

LETTERS

Neutral metacommunity models predict fish diversity patterns in Mississippi–Missouri basin

Rachata Muneeppeerakul¹, Enrico Bertuzzo^{1,2}, Heather J. Lynch³, William F. Fagan³, Andrea Rinaldo^{2,4} & Ignacio Rodriguez-Iturbe¹

River networks, seen as ecological corridors featuring connected and hierarchical dendritic landscapes for animals and plants, present unique challenges and opportunities for testing biogeographical theories and macroecological laws¹. Although local and basin-scale differences in riverine fish diversity have been analysed as functions of energy availability and habitat heterogeneity², scale-dependent environmental conditions³ and river discharge^{4,5}, a model that predicts a comprehensive set of system-wide diversity patterns has been hard to find. Here we show that fish diversity patterns throughout the Mississippi–Missouri River System are well described by a neutral metacommunity model coupled with an appropriate habitat capacity distribution and dispersal kernel. River network structure acts as an effective template for characterizing spatial attributes of fish biodiversity. We show that estimates of average dispersal behaviour and habitat capacities, objectively calculated from average runoff production, yield reliable predictions of large-scale spatial biodiversity patterns in riverine systems. The success of the neutral theory in two-dimensional forest ecosystems^{6–8} and here in dendritic riverine ecosystems suggests the possible application of neutral metacommunity models in a diverse suite of ecosystems. This framework offers direct linkage from large-scale forcing, such as global climate change, to biodiversity patterns.

The Mississippi–Missouri River System (MMRS) is an invaluable resource of great biotic diversity, including freshwater fish. Its vast extent spans diverse habitat types operating under varying environmental conditions (such as climate, hydrological regime, primary productivity and human disturbance); these diverse habitats are connected to each other by one river network. An analysis that adequately captures major spatial biodiversity patterns in such a system is therefore noteworthy.

In recent years, the neutral theory of biodiversity⁶, with its minimal set of assumptions and parameters, has proven both influential^{7–11} and controversial^{12–14} as an explanation of biodiversity patterns. However, the theory has been tested mainly with ecosystems in two-dimensional landscapes or a mean-field context, to which spatial aspects contribute only weakly^{6–8,14–17}. Only recently have the contributions of landscape spatial structure¹⁸, for example, to biodiversity patterns in river networks¹, been investigated. Furthermore, implications of hydrological controls placed by river networks as ecological corridors have recently been explored^{19,20}. Here we analyse a large database of fish diversity in the MMRS to compare empirical biodiversity patterns against those predicted by a neutral metacommunity model (see Methods). The data analysis provides significant insights in its own right, and the comparison with model results allows us to investigate the extent to which a

neutral model captures observed patterns and extends inferences from the database.

In the following analysis, the 824 direct tributary areas comprising the MMRS are populated with occurrence data of 433 freshwater fish species from a database compiled by NatureServe²¹ (see Methods). Here, a direct tributary area (DTA) is a geographical region directly draining to a group of streams (that is, not including areas upstream of it); the DTAs correspond to the United States Geological Survey (USGS) HUC8-scale sub-basins as defined in US National Hydrography Database Plus²² (NHDPlus; see also Methods). Occurrence data and river network structure can be combined and analysed for several biodiversity patterns. We consider three patterns: first, the distribution of local species richness (LSR), or α diversity; second, species occupancies; and third, between-community (β) diversity. LSR is the number of species found in a randomly selected DTA. The occupancy of a species, in this case, is simply the number of DTAs in which that species is reported as present. To characterize β diversity, we consider the overall spatial decay of Jaccard's similarity index⁷ (JSI). JSI of any pair of DTAs is defined as $S_{ij}/(S_i + S_j - S_{ij})$, where S_{ij} is the number of

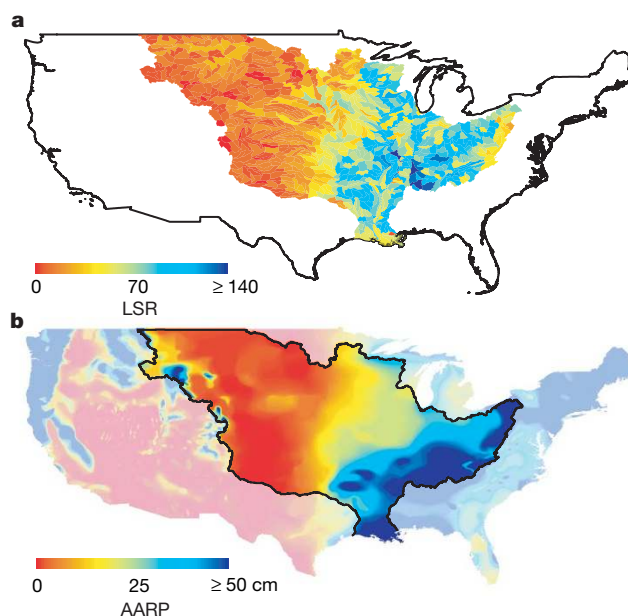


Figure 1 | Maps of freshwater fish diversity and AARP in the MMRS. **a**, Local species richness (LSR), or α diversity, of the freshwater fish in each DTA (that is, at the USGS HUC8 scale; see the text) of the MMRS. **b**, The AARP of the MMRS.

¹Department of Civil and Environmental Engineering, E-Quad, Princeton University, Princeton, New Jersey 08544, USA. ²Laboratory of Ecohydrology, Faculté ENAC, École Polytechnique Fédérale, CH-1015 Lausanne, Switzerland. ³Department of Biology, University of Maryland, College Park, Maryland 20742, USA. ⁴Dipartimento di Ingegneria Idraulica, Marittima, Ambientale e Geotecnica (IMAGE) and Centro Internazionale di Idrologia 'Dino Tonini', Università di Padova, via Loredan 20, I-35131, Padua, Italy.

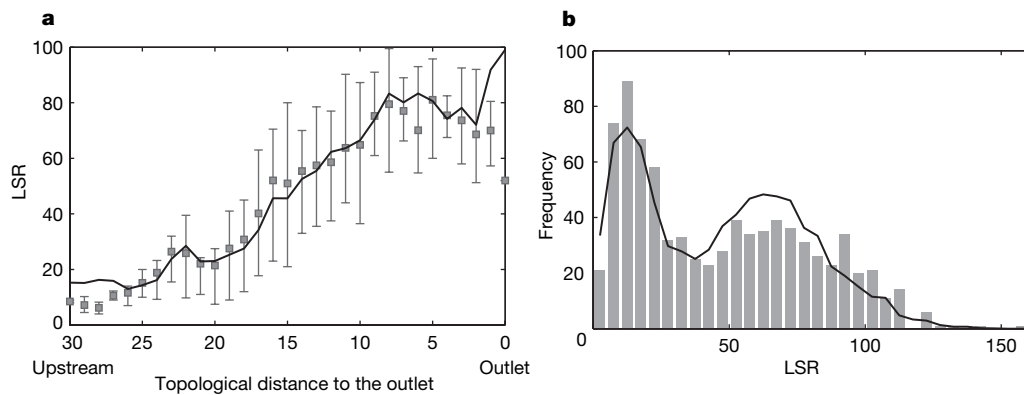


Figure 2 | Patterns of local species richness. **a**, LSR profile as a function of the topological distance (see Methods) from the outlet located in Atchafalaya, Louisiana (NHDPlus sub-basin number 08080101). **b**, Frequency distribution of LSR. The squares (average values) with error bars (ranging from the 25th to the 75th quantile) and bar plots represent the empirical data, and the lines represent the average values of the model results.

species present in both DTAs i and j , and S_i is the total number of species in DTA i . To achieve reliable statistics, we consider only topological distances (see Methods) for which more than 500 DTA pairs exist.

The map of LSR is shown in Fig. 1a. The DTA with the maximum LSR is Pickwick Lake (156 species) at the borders of the states of Alabama, Mississippi and Tennessee (NHDPlus sub-basin number 06030005). The map can be divided roughly into the western, species-poor half and the eastern, species-rich half. The sharp decrease in the species richness occurs around the 100° W meridian, which is also known to be the location of sharp gradients of annual precipitation²³ and runoff production^{24,25} (Fig. 1b). Although these gradients partly explain the arid climate and low fish diversity in the western half^{26,27}, we argue that the western DTAs are low in fish diversity both because their climate is dry and because they are upstream portions of the river network (see Supplementary Information). If they were located downstream, they might receive enough water supply and have access to a larger species pool from their wetter upstream sub-networks to maintain high fish-habitat capacities and fish diversity.

Figure 2a shows the LSR as a function of the topological distance from the network outlet. The distance zero corresponds to Atchafalaya, Louisiana (NHDPlus sub-basin number 08080101). The LSR profile shows a significant increase in the downstream direction, except at the very end in Louisiana, where we suggest that the freshwater fish-habitat capacities are significantly reduced by salinity, co-occurrence/intrusion by some freshwater-tolerant estuarine or coastal fish species, human disturbance and pollution. The overall downstream increase in richness results from the converging character of the river network²⁸ and is steepened by the dry-wet climatic gradient mentioned above (see Supplementary Information). Figure 2b presents the frequency distribution of LSR, whose two peaks at low and high values reflect the difference between the western and eastern halves of the MMRS.

The species occupancies are presented in Fig. 3 as a rank-occupancy curve, in which the fish species are ranked by their occupancies. The rank-occupancy curve (akin to the familiar rank-abundance curves⁶) yields a straight line on a semilogarithmic scale, a pattern reminiscent of the rank-abundance curves predicted by the neutral theory^{6,9,16}. Figure 4 shows that the JSI decreases as the topological distance between DTA pairs under consideration increases, an expected trend for β diversity. However, the JSI does not vanish even for DTA pairs that are very far apart. Such long-distance similarity in species composition is probably maintained by species with extremely large occupancies, for example *Ictalurus punctatus* (channel catfish), *Ameiurus melas* (black bullhead) and *Ameiurus natalis* (yellow bullhead).

As alluded to above, the neutral metacommunity model is a promising candidate for modelling the general spatial biodiversity patterns of the MMRS's freshwater fish. Here we show that by implementing the neutral model in the MMRS and incorporating the effect of average annual runoff production (AARP) on fish-habitat capacities, we can effectively reproduce a wide spectrum of observed

biodiversity patterns. For instance, in addition to the general trend and magnitude, the model also captures fine-structured fluctuations of the LSR profile (Fig. 2a). The fits to the LSR frequency distribution and β diversity pattern are also very good (Figs 2b and 4). The straight-line character of the rank-occupancy curves is evident for both the data and the model result (Fig. 3). Simultaneous fits of these diverse patterns (and others, such as species-area relationship) are a very stringent test for a model²⁹, especially a model with only four parameters as in this case (see Supplementary Information). The model also permits additional inferences to be drawn. The parameters corresponding to the best fits imply that the spread of the average fish species is quite symmetrical; that is, significantly biased in neither the upstream nor the downstream direction ($w_u = 1$; see Methods). The model results also suggest that, on average, most fish disperse locally (that is, to nearby DTAs) but a non-negligible fraction travel very long distances (see Supplementary Information).

Given the diverse environmental conditions covered by the MMRS, our demonstration that a simple neutral metacommunity model coupled with an appropriate habitat capacity distribution and dispersal kernel can simultaneously reproduce several major observed biodiversity patterns has far-reaching implications. These results suggest that only parameters characterizing average fish behaviour—as opposed to those characterizing biological properties of all different fish species in the system—and habitat capacities and connected structure suffice for reasonably reliable predictions of large-scale biodiversity patterns to be obtained. The neutral metacommunity model also provides a null model against which more biologically realistic models may be compared, and further developments in our understanding

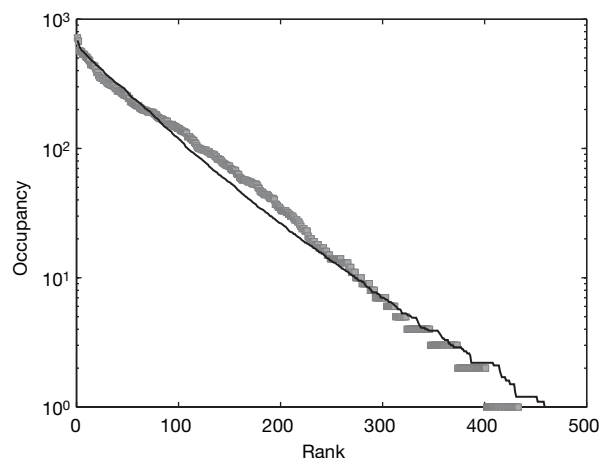


Figure 3 | Rank-occupancy curve. The squares (which are densely placed and appear as grey stripes) represent the data and the line represents the model result. Here, the occupancy of a freshwater fish species is simply the number of DTAs in which that species is reported as present. Note that the straight-line character in the semilogarithmic scale is shared by the familiar rank-abundance curves predicted by the neutral theory of biodiversity^{6,9,16}.

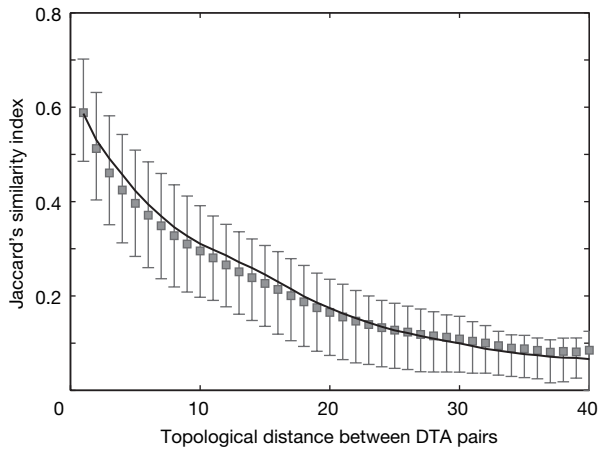


Figure 4 | The Jaccard's similarity index (JSI) as a function of topological distance between DTA pairs. The overall decay of JSI characterizes β diversity⁷; the squares with error bars represent the average values with the range between the 25th and 75th quantiles of the empirical data, and the line represents the average values of the model results. Note that the JSI does not vanish, even for widely separated DTA pairs.

of riverine networks and fish movement will permit a continued improvement between model and data. Indeed, although this modelling approach has been shown here to be useful for investigating key spatial patterns, it is crucial to recognize that “neutral pattern does not imply neutral process”⁹. Different approaches will therefore be necessary for predicting transient dynamics of the system or for understanding patterns and dynamics of specific species.

Finally, because mobile fish in a river network differ drastically from sessile trees in a forest, it is remarkable that the neutral theory can reproduce key biodiversity patterns of both sets of organisms quite well. This suggests that patterns predicted by the neutral metacommunity model—with appropriate habitat capacity distribution and dispersal kernel—may be broadly applicable across diverse ecosystems. It also offers a general, parsimonious modelling approach that acts as a coherent framework for studying several large-scale spatial biodiversity patterns simultaneously. This framework permits direct linkages to be made from various environmental changes to biodiversity patterns. For example, changes in precipitation patterns, perhaps as a result of global climate change, can now be mapped to changes in habitat capacities in the model; changes in connectivity among local communities, for example flow rerouting or damming in the case of fish, can be characterized by modifying the dispersal kernel. These linkages in turn enable us to make reliable predictions of a comprehensive set of altered biodiversity patterns, with significant implications for conservation campaigns and large-scale resource management.

METHODS SUMMARY

The biogeographical data on fish used in the analysis were obtained from the NatureServe²¹ database of US freshwater fish distributions, which summarizes museum records, published literature and expert opinion about fish species distribution in the United States, except Alaska, and is tabulated at the USGS HUC8 scale²². Owing to the present lack of availability of data, the Canadian portions of the MMRS are not included in the analysis, but we do not expect this to affect the key results and conclusions reported here. The data were then analysed to produce spatial biodiversity patterns (see Methods).

Our model is of a structured metacommunity type. The neutral theory of biodiversity is implemented in the MMRS, using its network as the structure of the metacommunity. Each DTA is a local community in that metacommunity and has a different fish-habitat capacity, H , defined as the number of ‘fish units’ sustainable by resources in that particular DTA; a fish unit can be thought of as a subpopulation of fish of the same species. H is assumed to be proportional to the product of the DTA watershed area and AARP²⁵, an indicator of the quantity of resources available for fish^{2,4}. The model uses the topological, rather than euclidean, distances between DTAs because they are representative of how far fish

travel. The model captures basic ecological processes: birth, death, dispersal, colonization and diversification. The simulations are run until the system reaches a steady state; the biodiversity patterns of interest are then determined and compared with the empirical patterns.

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

Received 14 November 2007; accepted 1 February 2008.

- Muneepeerakul, R., Weitz, J. S., Levin, S. A., Rinaldo, A. & Rodriguez-Iturbe, I. A neutral metapopulation model of biodiversity in river networks. *J. Theor. Biol.* **245**, 351–363 (2007).
- Guégan, J.-F., Lek, S. & Oberdorff, T. Energy availability and habitat heterogeneity predict global riverine fish diversity. *Nature* **39**, 382–384 (1998).
- Angermeier, P. L. & Winston, M. R. Local vs. regional influences on local diversity in stream fish communities of Virginia. *Ecology* **79**, 911–927 (1998).
- Oberdorff, T., Guégan, J.-F. & Hugué, B. Global scale patterns of fish species richness in rivers. *Ecography* **18**, 345–352 (1995).
- Xenopoulos, M. A. & Lodge, D. M. Going with the flow: using species-discharge relationships to forecast losses in fish diversity. *Ecology* **87**, 1907–1914 (2006).
- Hubbell, S. P. *The Unified Neutral Theory of Biodiversity and Biogeography* (Princeton Univ. Press, Princeton, NJ, 2001).
- Condit, R. *et al.* Beta-diversity in tropical forest trees. *Science* **295**, 666–669 (2002).
- Volkov, I., Banavar, J. R., Hubbell, S. P. & Maritan, A. Neutral theory and relative species abundance in ecology. *Nature* **424**, 1035–1037 (2003).
- Purves, D. W. & Pacala, S. W. in *Biotic Interactions in the Tropics* (eds Burslem, D., Pinard, M. & Hartley, S.) 107–138 (Cambridge Univ. Press, Cambridge, 2005).
- Holt, R. D. Emergent neutrality. *Trends Ecol. Evol.* **21**, 531–533 (2006).
- Alonso, D., Etienne, R. S. & McKane, A. J. The merits of neutral theory. *Trends Ecol. Evol.* **21**, 451–457 (2006).
- McGill, B. J. A test of the unified neutral theory of biodiversity. *Nature* **422**, 881–885 (2003).
- Chave, J. Neutral theory and community ecology. *Ecol. Lett.* **7**, 241–253 (2004).
- Dornelas, M., Connolly, S. R. & Hughes, T. P. Coral reef diversity refutes the neutral theory of biodiversity. *Nature* **440**, 80–82 (2006).
- Etienne, R. S. & Olff, H. Confronting different models of community structure to species-abundance data: a Bayesian model comparison. *Ecol. Lett.* **8**, 493–504 (2005).
- Walker, S. C. & Cyr, H. Testing the standard neutral model of biodiversity in lake communities. *Oikos* **116**, 143–155 (2007).
- Volkov, I., Banavar, J. R., Hubbell, S. P. & Maritan, A. Patterns of relative species abundance in ecology. *Nature* **450**, 45–49 (2007).
- Economu, E. P. & Keitt, T. H. Species diversity in neutral metacommunities: a network approach. *Ecol. Lett.* **10**, 52–62 (2007).
- Campos, D., Fort, J. & Mendez, V. Transport on fractal river networks: Application to migration fronts. *Theor. Popul. Biol.* **69**, 88–93 (2006).
- Bertuzzo, E., Maritan, A., Gatto, M., Rodríguez-Iturbe, I. & Rinaldo, A. River networks and ecological corridors: reactive transport on fractals, migration fronts, hydrochory. *Wat. Resour. Res.* **41**, W04419 (2007).
- NatureServe. *Distribution of Native U.S. Fishes by Watershed* (www.natureserve.org/getData/dataSets/watershedHuucs/index.jsp) (2004).
- Seaber, P. R., Kapinos, F. P. & Knapp, G. L. U. S. Geological Survey, *Water Supply Paper 2294—Hydrologic Unit Maps* (<http://pubs.usgs.gov/wsp/wsp2294/html/pdf.html>) (1994).
- Dingman, S. *Physical Hydrology* 2nd edn (Prentice Hall, Upper Saddle River, New Jersey, 2002).
- Henning, D. *Atlas of the Surface Heat Balance of the Continents* (Gebrüder Borntraeger, Berlin, 1989).
- Gebert, W. A., Graczyk, D. J. & Krug, W. *Average Annual Runoff in the United States, 1951–80* (<http://aa179.cr.usgs.gov/metadata/wrdmeta/runoff.htm>) (1987).
- Master, L. L., Flack, S. R. & Stein, B. A. *Rivers of Life: Critical Watersheds for Protecting Freshwater Biodiversity* (Nature Conservancy, Arlington, VA, 1998).
- Matthews, W. J. *Patterns in Freshwater Fish Ecology* (Kluwer Academic, Norwell, MA, 1998).
- Fernandes, C. C., Podos, J. & Lundberg, J. G. Amazonian ecology: Tributaries enhance the diversity of electric fishes. *Science* **305**, 1960–1962 (2004).
- Grimm, V. *et al.* Pattern-oriented modelling of agent-based complex systems: lessons from ecology. *Science* **310**, 987–991 (2005).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank L. Master, who facilitated our access to the fish database, T. Mueller for help in data preparation, and S. Azele for comments. We acknowledge NatureServe²¹ as the author/compiler of the fish biogeography database on which IndicatorMatrix.txt (see Methods) is based and for authorizing us to include the matrix as part of the Supplementary Information. R.M., H.J.L., W.F.F. and I.R.-I. acknowledge the support of the James S. McDonnell Foundation.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to R.M. (rmuneepe@princeton.edu) or I.R.-I. (irodrigu@princeton.edu).

METHODS

Model simulations. Every DTA is assumed to always be saturated at its capacity; that is, no available resources are left unexploited. At each time step, a fish unit, randomly selected from all fish units in the system, dies and the resources that previously sustained the unit are freed and available for sustaining a new fish unit. With probability ν , the diversification rate, the new unit will represent a new species (the diversification is a rate per birth and is due to speciation, to external introduction of non-native species, or to immigration (and reimmigration) of a new species from outside the MMRS); with probability $1 - \nu$, the new unit will belong to a species already existing in the system (the MMRS). In the latter case, the probability P_{ij} that an empty unit in DTA i will be colonized by a species from DTA j is determined as follows (including the probability $1 - \nu$):

$$P_{ij} = (1 - \nu) \frac{K_{ij}H_j}{\sum_{k=1}^N K_{ik}H_k}$$

where K_{ij} is the dispersal kernel (see below), H_k is the habitat capacity of DTA k , and N is the total number of DTAs (here, $N = 824$). All the fish units in DTA j have the same probability of colonizing the empty unit in DTA i where the death took place. The reported model results are the average patterns after the system reaches a statistically steady state.

A dispersal kernel determines how the fish units move within the river network. Here, it is assumed to take the form of a combination of back-to-back exponential and Cauchy distributions; note that a combination of several theoretical dispersal kernels has been used to achieve a good representation of real dispersal kernels³⁰ (see also Supplementary Information). The dispersal kernel in this model can be expressed as

$$K_{ij} = C \left[a^{L_{ij}} + \frac{b^2}{L_{ij}^2 + b^2} \right]$$

where K_{ij} is the probability that a fish unit produced at DTA j arrives at DTA i after dispersal; C is the normalization constant; L_{ij} is the effective distance, defined as $ND_{ij} + w_u NU_{ij}$, where ND_{ij} and NU_{ij} are the numbers of downstream and upstream steps comprising the shortest path from DTA j to DTA i , and w_u is the weight factor modifying the upstream distance; $w_u > 1$ implies downstream-biased dispersal, thereby characterizing dispersal directionality; and a (less than 1) and b characterize the exponential and Cauchy decays, respectively. Here, C is determined numerically such that, for every DTA j , $\sum_i K_{ij} = 1$; that is, no fish can travel out of the network. At the upstream ends of the system this is obvious; it is also true at the downstream end of the system; namely the outlet to the Gulf of Mexico, a marine body that acts as a barrier to freshwater fish. Finally, the dispersal kernel of every species is assumed to be the same; this is perhaps a strong assumption because fish species obviously differ in their dispersal abilities. However, the 'functional equivalence' between species is a key way in which the neutral theory of biodiversity departs from classical ecological models. We assume the species equivalence to study just how good a fit the neutral

metacommunity model can produce to our data in the absence of detailed, species-specific information.

Average annual runoff production (AARP). Runoff is the portion of precipitation that is drained by the river network. It depends on precipitation, evapotranspiration and infiltration. The map in Fig. 1b is estimated from the streamflow data of small tributaries collected from about 12,000 gauging stations averaged over the period 1951–80. For details see ref. 25.

Direct tributary area (DTA). The DTAs in the present analysis correspond to the HUC8-scale sub-basins designated by the USGS (available from www.horizon-systems.com/nhdplus/index.php). Details of how their boundaries are designated are given in ref. 22.

Habitat capacity, H_i . Habitat capacity of DTA i , H_i , is determined by

$$H_i = C_H N \frac{\text{AARP}_i \times \text{WA}_i}{\sum_{i=1}^N (\text{AARP}_i \times \text{WA}_i)}$$

rounded to the nearest integer. WA denotes watershed area, N (which here is 824) the total number of DTAs, and C_H the estimate (due to rounding) of average habitat capacity in a DTA.

Topological distance. The topological distance is a measure of distance along the network. An increment in the topological distance occurs when one travels along the network and crosses from one DTA to another. In the present case, one unit of topological distance corresponds to a distance in the range 100–200 km.

Notes on supplementary data. Two matrices summarizing the data used in the analysis are provided in the Supplementary Information. The first matrix, *IndicatorMatrix.txt*, reports the occurrence data²¹ of each of the 433 fish species in each of the 824 DTAs included in the analysis. Its first column lists the identification numbers of the HUC8 sub-basins²² (DTAs in the present analysis). The remaining 433 columns consist exclusively of zeros and ones, representing the absence and presence of each species, respectively. No species names are given; they are not necessary for the analysis. The sum along each row (433 elements) therefore gives the local species richness (LSR) of the corresponding DTA, and the sum along each column (824 elements) gives the occupancy of the corresponding species. The second matrix, *TopologicalDistanceMatrix.txt*, reports the topological distance between each pair of the 824 DTAs, which we derived from the data available from ref. 22. Its diagonal elements are zeros: each DTA is at zero distance from itself. The rows and columns of this matrix correspond to the DTA (that is, the HUC8 sub-basin) numbers in the first column of *IndicatorMatrix.txt*. The outlet corresponds to DTA number 08080101, which is the 364th row of *IndicatorMatrix.txt*. These two matrices can thus be combined to produce the profile of LSR as a function of topological distance from the outlet (Fig. 1a) and the pattern of Jaccard's similarity index (JSI) as a function of topological distance between DTA pairs (Fig. 4).

30. Levin, S. A., Muller-Landau, H. C., Nathan, R. & Chave, J. The ecology and evolution of seed dispersal: a theoretical perspective. *Annu. Rev. Ecol. Syst.* **34**, 575–604 (2003).