LETTER

Universal species–area and endemics–area relationships at continental scales

David Storch[1,2], Petr Keil[3] & Walter Jetz[3]

Despite the broad conceptual and applied relevance of how the number of species or endemics changes with area (the species–area and endemics–area relationships (SAR and EAR)), our understanding of universality and pervasiveness of these patterns across taxa and regions has remained limited. The SAR has traditionally been approximated by a power law[1], but recent theories predict a triphasic SAR in logarithmic space, characterized by steeper increases in species richness at both small and large spatial scales[2–6]. Here we uncover such universally upward accelerating SARs for amphibians, birds and mammals across the world's major landmasses. Although apparently taxon-specific and continent-specific, all curves collapse into one universal function after the area is rescaled by using the mean range sizes of taxa within continents. In addition, all EARs approximately follow a power law with a slope close to 1, indicating that for most spatial scales there is roughly proportional species extinction with area loss. These patterns can be predicted by a simulation model based on the random placement of contiguous ranges within a domain. The universality of SARs and EARs after rescaling implies that both total and endemic species richness within an area, and also their rate of change with area, can be estimated by using only the knowledge of mean geographic range size in the region and mean species richness at one spatial scale.

The scale dependence of species richness has implications for all biodiversity patterns[1]. The SAR has been used to extrapolate species richness across spatial scales and also to estimate species extinctions after habitat loss[7,8] (but see refs 9, 10), typically relying on its particular universal properties. However, the universality of the shape of the SAR has been questioned[11]. The nested SARs (in which smaller sample areas are located within larger ones) are classically described as a power law across most spatial scales[1,12], but current theoretical approaches predict that species richness first increases steeply with area at a decelerating rate, then increases roughly linearly in logarithmic space, and accelerates upwards again when sample areas approach the size of individual species' geographic ranges[2–6]. In contrast to the well-documented curvature over small areas[13,14], data availability has so far hindered generalizations about the SAR at large scales. The EAR is the relationship between the area of a region and the number of species restricted (that is, endemic) to it. The EAR provides information on the number of species that may go extinct if parts of the area are destroyed or transformed[9,15,16], because being endemic to the area would imply that any local extinction is also global. Despite the potential of the EAR in biodiversity science and conservation[9,15,17], its empirical shape at biogeographic scales has remained largely undocumented. The slope of the EAR at smaller spatial scales is expected to be connected to the slope of the SAR at large scales, because an increase in species richness with increasing study plot area (the SAR) corresponds to a decrease in the number of species that are restricted (that is, endemic) to the remaining area[16] (that is, the area not included in the study plot; see Supplementary Discussion and Supplementary Figs 1 and 2).

Here we provide a construction of fully nested continental SARs and EARs for all amphibians, birds and mammals (see refs 18 and 19 for

data description and validation, and Methods and Supplementary Table 1 for details). SARs for all continents and taxa accelerate upward in log–log space (Fig. 1a–c). Differences in the SAR position along the y axis correspond to known differences in total species richness of individual continents and taxa[20,21]; for example, birds have consistently
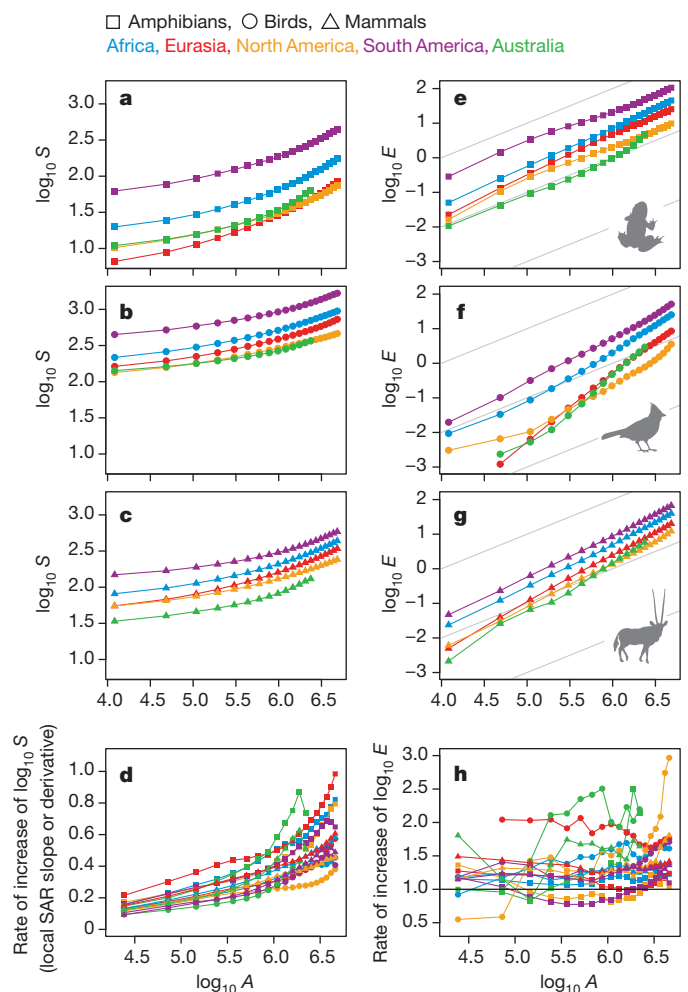

Figure 1 | SARs and EARs across five continents and three vertebrate classes. a–c, e–g, The SARs for amphibians (**a**), birds (**b**) and mammals (**c**) reveal an upward-accelerating shape for logarithmic axes, whereas EARs for amphibians (**e**), birds (**f**) and mammals (**g**) are more or less linear. **d, h,** Confirmation by plotting the local slopes (derivatives) of the relationships for each continent (**d** for SARs and **h** for EARs). All relationships were constructed by using a strictly nested quadrat design. Grey lines correspond to a power law with a slope of 1; that is, proportionality between area and the number of species. S is the mean number of species, E is the mean number of endemics, and A is the area in km^2.

[1]Center for Theoretical Study, Charles University and the Academy of Sciences of the Czech Republic, Jilská 1, 110 00, Praha 1, Czech Republic. [2]Department of Ecology, Faculty of Science, Charles University, Viničná 7, 128 44 Praha 2, Czech Republic. [3]Department of Ecology and Evolutionary Biology, Yale University, 165 Prospect Street, New Haven, Connecticut 06520-8106, USA.

Table 1 | Slopes of the EARs and SARs calculated by using the nested quadrat design

Taxon	Continent	EAR slope	SAR slope (lower half)	SAR slope (upper half)
Birds	Eurasia	1.92 (1.70–2.06)	0.21 (0.19–0.22)	0.41 (0.37–0.45)
	Africa	1.39 (1.22–1.65)	0.21 (0.19–0.23)	0.40 (0.34–0.45)
	N. America	1.25 (1.00–1.61)	0.19 (0.17–0.20)	0.29 (0.27–0.32)
	S. America	1.30 (1.15–1.44)	0.17 (0.16–0.19)	0.39 (0.35–0.43)
	Australia	1.96 (1.20–2.67)	0.15 (0.12–0.17)	n.a.
Mammals	Eurasia	1.36 (1.26–1.43)	0.26 (0.24–0.27)	0.48 (0.44–0.52)
	Africa	1.24 (1.15–1.39)	0.23 (0.21–0.25)	0.48 (0.42–0.54)
	N. America	1.26 (1.11–1.43)	0.21 (0.19–0.22)	0.40 (0.34–0.45)
	S. America	1.21 (1.10–1.36)	0.17 (0.16–0.18)	0.44 (0.41–0.46)
	Australia	1.46 (1.09–1.85)	0.21 (0.19–0.23)	n.a.
Amphibians	Eurasia	1.15 (1.05–1.31)	0.35 (0.31–0.40)	0.70 (0.54–0.86)
	Africa	1.13 (1.05–1.25)	0.29 (0.26–0.32)	0.64 (0.53–0.75)
	N. America	1.00 (0.84–1.16)	0.26 (0.22–0.31)	0.58 (0.43–0.70)
	S. America	0.93 (0.86–1.02)	0.26 (0.24–0.28)	0.58 (0.49–0.68)
	Australia	1.12 (0.80–1.53)	0.27 (0.23–0.33)	n.a.

The slopes were estimated by using linear regression on logarithms of the mean number of species for each area (also logarithmically transformed). The EAR slopes were calculated across the whole range of areas. The SAR slopes were calculated separately for the lower and upper half of the analysed areas (cut-off: $\log_{10}$ area = 6.1) to provide measures of both the lower and upper ends of the upward-accelerating SARs. For local slope estimates at each area see Fig. 1d. To give a general representation of the possible range of slopes that would be detected if biodiversity data were incomplete, we randomly selected only 10% of all possible positions of sampling windows, repeated the procedure 500 times and estimated the lower and upper 95% quantiles of the slopes obtained from the resampled data (see Methods).

higher species richness for a given area than do mammals, whereas amphibians typically show low richness. However, amphibians also show much steeper SARs than other taxa (Fig. 1d), and Eurasia has the steepest SARs for all taxa. An assessment of local slopes (derivatives) of the SARs illustrates the upward-increasing nature of the logarithmic SAR (Fig. 1d and Table 1). Considerable differences in this increase appear among taxa, most clearly between Eurasian amphibians and North American birds. We do not find evidence for the first phase of the triphasic SARs, which confirms the expectation that this phase occurs only when the number of individuals becomes limited[4,14]; that is, at scales considerably finer than those made possible by the current grain size of global distribution data[19].

The nonlinear shapes of SARs stand in striking contrast to those observed for EARs (Fig. 1e–g). All continents and taxa show a consistent and seemingly linear increase in number of endemics with increasing area in logarithmic space. Local slopes of EARs are reasonably invariant with scale, taxon and continent (Fig. 1h), although some show a slight increase at areas above $3 \times 10^6\,\mathrm{km}^2$. Except for generally steeper slopes in birds, EAR slopes tend to vary between 0.75 and 1.5 and are often close to 1 (Fig. 1h and Table 1), indicating that the number of endemic species increases more or less proportionally with area.

The increasing slope of the SAR at large spatial scales is predicted to be associated with increasing species spatial turnover as sample areas approach the sizes of species' geographic ranges[2–4] (see Supplementary Discussion). We contend that the SAR curvature may thus be dependent on an 'effective' range size equal to the mean range size. Because of its similar foundation, we predict that EARs will show similar range size dependence. We therefore rescaled all area axes such that one areal unit corresponds to the mean species geographic range for a given continent and taxon:

$$A_r = A/\overline{R}_{t,c} \qquad (1)$$

where A_r is the rescaled area, A is the area of the study plot and $\overline{R}_{t,c}$ is the mean range size for taxon t and continent c. In addition we rescaled the vertical axis to represent species richness proportional to the richness of an area equal to $\overline{R}_{t,c}$:

$$S_r = S_A/S_{\overline{R(t,c)}} \qquad (2)$$

and

$$E_r = E_A/E_{\overline{R(t,c)}} \qquad (3)$$

where S_r and E_r are the rescaled counts of species and endemics, S_A and E_A are mean counts for a given area, and $S_{\overline{R(t,c)}}$ and $E_{\overline{R(t,c)}}$ are mean richness values for the area that equals the mean geographic range size of a given taxon and continent. Under this transformation, the original SARs and EARs collapsed into an approximately single curve (Fig. 2a, b).

This collapse was also observed by using an alternative sampling design based on continental (self-similar) instead of quadratic shapes of study plots (Fig. 2c, d; see Methods). The steeper SARs observed for amphibians (Fig. 1a) can thus be attributed to their considerably smaller ranges: the SAR increases rapidly at smaller absolute areas, and the slope continues to increase, whereas the other two taxa with much larger range sizes never approach a similarly steep relationship (see Supplementary Fig. 11).
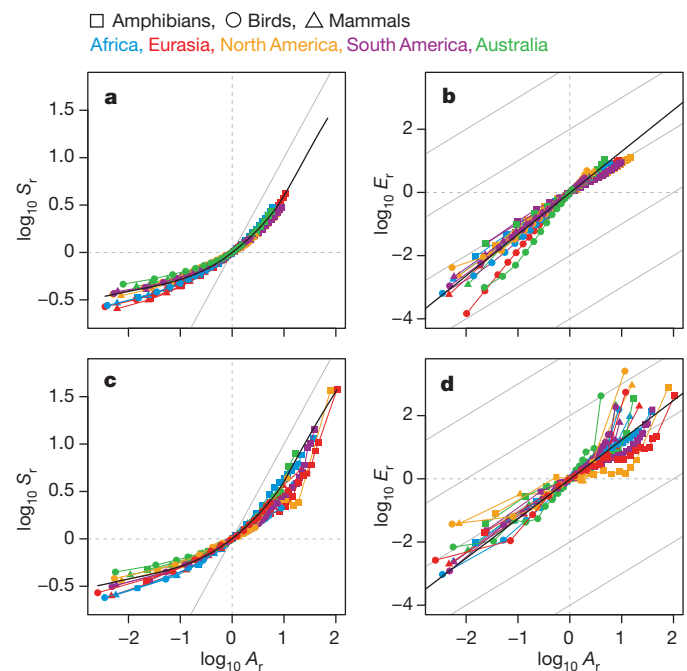

☐ Amphibians, ◯ Birds, △ Mammals
Africa, Eurasia, North America, South America, Australia

Figure 2 | SARs and EARs after rescaling. a, b, After expressing the area in units corresponding to mean range size and standardizing the vertical axis so that it represents species richness relative to mean richness for a given unit area, all the SARs and EARs collapse into one universal relationship, although some deviations exist, particularly in small areas. For EARs (**b**), birds in Eurasia and Australia represent the only considerable deviations. In these regions many endemics with small ranges occur at the edge of the continents, whereas the areas for which the EARs were calculated were taken predominantly from the centre of the continent. **c, d,** These universal relationships also exist for SARs and EARs constructed using the alternative, continental shape design, in which sample areas are not quadrats but keep the shape of the given continent (see Methods and Supplementary Discussion). Solid black lines refer to rescaled SARs and EARs predicted by simulations based on a random placement of simplified ranges (model 3; see Fig. 3 and Methods). Solid grey lines all have slope of 1. For explanations of A_r, S_r and E_r see equations (1), (2) and (3).

To test whether the observed continental relationships and their collapse can be predicted on the basis of simple assumptions concerning spatial distribution of species ranges, we developed several spatially explicit simulation models (see Methods). Specifically, we assessed the degree to which the observed SARs and EARs may be recovered by using placement of species ranges modelled as simple contiguous shapes. Although multiple evolutionary and ecological factors ultimately determine the exact location and sizes of geographic ranges (which in turn affect SAR and EAR), we find that just a few assumptions are sufficient to explain the observed scaling relationships (Fig. 3). Among our models those with independent (models 3 and 4; see Fig. 3 and Methods for details), as opposed to clumped (models 1 and 2), placement of geographic ranges produce SAR and EAR shapes as well as collapses of all curves that are essentially identical to those observed. The resulting patterns are not particularly sensitive to the exact shape of the frequency distribution of range sizes, because model 3 (which retained the observed range size distributions) provided very

similar patterns to those of model 4, which did not retain them (see Supplementary Discussion). In contrast with models 3 and 4, the observed geographic range locations and sizes do show spatial non-independence, but much less so than in model 2, and the effect on the observed collapse is minimal (Supplementary Figs 13–15 and Supplementary Discussion). The empirical patterns (Fig. 2) are therefore expected whenever a species distribution is represented by more or less independently located contiguous ranges, with the mean range size of species being the only biologically relevant variable affecting the exact properties of the patterns.

The universality of SARs and EARs after rescaling implies that a knowledge of mean species richness (of either endemics or all species) at one scale allows the estimation of the whole SARs and EARs with only one additional piece of information: mean range size for species in the region. Although this information may not usually be available without knowledge of the geographic distribution of all species (and thus also the whole SARs and EARs), in some cases it may be estimated
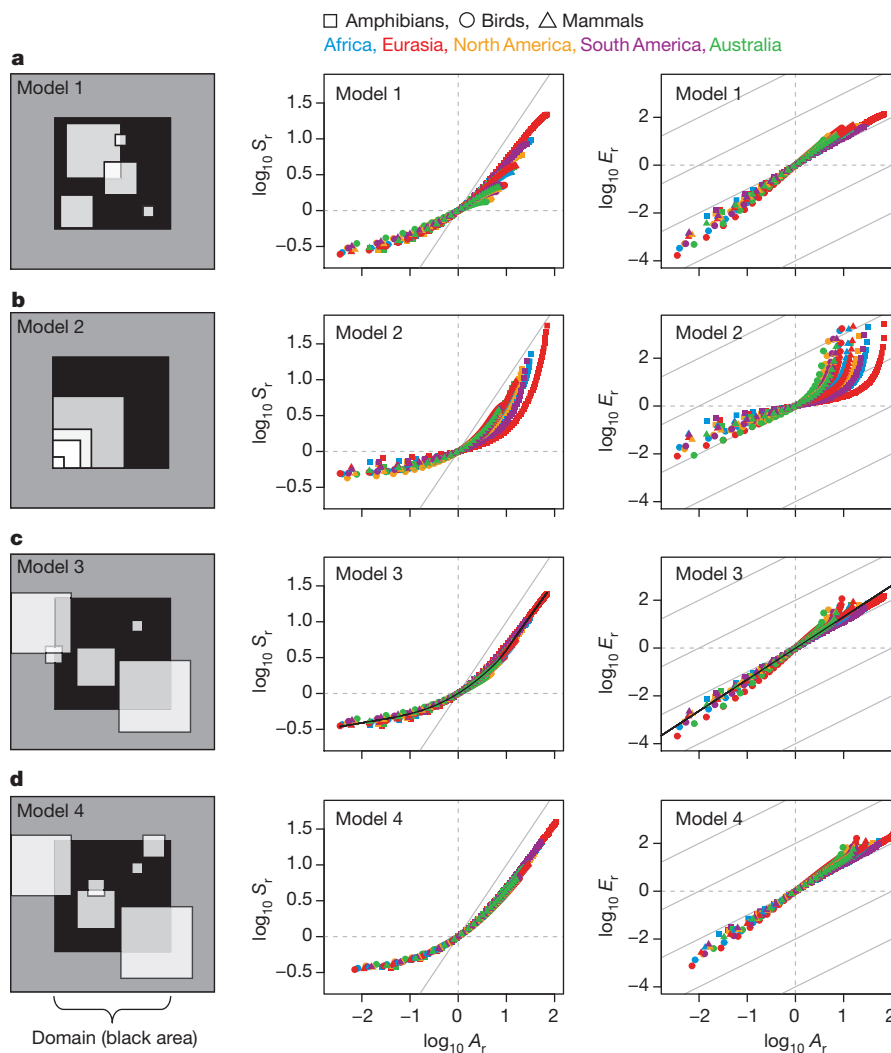

Figure 3 | Rescaled SARs and EARs predicted by four simulation models of range placement. Range sizes were drawn from the empirical frequency distributions of each taxon and domain (black areas in Supplementary Fig. 3) and were placed into a domain with a size equal to that of the regions analysed for the original SARs and EARs using the strictly nested quadrat design (Supplementary Fig. 3 and Methods; for the results based on the regions analysed using the continent shape design see Supplementary Fig. 12). **a**, Model 1 is based on a random placement of square ranges within the domain, producing a higher concentration of range midpoints in the centre of the domain (mid-domain effect[30]). **b**, Model 2 places all the square ranges in a corner of the domain, to illustrate the role of non-random range position.

c, Model 3 is based on a random placement of square ranges but minimizes the mid-domain effect by allowing model ranges to overlap the domain only partly. The observed frequency distribution of range sizes is retained, but resulting range shapes within the domain become variable. **d**, Model 4 is similar to model 3 and completely avoids the mid-domain effect but does not retain the originally observed range size distribution (see Methods for details). We produce a fitted line for model 3 results to highlight its match with the empirical patterns (see Fig. 2): black lines represent the Lowess regression line for the rescaled SAR plot (smoothing span 0.2) and the linear regression line for the rescaled EAR plot. Solid grey lines all have slope of 1.

reasonably well from similar taxa or representative subtaxa. However, the larger deviations from the universal relationship in small areas (Fig. 2a) will probably limit the accuracy of richness estimates towards smaller spatial scales. This is expected, because the SAR and EAR at these scales are determined by complex shapes of individual geographic ranges[22], which go beyond simple differences in range sizes.

In contrast to the strongly nonlinear EARs previously reported for small spatial scales[9], the slope of the continental EARs assessed here was generally close to 1. Therefore, for the scales examined, the number of species predicted to go extinct is roughly proportional to the area destroyed. However, the uncertainty in any such predictions would be large (Table 1), and the shape of the EAR is unlikely to hold up at the finer spatial scales relevant to conservation[9]. In addition, the relationships addressed here comprise only the direct effect of shrinking habitable area and not the cascading effects of species interactions or other consequences of species loss. Nevertheless, for the scales analysed here, the relatively steep slopes of all EARs suggest large extinction rates from area loss.

Spatial biodiversity patterns, including the SAR and EAR, are affected by many factors, ranging from the spatial arrangements of continental masses and biomes and the patterns of diversification and dispersal playing out within and among them, to population dynamics and interspecific interactions[23,24]. However, all of these processes ultimately translate into patterns of geographic distribution of individual species, which then represent a proximate driver of spatial macroecological patterns[22,25]. We have shown that species range sizes have a key role in large-scale upward-accelerating SARs, and consequently also EARs at specific spatial scales. Because local SAR slopes are mathematically related to species spatial turnover[14,26,27], the determinants of species range sizes predictably determine global patterns in species spatial turnover, the SAR and the EAR. These findings suggest that an integrated evolutionary and ecological understanding of just a few attributes of regional biota can enable far-reaching predictions to be made about the scaling of biodiversity.

METHODS SUMMARY

We calculated mean species richness across all possible quadrats of given size by using the strictly nested quadrat method[28,29]. We calculated SARs for continental regions able to accommodate quadrats encompassing 20×20 grid cells, except for Australia where because of the smaller size we used 14×14 (Supplementary Fig. 3). This avoids potential biases resulting from different species richness in marginal areas that could not be sampled by large quadrats (see Supplementary Fig. 4). In a second sampling design we adjusted plot boundaries to mimic continental shapes (Supplementary Fig. 5), which increased the amount of edge area that could be included in regions such as North America with more complicated geometry. This procedure yielded qualitatively similar, but noisier, results (see Methods, Supplementary Discussion, Supplementary Fig. 6 and Supplementary Table 2 for details). We used four simulation models of range placement within a domain to examine the effects of range position (random or spatially clumped) and shape (constant or varying) on resulting SARs and EARs (see Fig. 3 for more details on the models).

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 5 December 2011; accepted 10 May 2012.
Published online 24 June 2012.

1. Rosenzweig, M. L. *Species Diversity in Space and Time* (Cambridge Univ. Press, 1995).
2. Allen, A. P. & White, E. P. Effects of range size on species–area relationships. *Evol. Ecol. Res.* **5,** 493–499 (2003).
3. McGill, B. & Collins, C. A unified theory for macroecology based on spatial patterns of abundance. *Evol. Ecol. Res.* **5,** 469–492 (2003).
4. Hubbell, S. P. *The Unified Theory of Biodiversity and Biogeography* (Princeton Univ. Press, 2001).
5. Rosindell, J. & Cornell, S. J. Species–area relationships from a spatially explicit neutral model in an infinite landscape. *Ecol. Lett.* **10,** 586–595 (2007).
6. O'Dwyer, J. P. & Green, J. L. Field theory for biogeography: a spatially explicit model for predicting patterns of biodiversity. *Ecol. Lett.* **13,** 87–95 (2010).
7. Lawton, J. H. & May, R. (eds) *Extinction Rates* (Oxford Univ. Press, 1995).
8. Pimm, S. L. & Raven, P. Biodiversity: extinction by numbers. *Nature* **403,** 843–845 (2000).
9. He, F. & Hubbell, S. P. Species–area relationships always overestimate extinction rates from habitat loss. *Nature* **473,** 368–371 (2011).
10. Smith, A. B. Caution with curves: caveats for using the species–area relationship in conservation. *Biol. Conserv.* **143,** 555–564 (2010).
11. Drakare, S., Lennon, J. J. & Hillebrand, H. The imprint of the geographical, evolutionary and ecological context on species–area relationships. *Ecol. Lett.* **9,** 215–227 (2006).
12. Connor, E. F. & McCoy, E. D. The statistics and biology of the species–area relationship. *Am. Nat.* **113,** 791–833 (1979).
13. Harte, J., Smith, A. B. & Storch, D. Biodiversity scales from plots to biomes with a universal species–area curve. *Ecol. Lett.* **12,** 789–797 (2009).
14. Šizling, A. L., Kunin, W. E., Šizlingová, E., Reif, J. & Storch, D. Between geometry and biology: the problem of universality of the species–area relationship. *Am. Nat.* **178,** 602–611 (2011).
15. Kinzig, A. P. & Harte, J. Implications of endemics–area relationships for estimates of species extinctions. *Ecology* **81,** 3305–3311 (2000).
16. Pereira, H. M., Borda-de-Água, L. & Martins, I. S. Geometry and scale in species–area relationships. *Nature* **482,** E3–E4 (2012).
17. Green, J. L. & Ostling, A. Endemics–area relationships: the influence of species dominance and spatial aggregation. *Ecology* **84,** 3090–3097 (2003).
18. Belmaker, J. & Jetz, W. Cross-scale variation in species richness–environment associations. *Glob. Ecol. Biogeogr.* **20,** 464–474 (2011).
19. Hurlbert, A. H. & Jetz, W. Species richness, hotspots, and the scale dependence of range maps in ecology and conservation. *Proc. Natl Acad. Sci. USA* **104,** 13384–13389 (2007).
20. Storch, D. et al. Energy, range dynamics and global species richness patterns: reconciling mid-domain effects and environmental determinants of avian diversity. *Ecol. Lett.* **9,** 1308–1320 (2006).
21. Buckley, L. B. & Jetz, W. Environmental and historical constraints on global patterns of amphibian richness. *Proc. R. Soc. Lond. B* **274,** 1167–1173 (2007).
22. Storch, D. et al. The quest for a null model for macroecological patterns: geometry of species distributions at multiple spatial scales. *Ecol. Lett.* **11,** 771–784 (2008).
23. Losos, J. B. & Schluter, D. Analysis of an evolutionary species–area relationship. *Nature* **408,** 847–850 (2000).
24. Ricklefs, R. E. A comprehensive framework for global patterns in biodiversity. *Ecol. Lett.* **7,** 1–15 (2004).
25. McGill, B. J. Towards a unification of unified theories of biodiversity. *Ecol. Lett.* **13,** 627–642 (2010).
26. Harte, J. & Kinzig, A. P. On the implications of species–area relationships for endemism, spatial turnover, and food web patterns. *Oikos* **80,** 417–427 (1997).
27. Arita, H. T. & Rodríguez, P. Geographic range, turnover rate and the scaling of species diversity. *Ecography* **25,** 541–550 (2002).
28. Scheiner, S. M. Six types of species–area curves. *Glob. Ecol. Biogeogr.* **12,** 441–447 (2003).
29. Leitner, W. A. & Rosenzweig, M. L. Nested species–area curves and stochastic sampling: a new theory. *Oikos* **79,** 503–512 (1997).
30. Colwell, R. K. & Lees, D. C. The mid-domain effect: geometric constraints on the geography of species richness. *Trends Ecol. Evol.* **15,** 70–76 (2000).

METHODS

Description of the strictly nested quadrat (SNQ) design. We calculated mean species richness across all possible quadrats of given size by using the strictly nested quadrat (SNQ) method, which is a Type I curve in Scheiner's[28] terminology. It was implemented by using a moving-window algorithm[29,31]. SNQ implies mutual dependence of species richness at different spatial scales, as the species richness of larger areas encompasses all the species of the smaller plots within them. In this design, species richness of small areas can thus never be higher than the richness of the larger areas within which they sit. The overlapping nature of SNQ could be criticized for introducing some pseudoreplication as each point in space is sampled repeatedly by many samples of a given area. However, every SAR construction method has its limitations, and the SNQ design has several advantageous properties: first, it keeps the spatial extent[32] and shape[33] of the sampling window identical at all scales; second, it provides the most accurate estimate of expected species richness for a randomly located plot of a given area; and third, it is the only design in which local slope can be directly related to β diversity and patterns of species' spatial aggregation[14,22].

We started the SNQ procedure with the largest sampling window, which we moved continually across the world grid and counted the number of species captured within each window position. For the SAR construction we selected only areas that could contain the largest window without also including any sea or major water bodies (Supplementary Fig. 3). These are the black areas in Supplementary Fig. 3. Very large window sizes enabled us to explore SARs for a large range of areas, but only for limited proportion of a continent. In contrast, small window sizes fitted a larger proportion of continents, but the resulting SARs encompassed only a limited range of areas. We therefore initially set the largest window size to all values between 5×5 and 35×35 grid cells, and subsequently chose 20×20 grid cells because they were sufficiently representative of both continental coverage and range of areas in the SAR. However, for the continent of Australia we used 14×14 grid cells instead.

We then reduced the size of the sampling window to 19×19 and moved it continually within black areas, counting species within each possible window position. We repeated this procedure until the size of the sampling window reached 1×1 grid cell. The mean species richness for a given area (S_A) was calculated as

$$S_A = \frac{1}{n} \sum_{i=1}^{n} S_{A,i}$$

where n is number of all possible positions of the sampling window of area A within the black area (Supplementary Fig. 4).

Description of continent shape (CS) design. We developed a novel and alternative SAR construction design, which we call the continent shape (CS) design. It has the advantage that, unlike SNQ, it can show SARs that include areas as large as whole continents. It also keeps the shape of the sampling window approximately constant. The only disadvantage is that it is not strictly nested because the complex shapes of sampling windows cannot be placed everywhere within the complex shape of a given continent (Supplementary Figs 4 and 5). Therefore the coverage (the black area in Supplementary Fig. 3) for different sizes of sampling windows can vary, and hence different places within a given continent are not equally represented in plots of different areas.

The CS design works as follows. We first counted the number of species in a whole continent (such as Africa). We then multiplied the coordinates of each grid cell within Africa by a constant k ($0 < k < 1$) to obtain an approximated smaller representation of the African continental shape, which we then used as a moving sampling window in a same way as in the SNQ algorithm. We used this approach for five major land masses: North America, South America, Africa, Australia and Eurasia (Europe and Asia combined). We manually excluded most islands. The principle of the algorithm is further described in Supplementary Figs 4 and 5. The black area covered by the CS is illustrated in Supplementary Fig. 3.

Quantifying variation of SARs and EARs. Our results are based on mean values of the number of species or endemics for each area. Because we have analysed all possible plots on the whole Earth, it is not straightforward to express the variation around these 'mean' curves. We recognize that it is impossible to use standard statistical tools that assume that measured values represent samples from some larger universe (population) and to calculate an error of the estimated mean values when we have invoked the whole population, not just samples of it. Thus, we only calculate characteristics concerning the distribution of values, namely percentiles (Supplementary Figs 7–10). It is impossible to use standard regression tools for the same reasons, so to estimate some statistics concerning the curves themselves, we resampled the values and estimated the possible range of slopes by randomly selecting 10% of the possible positions of sampling windows, repeating the

procedure 500 times and estimating the lower and upper 95% quantiles of the slopes obtained from these resampled data (Table 1 and Supplementary Table 2).

Estimation of $S_{\overline{R(t,c)}}$ and $E_{\overline{R(t,c)}}$. In equation (2) we needed mean species richness of an area equal to the mean range size ($S_{\overline{R(t,c)}}$).The use of a sampling window of constant shape in a gridded data poses the problem that mean range size ($\overline{R}_{t,c}$) is mostly not exactly equal to any area (A) of the sampling window. We therefore took mean species richnesses S_{A1} and S_{A2} in sampling windows of areas A_1 and A_2 that were closest to $\overline{R}_{t,c}$ and satisfied $A_1 < \overline{R}_{t,c} < A_2$. We then calculated $S_{\overline{R(t,c)}}$ from a local power-law approximation of the SAR curve. Scaling exponent of the power law was $(\log S_{A2} - \log S_{A1})/(\log A_2 - \log A_1)$. The and hence $S_{\overline{R(t,c)}} = \exp[\log S_{A1} + (\log \overline{R}_{t,c} - \log A_1) \times (\log S_{A2} - \log S_{A1})/(\log A_2 - \log A_1)]$. The same approach was used to calculate $E_{\overline{R(t,c)}}$.

Models of range placement. We sought to explore how the rescaled SARs and EARs can be influenced by the spatial position of species' ranges (random or aggregated), the shape of ranges (uniform or variable) and the range-size frequency distribution. We developed four models in which square geographic ranges of species were placed on an artificial square continent (Fig. 3)—an approach similar to that in refs 2 and 29. The sizes of these ranges were drawn from the empirical distribution of a given taxon within a given domain. The domain is the black area for which SARs and EARs were explored (Fig. 3). Each model had a domain of approximately the same size as the domains within the real continents (the black areas in Supplementary Fig. 3). We replicated our simulations by using the domain size and empirical range-size frequency distribution for both SNQ design (smaller domains) and CS design (larger domains). We performed a simulation for each taxon, each continent and each model and plotted the results in the form of rescaled SAR and EAR (identically to Fig. 2). The models were characterized as follows.

Model 1: random placement, uniform shape of ranges. This model randomly places species' ranges of uniform (square) shape strictly inside the domain. This model incurs a strong mid-domain effect, leading to both higher species richness in central areas and relatively higher species richness for larger areas. The reason is that for uniform range shapes, large ranges will necessarily reach the central region of the domain and will therefore be necessarily sampled by larger sample windows (that is, there is no possibility of avoiding them by any position of the sampling window[34]).

Model 2: non-random placement, uniform shape of ranges. This model is very similar to model 1. The only difference is that instead of being placed randomly, all ranges are placed into one corner of the domain. Although the placement of range midpoints is then more balanced with respect to their central–peripheral position, this model still leads to the over-representation of large ranges in the central area of the domain as a result of the uniform shape of all ranges, including large ones, which cannot be avoided by any placement of the sampling window.

Model 3: random placement, variable shape of ranges. Here randomly placed species' ranges may extend beyond the continental domain. This weakens the mid-domain effect (but does not eliminate it completely); however, as a consequence, the shapes of ranges that are inside the domain are no longer uniform (they are cut by the domain boundary). The algorithm operates as follows. (1) Draw a range size from a given distribution of range sizes. (2) Place the range randomly into the continent so that it overlaps at least one grid cell within the domain. (3) Calculate the species range size (R_{domain}) as the part of the placed range that lies within the domain. (4) Try to find one value in the empirical range-size distribution ($R_{empirical}$) that is closest to R_{domain} and at the same time lies in the interval between $R_{domain} - \sqrt{R_{domain}}$ and $R_{domain} + \sqrt{R_{domain}}$. If such a value is found, the species is accepted to exist in the domain and the $R_{empirical}$ value is eliminated from the empirical range size distribution. If an acceptable $R_{empirical}$ is not found, steps 2 and 3 are repeated. (5) The procedure is repeated until all values of $R_{empirical}$ have been eliminated from the empirical distribution of range sizes (that is, all species have been placed into the domain). This model retains the observed distribution of range sizes in the domain very close to the original range size distribution (it very well preserves its mean and variance), and it partly eliminates the mid-domain effect because large ranges may have various shapes and thus do not necessarily reach the central areas.

Model 4: random placement, variable shape of ranges. As in the previous model, this model randomly places species' ranges at least partly inside the domain, so that the areas outside the domain are cut off. However, there is no algorithm that would ensure that the resulting distribution of range sizes within the domain is similar to the original range size distribution. There is therefore no control over the mean and variance of the resulting distribution of range sizes within the domain, but this model does completely eliminate any mid-domain effect.

Spatial patterns in range location and size. We performed an additional analysis (Supplementary Figs 13–15) to assess the similarity of spatial patterns of range location and size between simulation models and empirical data. Specifically, we investigated whether the spatial distribution of empirical ranges resembles a

random placement (models 1, 3 and 4) or is closer to the clumped distribution produced by model 2.

To assess the magnitude of randomness in range locations we first calculated the density (that is, count or 'richness') of species geographic range centroids (centres of gravity) for all CS continents and all three taxa. Range centroids were estimated by using gridded species' ranges rather than original range polygons. Centroids falling in between two grid cells were assigned randomly to one or the other. We then calculated spatial correlograms by plotting Moran's I of grid cell centroid count against geographic distance. We repeated the same procedure for the simulation models 1–4. The sizes of the artificial square continents required in these models were identical to those of the continents explored using the CS design (Supplementary Fig. 3). In these simulations we used the empirical distribution of range sizes of each taxon in the continents explored using the CS design. We ran 100 simulations for each continent and taxon combination (1,500 in total) and calculated mean correlograms of the simulations together with 95% confidence intervals.

We additionally assessed the spatial randomness of species geographic range sizes; that is, whether ranges of similar sizes tend to be clumped together or not. We used the same data and simulations as described above (CS design continents, 100 simulations for each continent, taxon and model) and calculated correlograms of range sizes; that is, the autocorrelation of range sizes (measured as Moran's I) plotted against the geographic distance between range centroids. For this we used all ranges, including those with centroids off the mainland.

31. Lennon, J. J., Koleff, P., Greenwood, J. J. D. & Gaston, K. J. The geographical structure of British bird distributions: diversity, spatial turnover and scale. *J. Anim. Ecol.* **70,** 966–979 (2001).
32. Nekola, J. C. & White, P. S. The distance decay of similarity in biogeography and ecology. *J. Biogeogr.* **26,** 867–878 (1999).
33. Kunin, W. E. Sample shape, spatial scale and species counts: implications for reserve design. *Biol. Conserv.* **82,** 369–377 (1997).
34. Šizling, A. L. & Storch, D. Power-law species–area relationships and self-similar species distributions within finite areas. *Ecol. Lett.* **7,** 60–68 (2004).