

ARTICLE

<https://doi.org/10.1038/s41467-019-14267-y>

OPEN

# Dispersion fields reveal the compositional structure of South American vertebrate assemblages

Michael K. Borregaard<sup>1\*</sup>, Gary R. Graves<sup>1,2</sup> & Carsten Rahbek<sup>1,3,4</sup>

The causes of continental patterns in species richness continue to spur heated discussion. Hypotheses based on ambient energy have dominated the debate, but are increasingly being challenged by hypotheses that model richness as the overlap of species ranges, ultimately controlled by continental range dynamics of individual species. At the heart of this controversy lies the question of whether species richness of individual grid cells is controlled by local factors, or reflects larger-scale spatial patterns in the turnover of species' ranges. Here, we develop a new approach based on assemblage dispersion fields, formed by overlaying the geographic ranges of all species co-occurring in a grid cell. We created dispersion fields for all tetrapods of South America, and characterized the orientation and shape of dispersion fields as a vector field. The resulting maps demonstrate the existence of macro-structures in the turnover of biotic similarity at continental scale that are congruent among vertebrate classes. These structures underline the importance of continental-scale processes for species richness in individual assemblages.

<sup>1</sup>Center for Macroecology, Evolution and Climate, GLOBE Institute, University of Copenhagen, Universitetsparken 15, 2100 Copenhagen, Denmark.

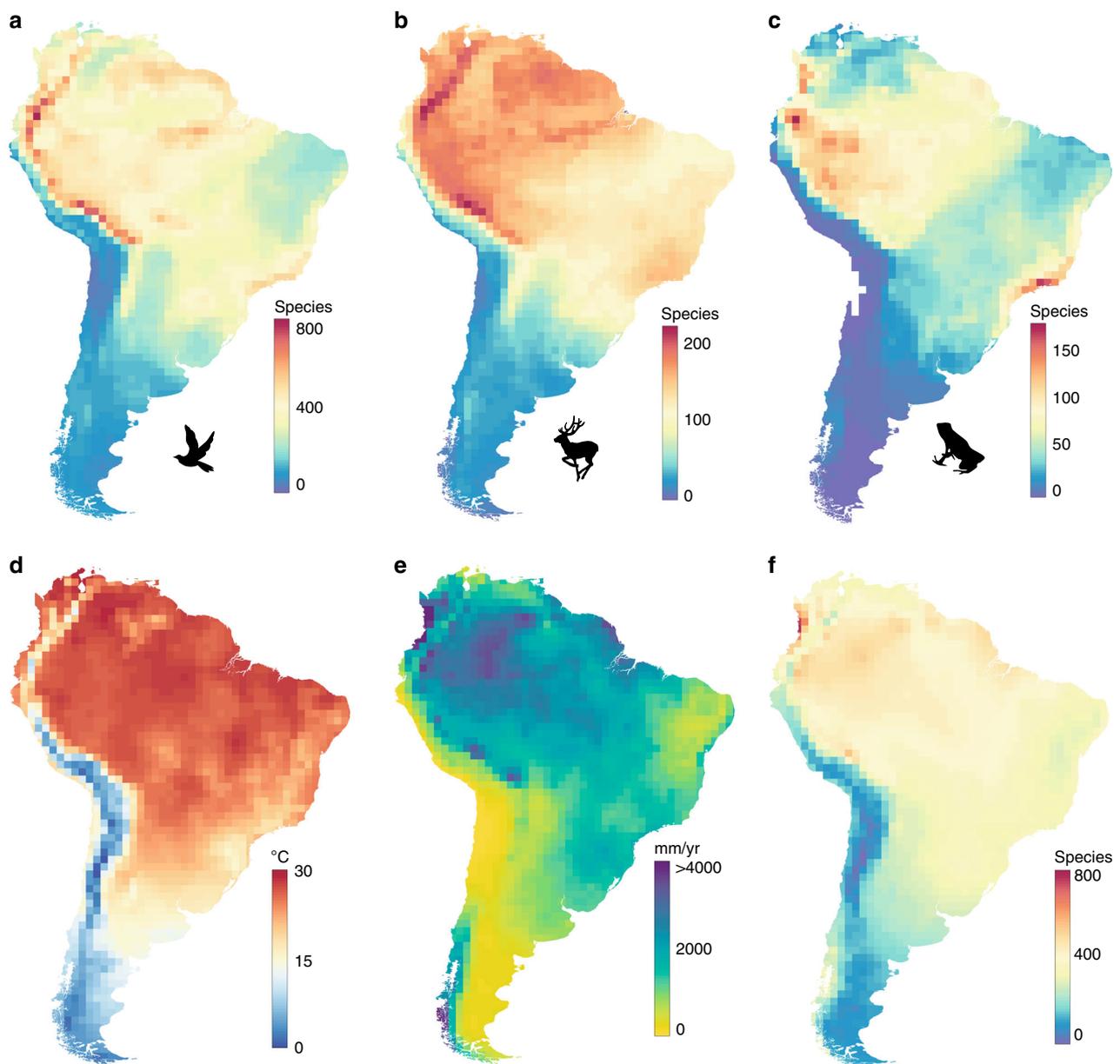
<sup>2</sup>Department of Vertebrate Zoology, National Museum of Natural History, Smithsonian Institution, Washington, DC 20013, USA. <sup>3</sup>Imperial College London, Silwood Park, Buckhurst, Road, Ascot, Berkshire SL5 7PY, UK. <sup>4</sup>Danish Institute for Advanced Study, University of Southern Denmark, Odense M 5230, Denmark. \*email: [mkborregaard@bio.ku.dk](mailto:mkborregaard@bio.ku.dk)

The cause of continental patterns of species richness is one of the most debated questions in ecology<sup>1–3</sup> and was for years considered the keystone question of macroecology. Though species richness of lowland regions generally correlates well with contemporary climate<sup>4</sup>, with mountainous regions correlating less well (refs. 5,6, Fig. 1), the mechanistic processes responsible for this correlation remain unresolved<sup>7,8</sup>. Hypotheses based on a direct local effect of ambient energy on species richness have dominated the debate, countered by theories focusing on species' niches and range dynamics, positing that species richness is an emergent property of species range overlap<sup>9–13</sup> that occurs at scales much greater than the grid cells used in most analyses. An example of the latter is the idea of tropical niche conservatism, which posits that climate–richness correlations arise because many species are evolutionarily adapted to warmer and more productive environments and thus have ranges extending into, or confined to, these areas. Since these hypotheses

make similar predictions for spatial variation in species richness, a resolution to this controversy has not been found, and in recent years relatively little progress has been made on the processes underlying patterns of large-scale species richness.

Though mechanisms based on (1) energy limiting rates of species origination<sup>14</sup>, (2) energy limiting local co-occurrence<sup>15</sup>, or (3) richness arising as an emergent property of species' niche overlap<sup>16</sup> will all lead to the observed correlation between climate and richness, they represent three different underlying pathways for the generation of continental species richness patterns. We argue here that these differences should be reflected in the continental pattern of species' turnover among grid cells, making it possible to distinguish between them.

Energetic limitation of speciation rates has generally been tested by correlating species richness with local energy levels<sup>4</sup> or by phylogenetically estimating the past diversification rates of all species whose current ranges overlap a given grid cell. Both of



**Fig. 1** Species richness and environmental variables in South America mapped in  $1^\circ \times 1^\circ$  grid cells. **a** Species richness of birds ( $n = 2869$ ), **b** mammals ( $n = 1146$ ), and **c** amphibians ( $n = 2265$ ). **d** Mean annual temperature ( $^\circ\text{C}$ ). **e** Annual precipitation (mm). **f** Predicted richness of birds based on a linear model of temperature and precipitation ( $R^2 = 0.63$ ).

these approaches will lead to biased results in the presence of spatio-temporal range dynamics. Because speciation in terrestrial vertebrates is usually allopatric<sup>17</sup>, newly geminate species will generally not co-occur, and thus the species richness of grid cells only increases where the species come into secondary contact. This may accumulate species in the area of origination, but if range expansion is not spatially random species richness will accumulate in the area that ranges expand into, which will greatly weaken the expected spatial link between local speciation rates and species richness. An empirical example of this is the high salamander diversity of Amazonia, which is mainly generated by species diversification in the Andes<sup>18</sup>. Thus speciation rate variation will only lead to strong energy–richness correlations when range expansion is relatively spatially symmetric around the area of origination.

Energetic limitation of species co-occurrence at large spatial scales<sup>19</sup> is an alternative local species–energy mechanism. This hypothesis is based on the theoretical premise that range expansion occurs more easily in areas with high energy and many resources and that local extirpation and range contraction is more likely in areas with less available energy. This is predicted to lead to range dynamics that are similar across different species, with high-energy areas as foci for all species range expansions.

Hypotheses based on range overlap, such as Tropical Niche Conservatism, in contrast, explicitly posit that range expansion history is determined by the spatial configuration of the landscape and underlies species richness patterns. Under the mechanism of niche conservatism, species with similar adaptations are likely to co-occur deterministically across sites, leading to large-scale emergent patterns of range overlap and geographically structured assemblages whose spatial distribution follow the asymmetric configuration of continental structures, such as the turnover of ecoregions characterized by different biomes<sup>20–22</sup>. These processes are predicted to leave a distinct biogeographical signature in the continental pattern of species distributions. Because these patterns result from a deterministic interaction with the physical environment, the pattern of disequilibrium should be predictable and consistent among taxa with shared habitat affinities.

The existence of geographically structured assemblages with a clear signature of biomes would thus support a key prediction of niche conservatism. Note that the well-established existence of biogeographic regions in species distributions<sup>23</sup> do not in themselves invalidate species–energy theory; however, as outlined above, the energy-based mechanisms will predictably affect spatial patterns of biotic similarity, giving us a previously unexplored opportunity to reassess the mechanistic basis of climate–richness correlations.

Here we develop a new approach to reveal and visualize geographic structuring of assemblages. The approach is based on assemblage dispersion fields, which are formed by overlaying the geographic ranges of all species co-occurring in a grid cell. We created dispersion fields for all tetrapods of South America and visualized their geographic structures by expressing their orientation and shape of dispersion fields as a vector field. These symmetry diagrams support the existence of macrostructures in the turnover of biotic similarity at a continental scale. The structures are congruent among vertebrate classes, lending support to the predictions of niche conservatism as a control of continental diversity patterns.

## Results and discussion

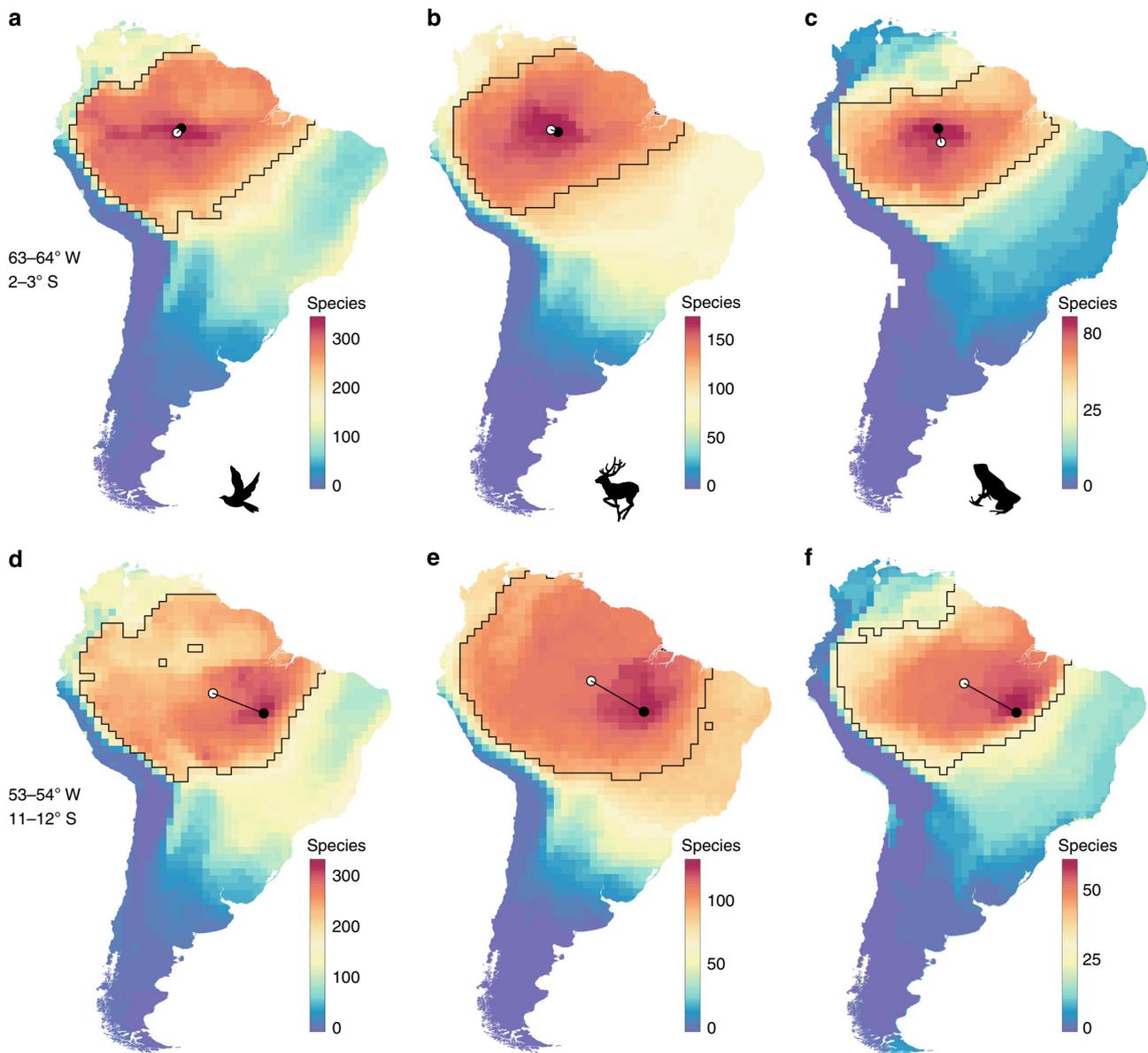
**Assemblage dispersion fields (ADFs).** We investigated the geographical pattern of compositional similarity for the mammals, amphibians, and freshwater and land birds in 1689 cells ( $1^\circ \times 1^\circ$  latitude–longitude blocks) in South America. To quantify patterns

of compositional similarity, we used the approach of ADFs, which are created by overlapping the geographic range maps of all species that occur in a specified grid cell<sup>24–27,28</sup>. The resulting contour map will peak at the focal cell and decline in all directions from this peak (Fig. 2, Supplementary Figs. 1 and 2). The contour slope along any vector illustrates the decay of compositional similarity with distance from the focal cell, revealing the biogeographical affiliation of the species assemblage of the focal cell. We generated ADFs for all  $1^\circ \times 1^\circ$  grid cells ( $n = 1689$ ) of the South American continent for all species of birds (2869 species), mammals (1146 species), and amphibians (2265 species).

Continental mapping of ADFs for individual grid cells reveals a clear geographical structure in the species composition of local assemblages<sup>24</sup> in all three vertebrate classes (Fig. 2). The size and shape of ADFs exhibited marked geographic variation, reflecting patterns of faunal turnover and the spatial configuration of major biomes. For example, the composition of assemblages of species in  $1^\circ \times 1^\circ$  grid cells occurring at opposite ends of the vast Amazonian ecoregion ( $\sim 5$  million  $\text{km}^2$ ) are more similar to one another than they are to geographically proximate assemblages in adjacent ecoregions that exhibit similar energy levels. ADFs sampled from the center of the Amazonian ecoregion are relatively symmetrical (Fig. 2a–c), whereas those sampled near the periphery of the ecoregion are asymmetrical (Fig. 2d–f). We quantified ADF symmetry by projecting a vector from the geographic center of each focal cell to the respective center of gravity for each ADF (Fig. 2; see “Methods”). The vector toward the center of gravity is a simple representation of the orientation and degree of asymmetry of the ADF. The direction of the vector indicates the major axis of asymmetry of the ADF, whereas vector length provides a direct measure of the degree of ADF asymmetry. In this context, asymmetry refers to a deviation from radial symmetry of individual dispersion fields.

**Symmetry diagrams.** We then mapped ADF symmetry diagrams for the continental lattice of cells using a color scale to illustrate the degree of ADF asymmetry and arrows to indicate the direction of the symmetry vectors (Fig. 3). This continental lattice of ADFs revealed a hidden biogeographical structure in the species composition of local assemblages that is not readily discerned from patterns of species richness (Fig. 3). For all three vertebrate groups, ADF symmetry diagrams revealed distinct geographical patterns in ADF symmetry and assemblage affinity (Fig. 3a–c). ADF vectors for the majority of adjacent grid cells tend to parallel one another or nearly so. However, ADFs near vegetation ecotones were highly asymmetrical, and the vectors of ADF symmetry diagrams straddling ecotones often point in orthogonal directions away from the ecotones and toward the centers of abutting ecoregions. In general, the degree of asymmetry increased with distance from these centers (Fig. 4). Groups of grid cells with similar ADFs correspond roughly to the configuration of major vegetation ecoregions (Fig. 3d), revealing, e.g., a clear separation of the faunas of Amazonian from the savannah-like biomes of the Cerrado and Caatinga, and an extremely rapid turnover of assemblages along the eastern versant of the Andes Mountains.

Note that, though the arrows on the diagrams might intuitively be understood as species movement, and indeed should to a degree reflect general patterns in range expansion, they do not prove movement and are not hypotheses about historical range dynamics for the faunas. The ADF symmetry diagrams are a representation of the current pattern of biotic similarity, whereas patterns of range expansion of individual species and clades can only be revealed by focused biogeographical analyses (and in the presence of fossils). Instead, the ADF symmetry diagrams reveal



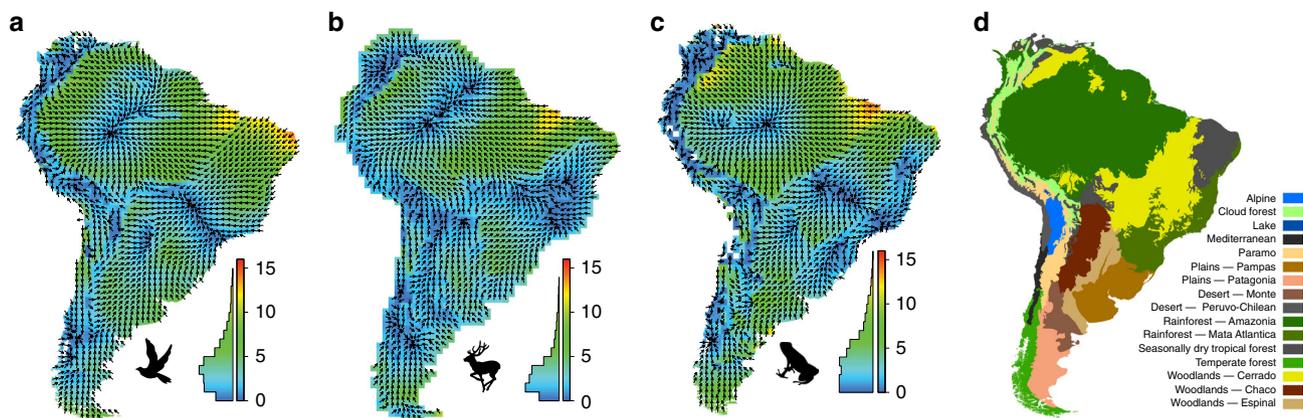
**Fig. 2** Asymmetry vectors and assemblage dispersion fields (ADFs) for selected  $1^{\circ} \times 1^{\circ}$  grid cells in Amazonia. The focal cell is marked by a black point. Colors indicate the number of species shared with the focal cell. The central region (see “Methods”) of the dispersion field is outlined by a black line. The asymmetry vector extends from the center of the focal cell to the center of gravity (white point) of the central region. ADFs of focal cells are relatively symmetrical near the center of the Amazonian ecoregion (a–c) and become increasingly asymmetrical near its periphery (d–f), for birds (a, d), mammals (b, e), and amphibians (c, f).

emergent structures in the current distribution of species that can be interpreted in the context of understanding the basis of the current pattern of species richness. The arrows reveal patterns at a larger scale than neighbor-based methods for compositional similarity, with the scale being determined by the species’ range sizes.

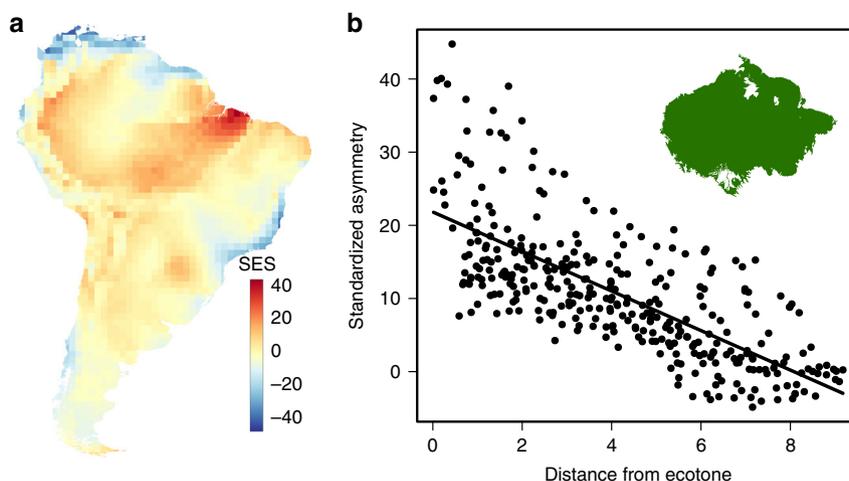
**Null model analysis.** We used a null model to correct for two properties of grid cells that are expected to affect ADF asymmetry in the absence of any ecological mechanism<sup>29,30</sup>: (1) The shape of the continent, which restricts the potential configuration of the ADF; and (2) the range–size frequency distribution of species in each grid cell, which will directly affect vector length, and thus asymmetry values. The null model assesses the expected ADF symmetry diagrams under random range placement (Fig. 5a)

constrained by the empirical species richness (Fig. 5b) and ecoregion boundaries (Fig. 5c) (see “Methods”; Supplementary Fig. 7).

As a basis for interpreting the empirical ADF symmetry diagrams, we also developed simple predictive models under three constraints to operationalize the verbal models presented above: (1) species originate in grid cells with higher energy levels and undergo random range expansion (Fig. 5a); (2) species originate randomly and range expansion is limited by a maximum level of co-occurring species (Fig. 5b); and (3) species originate randomly within their native ecoregion and undergo range expansion limited by ecoregion boundaries (Fig. 5c). The emergent pattern in Fig. 5c is broadly similar to the empirical pattern. Note that this type of predictive model is necessarily an over-simplification and will to a certain extent reflect the choices made in implementing the algorithm; we present the results here



**Fig. 3 ADF symmetry diagrams for South American birds, mammals, and amphibians and major vegetation biomes.** Arrows indicate the major axis of asymmetry of the ADFs for individual cells ( $n = 1689$ ). The degree of asymmetry is indicated by a color scale (green through red colors represent more asymmetrical ADFs). This figure summarizes the vector arrows shown in Fig. 2 for all grid cells on the continent. Histograms show the distribution of vector lengths. The panels show **a** birds, **b** mammals, **c** amphibians, and **d** vegetation ecoregions. ADF symmetry diagrams using different cutoff values for the central ADF region are shown in Supplementary Figs. 3 and 5. The map in **d** was provided by Tiina Särkinen, originally based on the map of the World's ecoregions developed by the WWF (ref. 48; under CC BY 4.0 license: <https://creativecommons.org/licenses/by/4.0/>).



**Fig. 4 Null model divergence of ADF vector symmetry.** **a** The deviation of the ADF symmetry vectors of birds from the null model (Supplementary Fig. 7), calculated as the standardized effect size (SES) deviation of vector lengths after 1000 repetitions of the null model. **b** The relationship between the deviation from the null model and the distance from the ecoregion boundary for the subset of grid cells ( $n = 342$ ) that are completely encompassed within Amazonia (Supplementary Fig. 9;  $R^2 = 0.52$ ). The least square regression explains 52% of the variation.

mainly as a context for discussing the effect of various constraints on ADF symmetry diagrams.

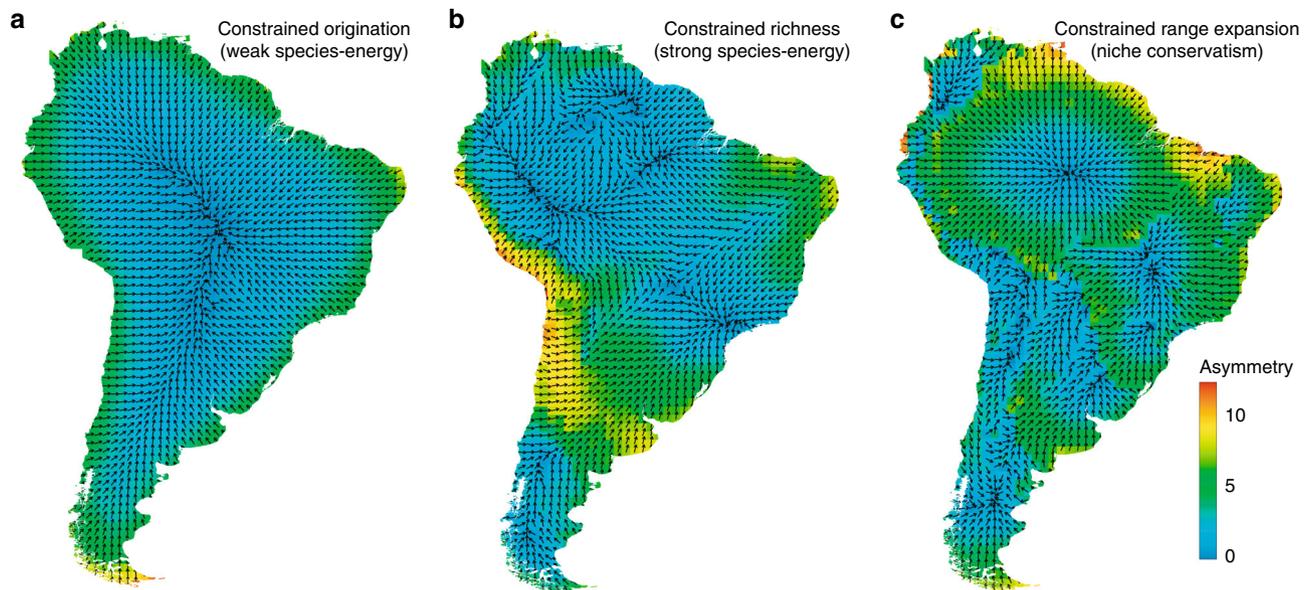
**Conclusion**

These results show that the species compositions of local assemblages are consistent with non-random range expansion limited by the geographical configuration of habitats, as suggested by niche theory. Though the patterns of ADF asymmetry are complex, they are consistent among three major vertebrate classes (Fig. 3). Thus range dynamics are not spatially random but instead result from deterministic processes that will weaken any spatial association between rates of speciation and species richness. Patterns of species richness appear instead to be created by an interaction between Grinnellian niches of species and geographical patterns of range overlap among species, which in turn depends on the configuration of major vegetation types. Furthermore, certain areas, such as the eastern Andes Mountains, exhibited very rapid faunal turnover, reflecting high turnover in species' habitats. Grid cells from this area are consistently

observed to support greater species richness than predicted purely from species-energy dynamics.

Our results support the idea that high species richness at the base of the Andes is due to high habitat heterogeneity and faunal turnover, rather than a simple response to elevated energy levels<sup>31</sup>. Our results show that, at the grain sizes of our analyses, the cohesive nature of the geographic ranges of species means that any process that affects species richness will also affect the species composition in grid cells and the large-scale congruence of species distributions. The processes behind species richness and species composition can thus not be decoupled at this scale, and importantly, employing classic regression methods that deal with spatial autocorrelation by imposing a symmetric neighborhood variance kernel (e.g., spatial autoregressive and conditional autoregressive) is not adequate to evaluate the effect of local processes on species richness. Other approaches, based on explicit mechanistic modeling of assembly processes, are likely to be necessary to resolve this.

The  $1^\circ \times 1^\circ$  represents the highest precision we can feasibly attain over such a large number of species but entails that many



**Fig. 5 Predicted assemblage dispersion field (ADF) symmetry diagrams based on predictive models.** The arrows reveal the pattern of biotic similarity. **a** The predicted ADF diagram from a model where species originate in grid cells according to local energy availability and then spread spatially randomly from there; **b** the predicted ADF diagram from a model where species richness is constrained to be identical to the empirical and ranges are constrained to be spatially cohesive; **c** the predicted ADF diagram from a model where each species originate in its original vegetation ecoregion and spreads spatially randomly with the constraint that ecoregion boundaries are crossed with low probability. Vegetation ecoregions are defined by the map shown in Fig. 3d.

areas within each occupied grid cells will not actually be occupied by a given species and that species co-occurring at the grid cell scale may never actually co-occur in local communities. This is particularly pronounced along ecotone boundaries, where many species range boundaries are expected to meet, making it hard to evaluate patterns, e.g., along the Andes.

It has long been clear that analyses of species richness must move beyond simple correlative analyses of species richness numbers and take account of the size, shape, and location of the geographic ranges of species. By doing so here, we are better able to evaluate the postulated mechanisms and predictions of competing hypotheses on the origin of species richness patterns, providing a needed corollary to the template predictions of energy—kinetic theory. The ADF diagrams provide a visually intuitive way of turnover in species composition and the relationship among different assemblages. The empirical results of our study are consistent with recent developments in ecological theory that argue that dynamics at the scale of complete ranges must be considered in an integrative theory for species richness<sup>32</sup>. This entails a revision of species richness theory that expands geographical grid-cell-based regression analyses of diversity patterns to integrate species niches and geographic range dynamics.

## Methods

**Distribution data.** The analyses were based on a distributional database of all species of birds, mammals, and amphibians in South America. The continent was divided into  $1^\circ \times 1^\circ$  grid cells corresponding to integer degree lines of latitude and longitude. All continental grid cells that contained land area inside were retained in the analysis, yielding a total of 1689  $1^\circ \times 1^\circ$  grid cells. Land-bridge islands (e.g., Trinidad) were excluded.

The bird distributions were extracted from the Copenhagen global bird database<sup>33</sup>, which is periodically updated. This database is based on museum specimens, published sight records, and expert opinion, following the approach outlined by Rahbek and Graves<sup>5,34</sup>. The distributions were mapped directly to the  $1^\circ \times 1^\circ$  latitude–longitude grid; hence, distributions are not based on a post hoc fitting of arbitrary-scale polygons to the grid cells (different from the mammal and amphibian datasets, which are based on IUCN data). The resulting high-quality dataset represents the best current knowledge of bird distributions in South America. To ensure comparability with other sources, the avian species delimitation followed the taxonomy published by the SACC<sup>35</sup> resulting in a dataset containing 2869 land and freshwater species. For a full list of the >1600 references

used to build the global dataset from which these distributions are derived, see the appendix of Holt et al.<sup>22</sup>

The mammal database was based on published range maps from the global mammal assessment<sup>36</sup>, as modified by Fritz and Purvis<sup>37</sup>. Domesticated species were excluded from the analysis, as well as portions of the geographic ranges of non-domesticated species labeled as historical, presence uncertain, introduced origin, or extinct. Species classified as Data Deficient, Extinct, or Extinct in the Wild in the IUCN database were excluded. The resulting database was updated to match the published phylogeny of mammals<sup>38</sup>. This entailed excluding 130 species that were absent from the phylogeny and adding 40 species from PanTHERIA<sup>39</sup> that were missing from the IUCN database. Finally, 56 species were added by splitting IUCN maps to reflect the phylogeny of Bininda-Emonds et al.<sup>40</sup>. Range polygons for the resulting 1146 mammalian species were overlaid on the  $1^\circ \times 1^\circ$  grid. A species was scored as present in any grid cell that overlapped the geographic range polygon for that species<sup>41</sup>.

The amphibian database was based on published range maps from the IUCN<sup>42</sup> and updated to follow the published taxonomy of amphibians<sup>43</sup>. Species labeled as incertae sedis, klepton, and undescribed taxa were excluded, as were species classified as introduced or uncertain/introduced. The final dataset consisted of range maps for 2265 amphibian species, which were overlaid on the  $1^\circ \times 1^\circ$  grid and scored as present or absent in each grid cell. We note that the amphibian dataset is not quite of the quality of the other two datasets and have been revealed to fit poorly with occupancy records<sup>44</sup>. We have chosen to include it here for consistency across the tetrapod groups, but conclusions based on the amphibian dataset should be interpreted carefully.

**Environmental data.** Records of mean annual temperature and annual precipitation were extracted from the mean monthly climatic database published by New and co-workers<sup>45</sup>, which was compiled globally at a  $0.5^\circ$  latitude–longitude resolution for the period 1961–1990 (>3,000,000 data points for each variable).

**Assemblage dispersion fields.** The occurrence of species can be summarized in a presence–absence matrix  $\mathbf{P}$ , where each row of the matrix represents a species, each column is a grid cell, and a 0 or 1 denotes absence or presence of a species within a grid cell<sup>46</sup>. The row sums represent the range sizes of species (measured as the total of occupied grid cells), and the column totals represent species richness values of grid cells. The ADF matrix  $\mathbf{D}$  is then defined as

$$\mathbf{D} = \mathbf{P}^T \mathbf{P}, \quad (1)$$

where  $T$  is the transpose operator.  $\mathbf{D}$  is a symmetrical site-by-site matrix that counts the number of species shared by any two grid cells. Each row (or column) represents the ADF for that cell and describes the number of species that a focal cell shares with any other cell. Mapping dispersion fields results in the maps presented in Fig. 2 and Supplementary Figs. 2 and 3, representing the geographical pattern of biotic similarity around the selected focal cells.

The row (and column) sums of the ADF matrix equal the ADF volume of cells. The ADF volume measures the sum of shared species between a specified cell and all other cells, or, equivalently, the sum of the range sizes of all species present within the focal cell<sup>26</sup>. Thus, if  $p_{i,j}$  refers to the  $i$ th row and  $j$ th column of  $\mathbf{P}$ , i.e., the presence of species  $i$  in cell  $j$ , and  $r_i$  is the range size of species  $i$ , defined as

$$r_i = \sum_{j=1}^N p_{i,j}, \quad (2)$$

then the dispersion field volume is

$$\text{df\_volume}_j = \sum_{i=1}^S p_{i,j} r_i, \quad (3)$$

where  $S$  is the total number of species and  $N$  is the total number of grid cells. It also quantifies the mean contribution of a cell to all ADFs, since cells with a higher ADF volume share more species with surrounding cells. Dividing the dispersion field volume with the species richness of sites yields the mean range of all species present at the site, what Smith<sup>47</sup> labeled the mean average cosmopolitanism of species.

**Measurement of ADF asymmetry.** The center of gravity of the dispersion field for a grid cell  $j$  was calculated as the point located at the average longitude ( $x$ ) and latitude ( $y$ ) of the ranges of all species present within cell  $j$ .

So the mean  $x$  and  $y$  coordinates of species  $i$  is

$$\bar{x}_i = \frac{1}{r_i} \sum_{j=1}^N p_{i,j} x_{i,j} \quad (4)$$

$$\bar{y}_i = \frac{1}{r_i} \sum_{j=1}^N p_{i,j} y_{i,j} \quad (5)$$

and then

$$\text{df\_center}X_j = \frac{1}{s_j} \sum_{i=1}^S p_{i,j} \bar{x}_i \quad (6)$$

$$\text{df\_center}Y_j = \frac{1}{s_j} \sum_{i=1}^S p_{i,j} \bar{y}_i \quad (7)$$

Although the contours and symmetry of ADFs exhibit marked geographic variation, the total extents of ADFs (i.e., areas that share at least one species with the focal cell) are similar because most assemblages of birds, mammals, and amphibians contain a few widespread species that occur over most of the South American continent. For example, the Neotropical cormorant (*Phalacrocorax brasilianus*) occurs in every cell ( $n = 1689$ ) in South America, so that all focal cells shared this species with all other grid cells. The pervasive influence of widespread species forces the center of gravity of ADFs toward the geographic center of South America and obscures regional signals of biotic similarity. This is particularly true in mammals (mean range size = 182.7 grid cells) and birds (mean range size = 186.0 grid cells) but less so in amphibians (mean range size = 37.5 grid cells).

To visualize the distinct pattern in ADFs, we trimmed away grid cells that shared only a few species with the focal cell. The cutoff value for trimming represents a balance between the ability to visualize core ADF asymmetries and the need to illustrate the distributional pattern of the entire species assemblage for a focal cell. Higher cutoff values reflect increasingly restrictive patterns of biotic similarity. We arbitrarily set the value at 0.5 for amphibians and increased the value to 0.6 for birds and 0.7 for mammals, which have larger average range sizes. We then calculated the center of mass and the ADF asymmetry vector for the trimmed ADF for each focal cell. The resulting spatial patterns of ADFs and asymmetry diagrams were qualitatively similar for all three vertebrate classes (as seen in Supplementary Figs. 4–6).

**Null and predictive models.** We created null models by randomly placing continuous ranges on the South American continent, using a simple spreading dye algorithm<sup>30</sup>. This algorithm starts by selecting a random starting cell anywhere on the continent. The algorithm then selects a random cell adjacent to that occupied and adds it to the range. This process continues until the target range size is reached. The algorithm is implemented in R.

To create the null model ADFs, we built a null distribution of random ranges for each grid cells. Ranges were selected from all possible random ranges in South America that intersect the focal cell, maintaining the empirical range–size frequency distribution of that grid cell. This was achieved by first building a sampling population of random ranges for each possible range size between 2 and 1689 grid cells over the entire domain (the continent), so that all grid cells were intersected by at least 100 different random ranges of each range size. For each grid cell, we then picked random ranges from this sampling population according to the empirical range–size distribution of that grid cell and constructed the ADF.

The resulting ADFs clearly demonstrated an effect of the geometry of the South American continent, which was most pronounced for the lowest regions of ADFs (Supplementary Fig. 7a, b). When focusing on the areas with more shared species

(yellow to red in the figures), the dispersion fields were highly symmetrical, with biotic similarity declining as a basic function of distance in all directions from the focal cell. The ADF symmetry diagram exhibited a simple pattern of all vectors pointing toward the center of the continent (Supplementary Fig. 7c). These patterns represent the baseline for comparing the results of ecological process.

To quantify the deviance of empirical ADFs from the null model, we calculated standardized effect sizes by repeating the sampling procedure 1000 times and calculating ADF symmetry values for all grid cells. The deviation between the empirical and simulated symmetry values was then calculated as (Eq. 7)  $(v - \mu(\lambda))/\sigma(\lambda)$ , where  $v$  is the empirical symmetry value,  $\lambda$  is a (mathematical) vector of symmetry values, and  $\mu$  and  $\sigma$  are the mean and standard deviation of the simulated symmetry values, respectively.

We also implemented three different predictive models for the expected ADF symmetry patterns expected under various constraints on range origination and expansion. In model 1, constrained origination, we placed the first grid cell of ranges on the continent according to a probability function of available energy and allowed range expansion into each surrounding grid cell to be random. To generate model 3, constrained range expansion, we placed the first grid cell on a random grid cell in a vegetation ecoregion chosen with a probability proportional to the occupancy of that ecoregion by a given species. Random range expansion was allowed, but a range expansion event into a grid cell across a boundary was 30× less likely than an expansion event into a grid cell in the same ecoregion (the number 30 was chosen to ensure that species' ranges expanded across boundaries occasionally, but rarely). Sampling from these models followed the protocol outlined above. Model 2, constrained species richness, was generated in a different way, which was designed to maintain the grid cell richness constant, reflecting a scenario where the richness of each grid cell is restricted to a maximum value by local resources. All species ranges were built simultaneously, at each step allowing one species to expand its range by one grid cell adjacent to its range, with a probability equal to the difference between the maximum (i.e., the empirical) species richness of the grid cell and the current richness at that point in the simulation. This allows building random ranges while both maintaining range cohesion and the empirical species richness. At the end of the simulation, absolute range cohesion would become impossible, and small gaps were allowed (as exists in the empirical ranges as well). ADF diagrams were built for all three null models (Fig. 5).

**Ecotone analysis for Amazonia.** It is a prediction of niche conservatism that the asymmetry of ADF vectors for cells within a major vegetation ecoregion should decrease with distance from the ecotone boundary. Because the scale of the distribution data ( $1^\circ \times 1^\circ$  latitude/longitude degrees) was too coarse to adequately assess this effect for many of the smaller ecoregion, e.g., along the Andes, we tested the conjecture for the largest coherent region, Amazonia. To avoid the confounding influence of disjunct faunas that come into contact in grid cells at ecotones, we restricted the analysis to only include cells that are completely within the lowland (<500 m above sea level) Amazonian ecoregion and also excluded a small number of cells that were isolated behind the highland Tepuis of the Guianan Shield. The cells that were included in the analysis, along with their distance to the outer boundary of Amazonia, are shown in Supplementary Fig. 8. Residuals from the regression analysis were symmetrical but departed significantly from normality according to Kolmogorov–Smirnov test ( $p < 0.01$ ).

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

## Data availability

All data depicted in figures here are available on request from the authors. The underlying databases with range maps on amphibians and mammals are available from the IUCN. The result of computer models are also available upon request.

## Code availability

All R code necessary to produce the results here will be made available upon acceptance. In addition, R and Julia codes necessary to use the ADF general computational framework are available as part of the *nodiv* package for R (on CRAN) (functions *dfield\_matrix*, *dfield\_explore*, *dfield\_calc*, and *arrowplot*) and the *SpatialEcology.jl* package for Julia (can be installed from the official repository, function *dispersionfield*), both openly available on GitHub on the MIT free and open license.

Received: 7 July 2019; Accepted: 19 December 2019;

Published online: 24 January 2020

## References

1. Klopfer, P. H. & MacArthur, R. H. On the causes of tropical species diversity: niche overlap. *Am. Nat.* **95**, 223–226 (1961).

2. Simpson, G. G. Species density of North American recent mammals. *Syst. Zool.* **13**, 57–73 (1964).
3. Pianka, E. R. Latitudinal gradients in species diversity: a review of concepts. *Am. Nat.* **100**, 33–46 (1966).
4. Currie, D. J. Energy and large-scale patterns of animal-and plant-species richness. *Am. Nat.* **137**, 27–49 (1991).
5. Rahbek, C. & Graves, G. R. Multiscale assessment of patterns of avian species richness. *Proc. Natl Acad. Sci.* **98**, 4534–4539 (2001).
6. Rahbek, C. et al. Humboldt's enigma: what causes global patterns of mountain biodiversity? *Science* **365**, 1108–1113 (2019).
7. Currie, D. J. et al. Predictions and tests of climate-based hypotheses of broad-scale variation in taxonomic richness. *Ecol. Lett.* **7**, 1121–1134 (2004).
8. Rohde, K. Latitudinal gradients in species diversity: the search for the primary cause. *Oikos* **65**, 514–527 (1992).
9. Ricklefs, R. E. History and diversity: explorations at the intersection of ecology and evolution. *Am. Nat.* **170**, S56–S70 (2007).
10. Wiens, J. J. & Donoghue, M. J. Historical biogeography, ecology and species richness. *Trends Ecol. Evol.* **19**, 639–644 (2004).
11. Hawkins, B. A., Diniz-Filho, J. A., Jaramillo, C. A. & Soeller, S. A. Climate, niche conservatism, and the global bird diversity gradient. *Am. Nat.* **170**, S16–S27 (2007).
12. Hawkins, B. A., Diniz-Filho, J. A., Jaramillo, C. A. & Soeller, S. A. Post-Eocene climate change, niche conservatism, and the latitudinal diversity gradient of New World birds. *J. Biogeogr.* **33**, 770–780 (2006).
13. Rangel, T. F. L. V. B., Diniz-Filho, J. A. F. & Colwell, R. K. Species richness and evolutionary niche dynamics: a spatial pattern-oriented simulation experiment. *Am. Nat.* **170**, 602–616 (2007).
14. Weir, J. T. & Schluter, D. The latitudinal gradient in recent speciation and extinction rates of birds and mammals. *Science* **315**, 1574–1576 (2007).
15. Wright, D. H. Species-energy theory: an extension of species-area theory. *Oikos* **41**, 496–506 (1983).
16. Nakazawa, Y. Niche breadth, environmental landscape, and physical barriers: their importance as determinants of species distributions. *Biol. J. Linn. Soc.* **108**, 241–250 (2013).
17. Mayr, E. *Animal Species and Evolution* (Harvard University Press, Boston, Massachusetts, 1963).
18. Santos, J. C. et al. Amazonian amphibian diversity is primarily derived from late Miocene Andean lineages. *PLoS Biol.* **7**, e1000056 (2009).
19. Rabosky, D. L. & Hurlbert, A. H. Species richness at continental scales is dominated by ecological limits. *Am. Nat.* **185**, 572–583 (2015).
20. Grinnell, J. Field tests of theories concerning distributional control. *Am. Nat.* **51**, 115–128 (1917).
21. Gleason, H. A. The individualistic concept of the plant association. *Bull. Torrey Bot. Club* **53**, 7 (1926).
22. Peterson, A. T. et al. *Ecological Niches and Geographic Distributions (MPB-49)* (Princeton University Press, 2011).
23. Holt, B. G. et al. An update of Wallace's zoogeographic regions of the world. *Science* **339**, 74–78 (2012).
24. Graves, G. R. & Rahbek, C. Source pool geometry and the assembly of continental avifaunas. *Proc. Natl Acad. Sci.* **102**, 7871–7876 (2005).
25. Arita, H. T., Christen, J. A., Rodríguez, P. & Soberón, J. Species diversity and distribution in presence-absence matrices: mathematical relationships and biological implications. *Am. Nat.* **172**, 519–532 (2008).
26. Borregaard, M. K. & Rahbek, C. Dispersion fields, diversity fields and null models: uniting range sizes and species richness. *Ecography* **33**, 402–407 (2010).
27. Soberón, J. & Ceballos, G. Species richness and range size of the terrestrial mammals of the world: biological signal within mathematical constraints. *PLoS ONE* **6**, e19359 (2011).
28. Villalobos, F. & Arita, H. T. The diversity field of New World leaf-nosed bats (Phyllostomidae). *Glob. Ecol. Biogeogr.* **19**, 200–211 (2010).
29. Gotelli, N. J. & Graves, G. R. *Null Models in Ecology* (Smithsonian Institution Press, Washington D.C., USA, 1996).
30. Jetz, W. & Rahbek, C. Geometric constraints explain much of the species richness pattern in African birds. *Proc. Natl Acad. Sci.* **98**, 5661–5666 (2001).
31. Rahbek, C. et al. Predicting continental-scale patterns of bird species richness with spatially explicit models. *Proc. R. Soc. B Biol. Sci.* **274**, 165–174 (2007).
32. Ricklefs, R. E. Disintegration of the ecological community. *Am. Nat.* **172**, 741–750 (2008).
33. Rahbek, C., Hansen, L. A. & Fjeldsø, J. *One Degree Resolution Database for the Global Distribution of Birds* (University of Copenhagen, Zoological Museum, 2012).
34. Rahbek, C. & Graves, G. R. Detection of macro-ecological patterns in South American hummingbirds is affected by spatial scale. *Proc. R. Soc. B Biol. Sci.* **267**, 2259–2265 (2000).
35. Remsen, J. V. J. et al. A classification of the bird species of South America. <http://www.museum.lsu.edu/~Remsen/SACCWordFiles/SACCBaseline01.html> (2010).
36. Schipper, J. et al. The status of the world's land and marine mammals: diversity, threat, and knowledge. *Science* **322**, 225–230 (2008).
37. Fritz, S. A. & Purvis, A. Phylogenetic diversity does not capture body size variation at risk in the world's mammals. *Proc. R. Soc. B Biol. Sci.* **277**, 2435–2441 (2010).
38. Wilson, D. E. & Reeder, D. A. M. *Mammal Species of the World. A Taxonomic and Geographic Reference* 3rd edn (Johns Hopkins University Press, Baltimore, MD, 2005).
39. Jones, K. E. et al. PanTHERIA: a species-level database of life history, ecology, and geography of extant and recently extinct mammals. *Ecology* **90**, 2648–2648 (2009).
40. Bininda-Emonds, O. R. P. et al. The delayed rise of present-day mammals. *Nature* **446**, 507–512 (2007).
41. Fritz, S. A., Jönsson, K. A., Fjeldsø, J. & Rahbek, C. Diversification and biogeographic patterns in four island radiations of passerine birds. *Evolution* **66**, 179–190 (2012).
42. IUCN, Conservation International, and NatureServe. 2008. An Analysis of Amphibians on the 2008 IUCN Red List.
43. Frost, D. R. Amphibian species of the world: An online reference. Version 5.3 (Available at: <http://research.amnh.org/herpetology/amphibia/index.php>, accessed Jun, 2009).
44. Ficetola, G. F. et al. An evaluation of the robustness of global amphibian range maps. *J. Biogeogr.* **41**, 211–221 (2014).
45. New, M., Hulme, M. & Jones, P. Representing twentieth-century space-time climate variability. Part I: Development of a 1961–90 mean monthly terrestrial climatology. *J. Clim.* **12**, 829–856 (1999).
46. Connor, E. F. & Simberloff, D. S. The assembly of species communities: chance or competition? *Ecology* **60**, 1132–1140 (1979).
47. Smith, C. H. Areographic representation of faunal characteristics through a “second-order” relational approach. *Evol. Theory* **6**, 225–232 (1983).
48. Olson, D. M. et al. Terrestrial ecoregions of the world: a new map of life on Earth. *Bioscience* **51**, 933 (2001).

## Acknowledgements

All authors acknowledge the support of the Danish National Research Foundation for the Center for Macroecology, Evolution and Climate (grant DNRF96). M.K.B. was supported by a DFF – Individual Postdoctoral grant (grant number 0602-02109B) from the Danish Councils for Independent Research and a Marie-Sklodowska Curie IF (IDEA-707968) (for the past 2 years) when undertaking this work. G.R.G. was supported by the Smoketree Trust.

## Author contributions

M.K.B., G.R.G. and C.R. developed ideas and designed research. M.K.B. performed research and wrote the first draft. M.K.B., G.R.G. and C.R. wrote the manuscript.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary information** is available for this paper at <https://doi.org/10.1038/s41467-019-14267-y>.

**Correspondence** and requests for materials should be addressed to M.K.B.

**Peer review information** *Nature Communications* thanks Jorge Soberon and the other anonymous reviewers for their contribution to the peer review of this work.

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

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



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

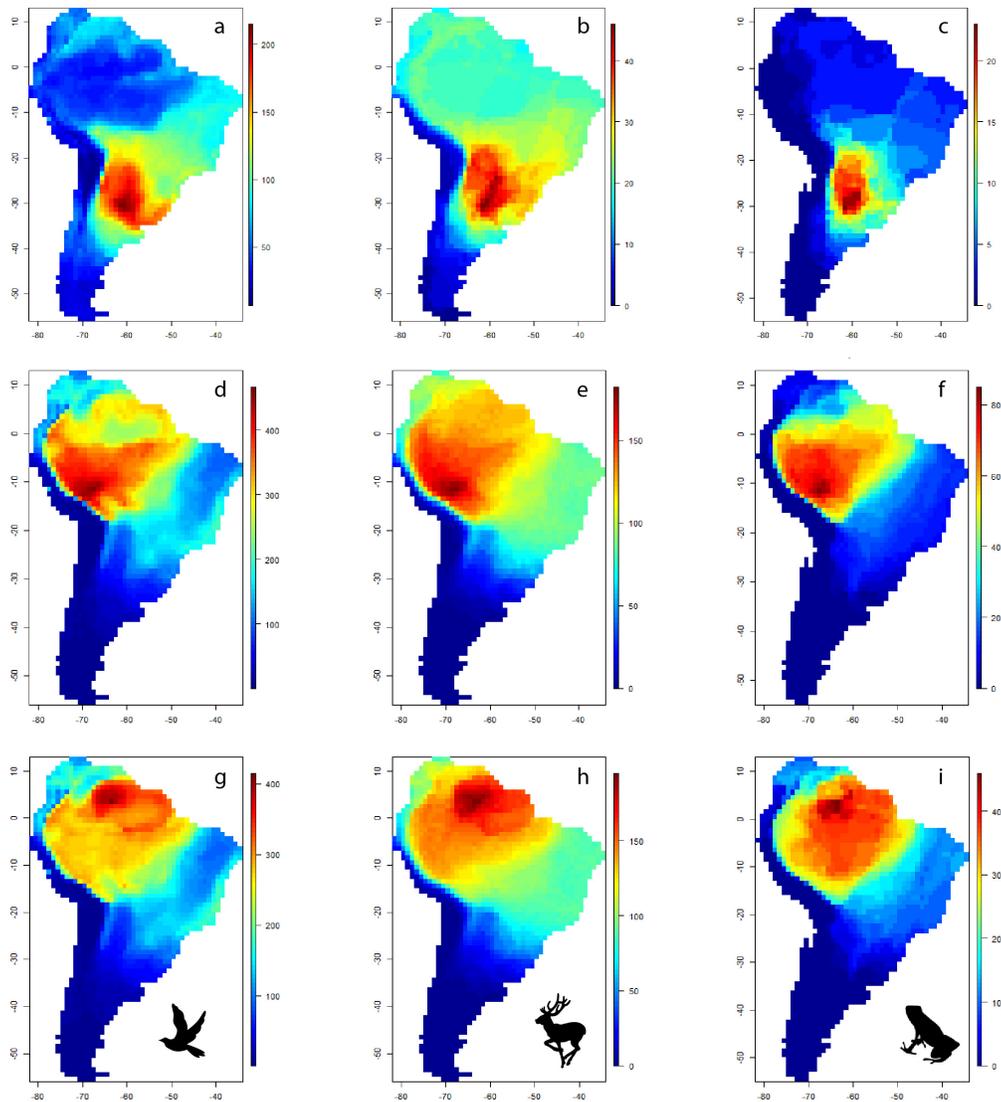
© The Author(s) 2020

“Assemblage dispersion fields reveal the spatial structure of biotic similarity among South American vertebrate distributions”

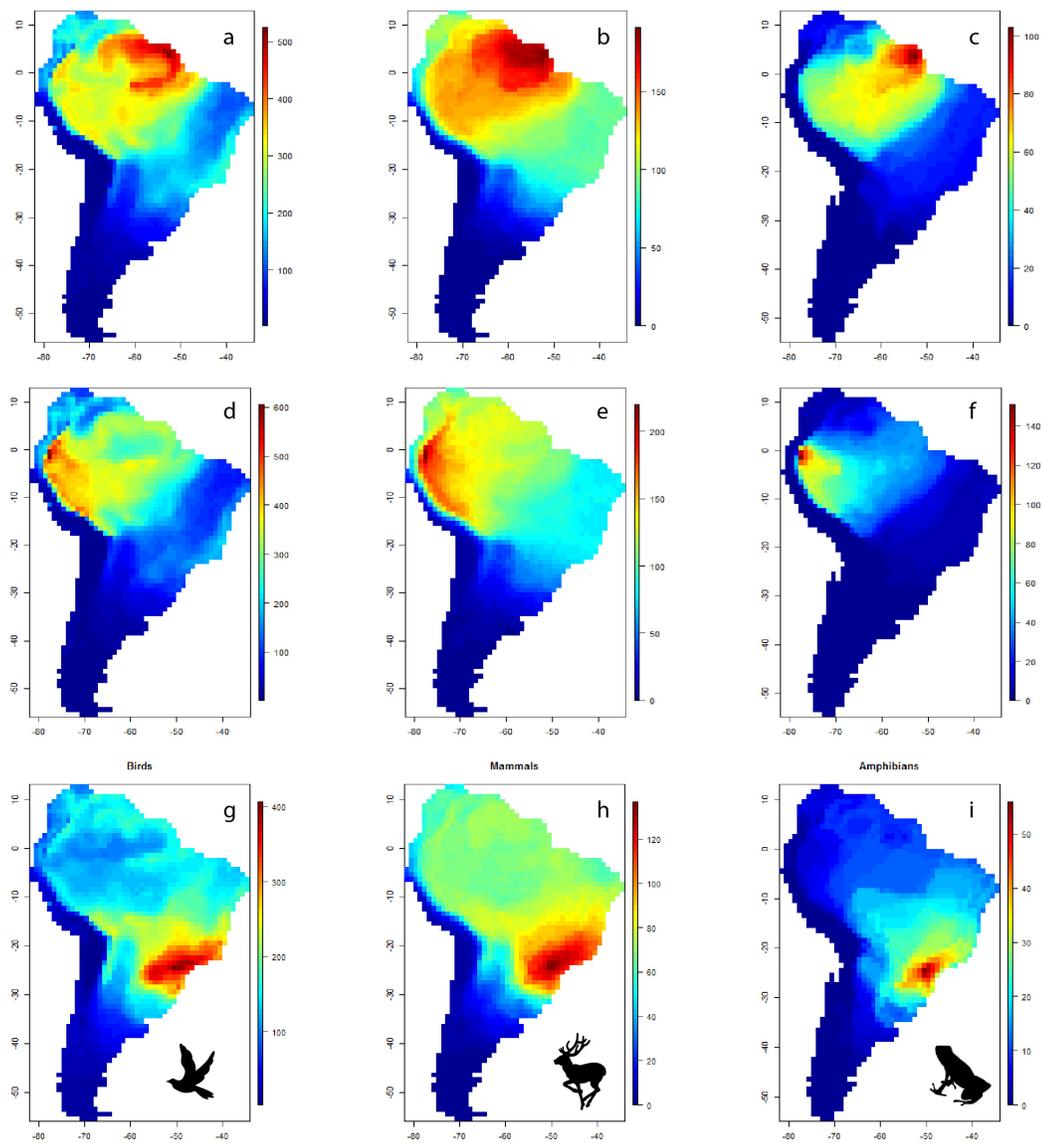
Borregaard, M.K., Graves, G. & Rahbek, C.

Submission to Nature Communications 2019

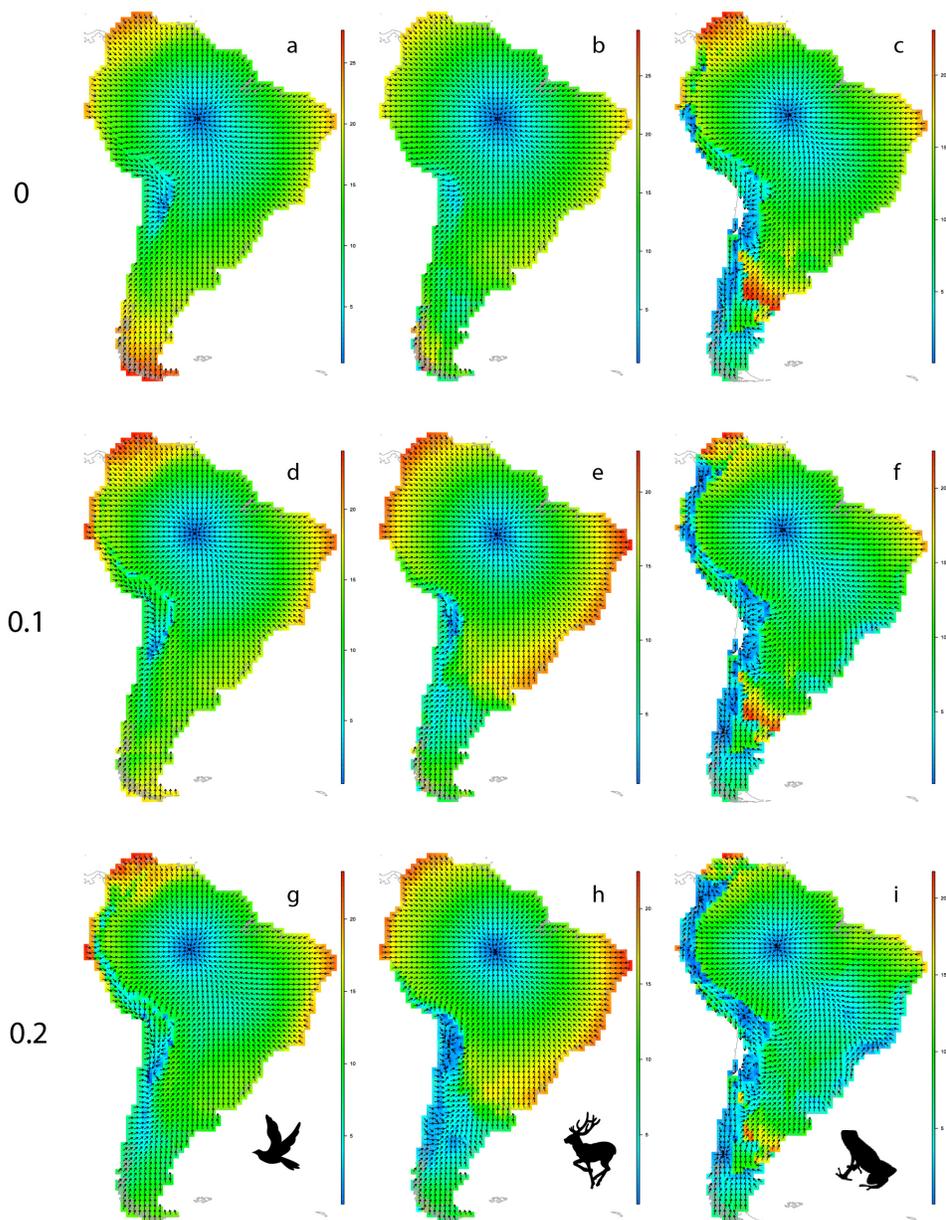
Supplementary Figures



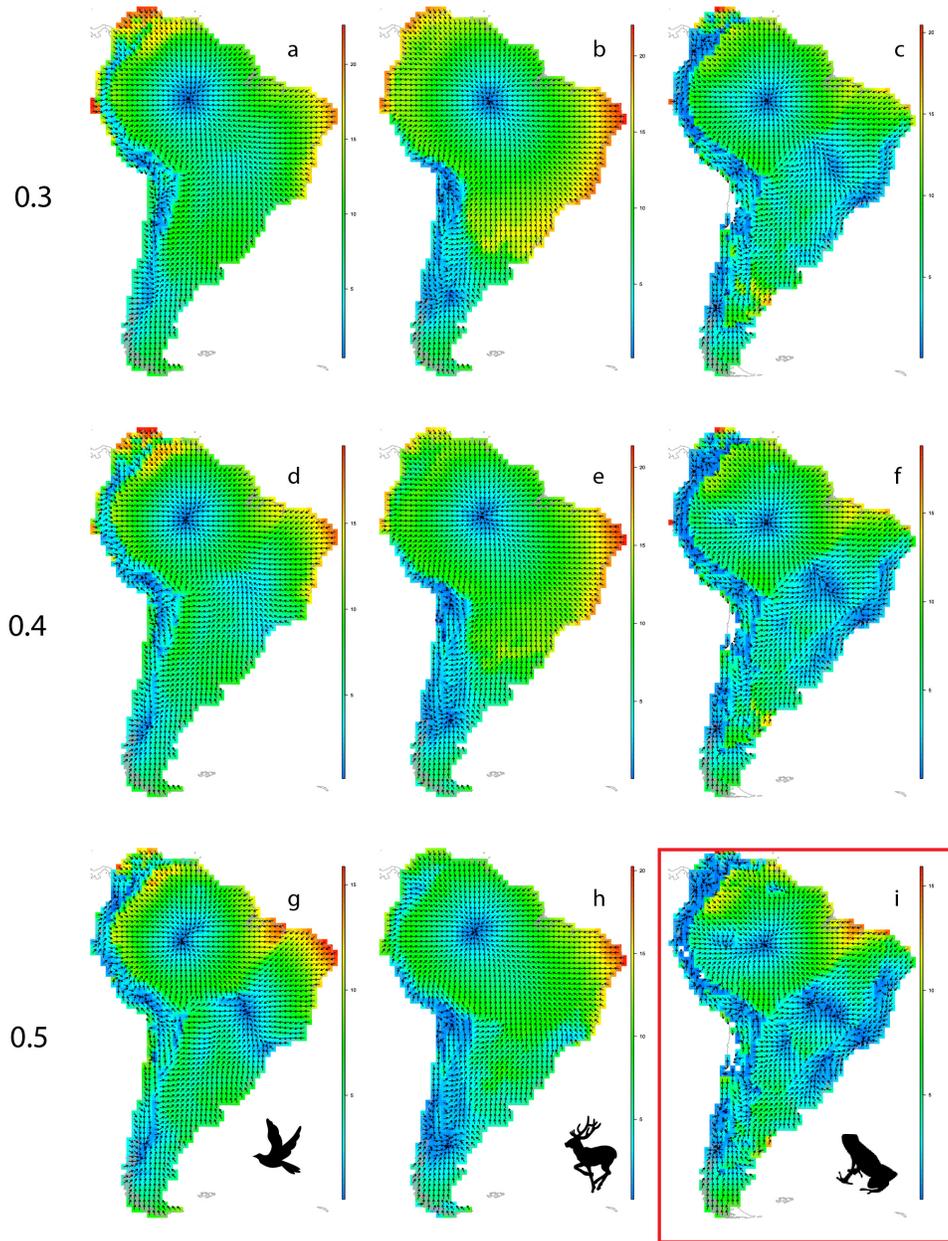
**Supplementary Figure 1. Examples of dispersion fields from different grid cells.** Each row shows the dispersion field for a given grid cell for birds, mammals and amphibians. The centroids of the focal cells are: row 1 (a-c): 61.5W 30.5S; row 2 (d-f): 63.5W 3.5N; row 3 (g-i): 68.5W 11.5S.



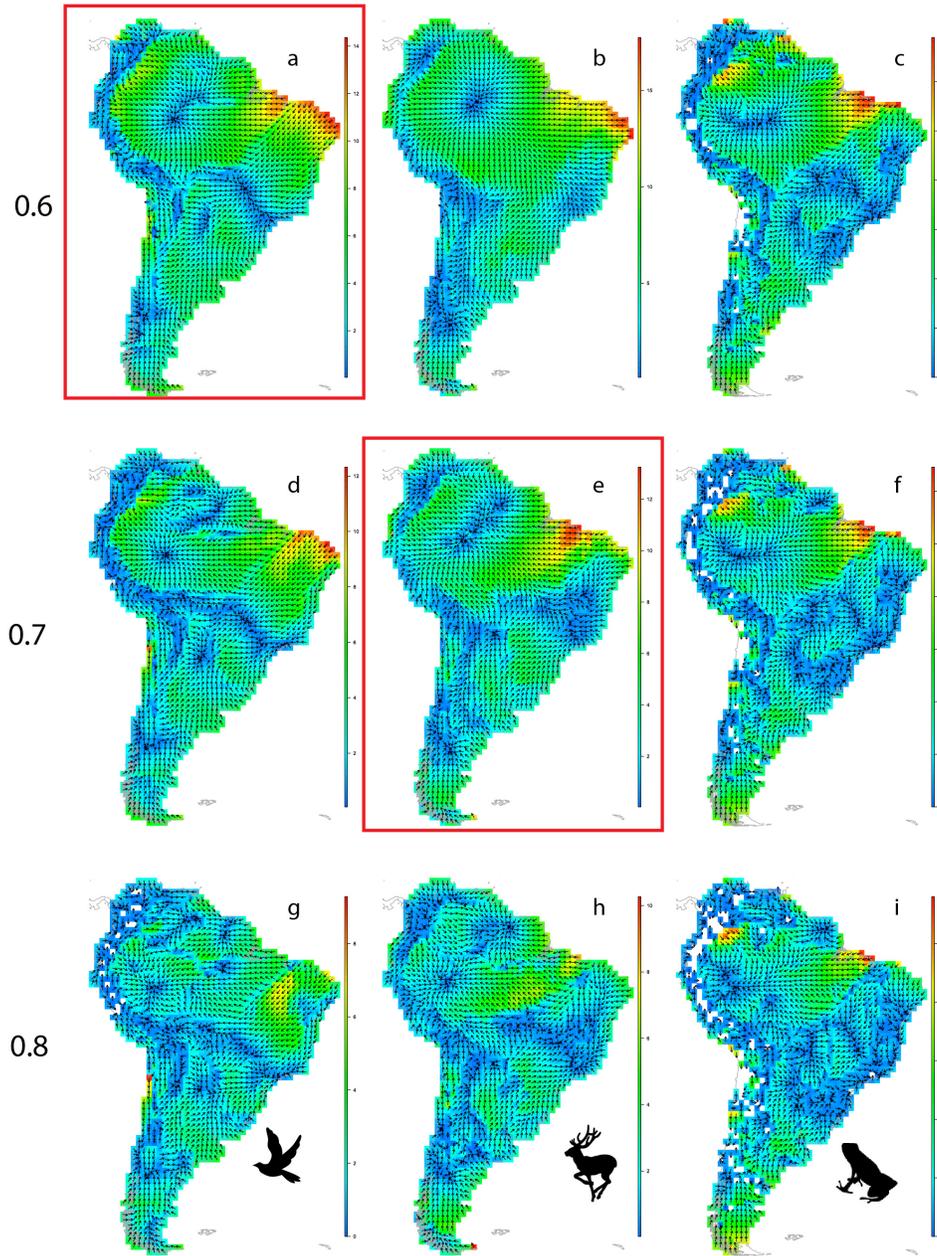
**Supplementary Figure 2. Further examples of dispersion fields from different grid cells.** Each row shows the dispersion field for a given grid cell for birds, mammals and amphibians. The centroids of the focal cells are: row 1 (a-c): 52.5W 3.5N; row 2 (d-f): 7.5W 1.5S; row 3 (g-i) 50.5W 24.5S.



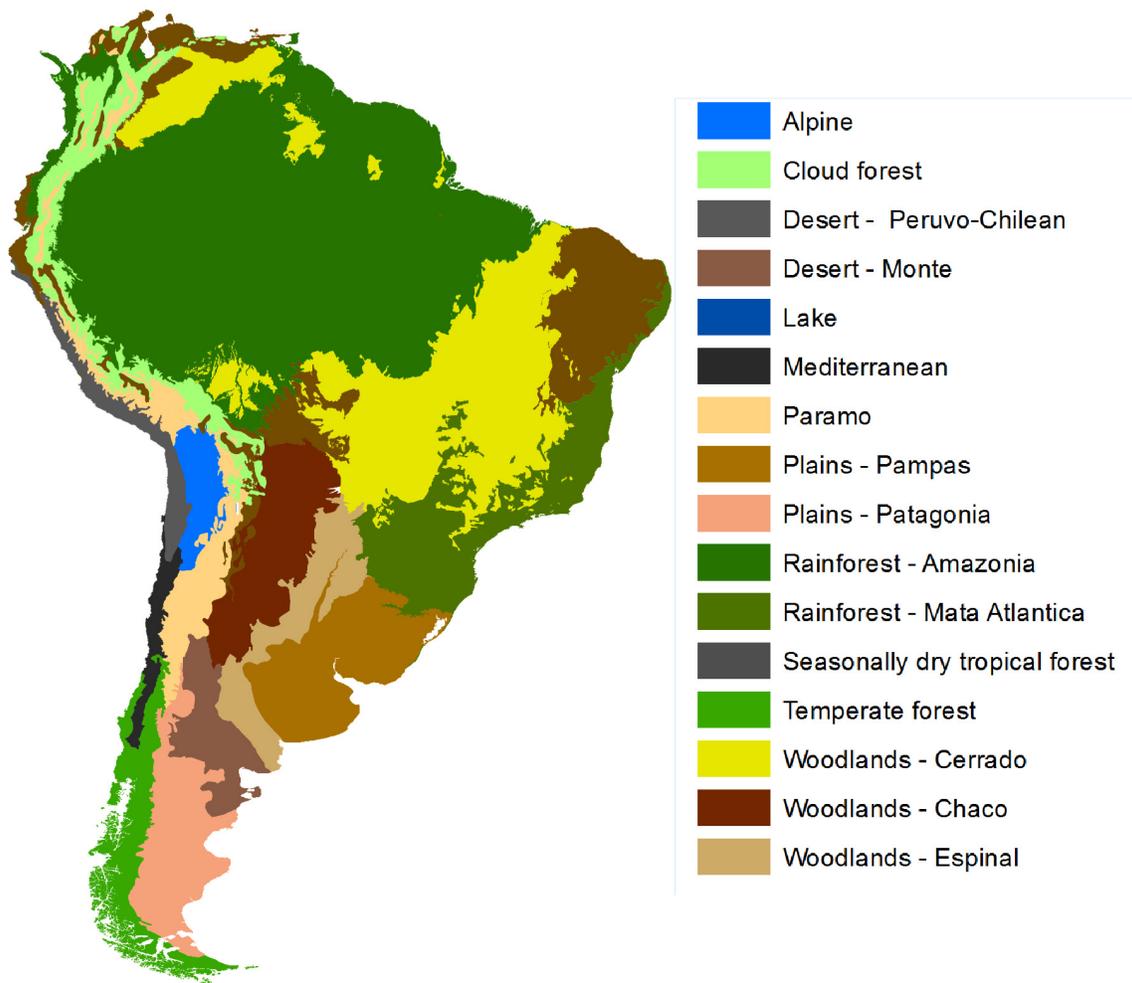
**Supplementary Figure 3. An illustration of the dependence of *ADF* symmetry diagrams on the cutoff value used to delimit the central region of the *ADF*.** The number at the left of each row indicates the cutoff value. A value of 0 indicates that all grid cells are used in calculating the centre of gravity of the *ADF*, whereas a value of 0.1 indicates that all grid cells that contain at least 10% of the species present in the focal cell are considered part of the central region. Each row shows the *ADF* symmetry diagrams for birds (a, d, g), mammals (b, e, h) and amphibians (c, f, i).



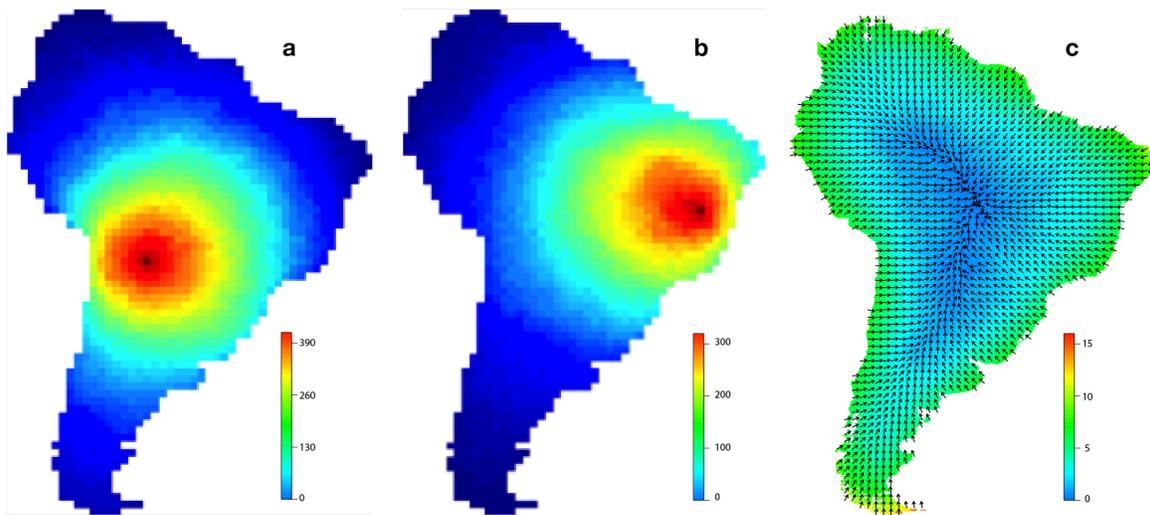
**Supplementary Figure 4. *ADF* symmetry diagrams for higher cutoff values.** The red box indicates the diagram that was included in the main text for this taxon. Legend as in Supplementary Fig 3.



**Supplementary Figure 5. ADF symmetry diagrams for the highest cutoff values.**  
 Legend as in Supplementary Fig 3.

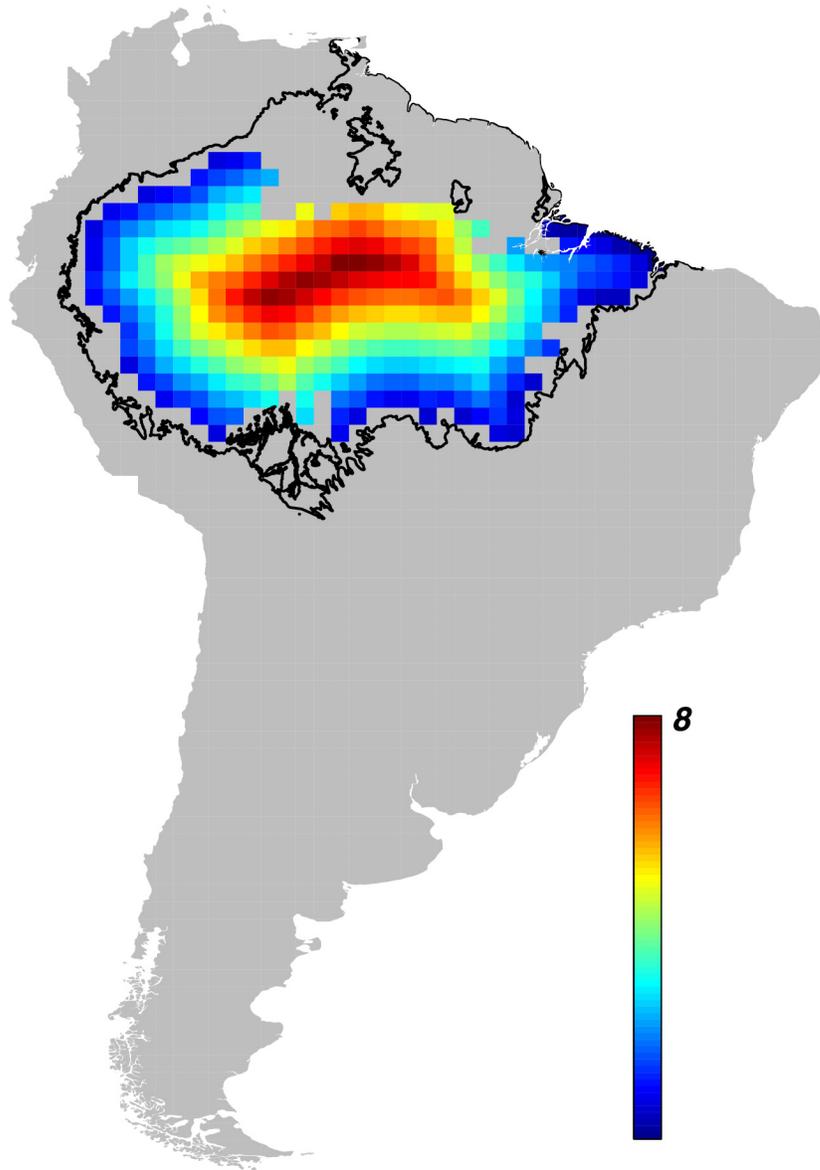


**Supplementary Figure 6. Larger map of South American biomes.** The map was provided by Tiina Särkinen, originally based on the map of the World's ecoregions developed by the WWF (under CC BY 4.0 license: <https://creativecommons.org/licenses/by/4.0/>).



**Supplementary Figure 7. Results from a basic spreading dye null model.**

(a-b) Examples of dispersion fields created by the spreading dye model. Biotic similarity is an approximate function of distance, similarly to the predictions of the species-energy model. (c) The *ADF* symmetry diagram expected under a null model of cohesive, non-interacting ranges. All symmetry vectors are oriented towards the centre of the continent.



**Supplementary Figure 8. The cells used for the analysis of the correlation between ecotone distance and *ADF* symmetry.** The ecotone around Amazonia is shown as a black line. All grid cells intersecting the ecotone are excluded, as well as all cells with a maximum altitude over 500meters, which marks a high-altitude fauna that is distinct from the Amazonian.