of life reveals clock-like speciation and diversification. *Molecular Biology and Evolution* 32, 835–45.
10. Marin J, Hedges SB. 2016 Time best explains global variation in species richness of amphibians, birds and mammals. *Journal of Biogeography* 43, 1069–79.
11. Rabosky DL, Grundler M, Anderson C, Shi JJ, Brown JW, Huang H, Larson JG. 2014 BAMMtools: an R package for the analysis of evolutionary dynamics on phylogenetic trees. *Methods in Ecology and Evolution* 5, 701–7.
12. Uetz P, Freed P, Hošek J. 2016 The reptile database. Available at: <http://www.reptile-database.org> (accessed 10 November 2016).

13. R Core Team. 2016 R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.
14. Bivand R, Lewin-Koh N. 2014 Maptools: tools for reading and handling spatial objects. R package version 0.8–29. Comprehensive R Archive Network, Vienna, Austria.
15. Hijmans RJ, van Etten J. 2014 Raster: geographic data analysis and modeling. R package version 2–2. Comprehensive R Archive Network, Vienna, Austria.
16. Bivand R, Keitt T, Rowlingson B, Pebesma E, Sumner M. 2013 Rgdal: bindings for the geospatial data abstraction library. R package version 0.8–9. Comprehensive R Archive Network, Vienna, Austria.
17. Olson DM, Dinerstein E, Wikramanayake ED, Burgess ND, Powell GV, Underwood EC, D'amico JA, Itoua I, Strand HE, Morrison JC, Loucks CJ. 2001 Terrestrial ecoregions of the world: a new map of life on earth. *BioScience* 51: 933–38.
18. Rabosky DL. 2010 Extinction rates should not be estimated from molecular phylogenies. *Evolution* 64, 1816–24.
19. Rabosky DL. 2016 Challenges in the estimation of extinction from molecular phylogenies: A response to Beaulieu and O'Meara. *Evolution* 70, 218–28.
20. Moore BR, Höhna S, May MR, Rannala B, Huelsenbeck JP. 2016 Critically evaluating the theory and performance of Bayesian analysis of macroevolutionary mixtures. *Proceedings of the National Academy of Sciences of the United States of America* 113, 9569–74.
21. Rabosky DL, Mitchell JS, Chang J. Is BAMM flawed? 2017 Theoretical and practical concerns in the analysis of multi-rate diversification models. *Systematic Biology* 66, 477–98.
22. Goldberg EE, Lancaster LT, Ree RH. 2011 Phylogenetic inference of reciprocal effects between geographic range evolution and diversification. *Systematic Biology* 60, 451–465.
23. Davis MP, Midford PE, Maddison W. 2013 Exploring power and parameter estimation of the BiSSE method for analyzing species diversification. *BMC Evolutionary Biology* 13, 38.
24. Rabosky DL, Goldberg EE. 2015 Model inadequacy and mistaken inferences of trait-dependent speciation. *Systematic Biology* 64, 340–55.
25. Maddison WP, Midford PE, Otto SP. 2007 Estimating a binary character's effect on speciation and extinction. *Systematic Biology* 56, 701–10.
26. Coops NC, Wulder MA, Iwanicka D. 2009 Demonstration of a satellite-based index to monitor habitat at continental-scales. *Ecological Indicators* 9, 948–58.
27. Hijmans RJ, Cameron S, Parra J, Jones PG, Jarvis A. 2005 WorldClim, version 1.3. University of California, Berkeley.

28. Jetz W, Fine PVA. 2012 Global gradients in vertebrate diversity predicted by historical area-productivity dynamics and contemporary environment. *PLoS biology* 10, e1001292.
29. Fine PV, Ree RH. 2006 Evidence for a time-integrated species-area effect on the latitudinal gradient in tree diversity. *The American Naturalist* 168, 796–804.
30. Carroll R. 1988 *Vertebrate paleontology and evolution*. New York: W.H. Freeman and Compagny.