



Drainage network position and historical connectivity explain global patterns in freshwater fishes' range size

Juan Carvajal-Quintero^{a,b,1}, Fabricio Villalobos^a, Thierry Oberdorff^b, Gaël Grenouillet^b, Sébastien Brosse^b, Bernard Hugueny^b, Céline Jézéquel^b, and Pablo A. Tedesco^b

^aLaboratorio de Macroecología Evolutiva, Red de Biología Evolutiva, Instituto de Ecología, A.C., El Haya, 91070 Xalapa, Veracruz, Mexico; and ^bResearch Unit 5174, Laboratoire Evolution et Diversité Biologique, CNRS, Institut de Recherche pour le Développement, Université Paul Sabatier, F-31062 Toulouse, France

Edited by Andrea Rinaldo, Swiss Federal Institute of Technology in Lausanne, Lausanne, Switzerland, and approved May 13, 2019 (received for review February 11, 2019)

Identifying the drivers and processes that determine globally the geographic range size of species is crucial to understanding the geographic distribution of biodiversity and further predicting the response of species to current global changes. However, these drivers and processes are still poorly understood, and no ecological explanation has emerged yet as preponderant in explaining the extent of species' geographical range. Here, we identify the main drivers of the geographic range size variation in freshwater fishes at global and biogeographic scales and determine how these drivers affect range size both directly and indirectly. We tested the main hypotheses already proposed to explain range size variation, using geographic ranges of 8,147 strictly freshwater fish species (i.e., 63% of all known species). We found that, contrary to terrestrial organisms, for which climate and topography seem preponderant in determining species' range size, the geographic range sizes of freshwater fishes are mostly explained by the species' position within the river network, and by the historical connection among river basins during Quaternary low-sea-level periods. Large-ranged fish species inhabit preferentially lowland areas of river basins, where hydrological connectivity is the highest, and also are found in river basins that were historically connected. The disproportionately high explanatory power of these two drivers suggests that connectivity is the key component of riverine fish geographic range sizes, independent of any other potential driver, and indicates that the accelerated rates in river fragmentation might strongly affect fish species distribution and freshwater biodiversity.

freshwater fishes | species distribution | global scale | connectivity | river networks

The factors that determine species' geographic range sizes are complex and interrelated, and disentangling this complexity represents a central concern in macroecology, biogeography, and conservation (1, 2). At broad geographical scales, the overlapping of species ranges throughout space and time determines the variation in species richness and structure of regional biotas from which local communities are assembled (3). This overlapping of species ranges ultimately drives the biodiversity patterns that we use as a primary source to define regions of high conservation importance (e.g., ref. 4). Further, species' range size is one of the most important criteria for assigning a species' conservation status [International Union for Conservation of Nature (IUCN) Red List classification (5)], given its negative relationship with extinction risk (2). Quantifying the determinants of range size is also pivotal for evaluating community sensitivity to anthropogenic environmental change (6) and predicting shifts in response to climate change (2, 7, 8). During the last decades, multiple ecological and evolutionary hypotheses have been proposed to explain the variation in species' range sizes (*SI Appendix, Table S1*), including intrinsic biological characteristics of species (e.g., niche breadth, body size, population abundance, dispersal ability), metapopulation dynamics (i.e., colonization and/or extinction dynamics), and current or historical environmental characteristics (e.g., habitat availability

and environmental variability) (1, 2). However, the factors and processes determining the size of species' geographic ranges at broad spatial scales are still poorly understood, as none has emerged as preponderant in explaining the extent of species' geographical distributions (9, 10). For terrestrial groups (mostly vertebrates and plants), climatic and topographic factors have been recently identified as important determinants of species' range size at continental or global scales, with widespread species having higher thermal tolerance and occurring in areas with higher current and historical climate variability and lower topographic heterogeneity (8, 11, 12). Although strictly freshwater species (i.e., obligate freshwater dispersal) also inhabit continental landscapes, the global or continental determinants of their range size variation have never been assessed and may greatly differ from those identified for terrestrial ones. Indeed, a growing body of evidence suggests that theories developed in open landscapes, such as terrestrial, may be inadequate to predict the properties of complex branching ecosystems, such as river networks (13, 14).

Strictly freshwater fishes are an ideal model to continue improving our knowledge about the factors and processes that determine species' geographic range sizes. Indeed, unlike vagile terrestrial organisms, movements and dispersal processes of freshwater fishes are constrained by the dendritic and isolated arrangements of riverine ecosystems at different spatial scales (15, 16). At the largest spatial scales, fish movements are limited

Significance

Species' geographic range size is a fundamental aspect of understanding and predicting changes in biodiversity patterns. Investigating the global drivers of geographic range size variation in freshwater fishes, we found clear evidence that current and historical connectivity are, by far, the main determinants of range size. More specifically, we found that, everything else being equal, species displaying basal position in the drainage network (i.e., lowland areas) and found in drainage basins that have had connections during Quaternary low-sea-level periods have larger range sizes than their counterparts. Our findings suggest that connectivity is the key component of riverine fish geographic range sizes. This may have important implications for evaluating the vulnerability of freshwater species to river fragmentation.

Author contributions: J.C.-Q., F.V., and P.A.T. designed research; J.C.-Q., F.V., T.O., G.G., S.B., B.H., C.J., and P.A.T. performed research; J.C.-Q. analyzed data; and J.C.-Q. and P.A.T. wrote the paper with assistance from all authors.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Published under the PNAS license.

¹To whom correspondence may be addressed. Email: juanchocarvajal@gmail.com.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1902484116/-DCSupplemental.

Published online June 17, 2019.

by their inability to cross oceans, high mountain ranges, or expansive lands (17). This implies, for instance, that geographic range expansions between drainage basins are restricted to geological and hydrological events, such as river captures (18) or the confluence of river systems during low-sea-level periods resulting from climatic changes (19). At smaller spatial scales (i.e., within drainage basins), fish movements are determined by a combination of biotic and abiotic factors, including species' dispersal capacities and behaviors (20), the degree of river network branching (15), the basin slope, and other barriers to dispersal (e.g., rapids and waterfalls) that vary longitudinally along the network (16) (Fig. 1). As a consequence, drainage basins are structured by gradients of hydrological connectivity, where the branching and type of habitats encountered by species depend on the species' position within the drainage network, determining the travel distances and dispersal costs for freshwater species (13, 21) (Fig. 1).

Here we identify the main drivers of geographic range size variation of riverine fishes at global and biogeographic scales. In addition, we explore the complex path system of interactions among these drivers that ultimately determines species' range size variation. To do so, we compiled the most comprehensive dataset available to date for riverine fish species distributions, including 8,147 species (i.e., 63% of all known strictly freshwater species; ref. 22) covering all continents. Using these distributions, we applied multilevel path models (MLPMs) to evaluate at a global scale the effect of several drivers encompassing the main explanations already proposed for the variation of species' geographic range size (*SI Appendix, Table S1*). This allowed us to determine the direct and indirect effects through which multiple drivers influence the geographic range size, while controlling for the effect of random factors (i.e., taxonomic relatedness and spatial dependence). We further examined the strength and consistency of these drivers and pathways among the different biogeographic realms of the world.

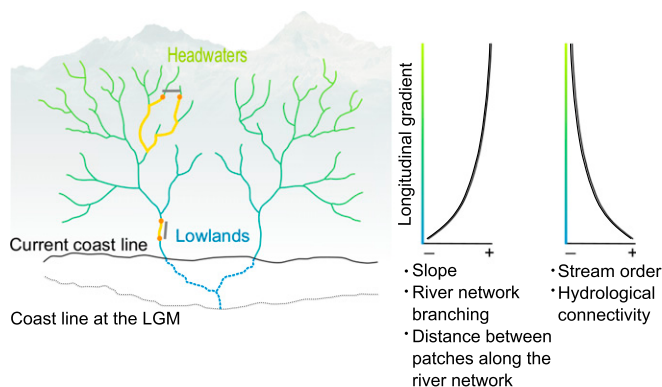


Fig. 1. Different features of hydrological connectivity across the longitudinal gradient of schematized river networks. Gradient solid lines represent two river drainages currently disconnected, but that formed a single paleobasin during a lower-sea-level period at the LGM (dashed blue lines). Solid black line shows the current seashore line and the dashed gray line the seashore during the LGM. In a downstream position of the river network, the branching degree is lower and the Euclidian distance between two localities (gray lines) is similar to the distance measured along the river network (yellow lines). As we move to more derivate positions toward headwaters, the dendritic branching increases and the Euclidian distance between two localities can be much shorter than the actual distance through the network (15). This increase in river branching toward headwaters is also accompanied by an increase in river slope that configures changes in habitats along a river drainage basin (16). This results in a longitudinal gradient of hydrological connectivity that determines the travel distances and dispersal costs for aquatic organisms. On the right side are graphically represented the hydrological connectivity features along the longitudinal gradient.

Results

The range sizes of freshwater fish species varied over six orders of magnitude, from 13 to 10,996,733 km², with a median of 77,322 km² (*SI Appendix, Fig. S1*). All MLPMs (for global and by biogeographic realms) yielded significant coefficients, indicating that the variation in geographic range size was well represented by our path models, and that no links among variables were missing (*SI Appendix, Tables S2–S8*). The R^2 values of range size for all MLPMs ranged between 0.739–0.909 and 0.758–0.921 for the marginal (R^2_m , fixed factors) and conditional (R^2_c , fixed plus random factors) variances, respectively (*SI Appendix, Table S9*).

Drainage network position (DNP; the average of the stream orders where a species occurs) and historical connectivity (a measure of past connections among drainage basins) were, by far, the most important drivers of range size variation in freshwater fish species at the global scale (Figs. 2 and 3), both with positive standardized path coefficients (SPC) followed by aridity and topographic heterogeneity showing negative coefficients. Glaciation history and body size were, respectively, the most important historical climatic and species biological trait variables associated with range size, both presenting a positive SPC. Other predictors (i.e., migratory behavior, swimming capacity, drainage basin area, and temperature anomaly and seasonality) showed the lowest coefficients, all of them being positive (Fig. 2). These general results remained stable when using diverse proxies for different predictors (*SI Appendix, Sensitivity Analysis*).

At the scale of biogeographic realms, four of the drivers found to be important at the global scale were also included in all MLP models: DNP, historical connectivity, topographic heterogeneity, and body size (Fig. 4 and *SI Appendix, Figs. S2–S7*), highlighting consistent results at both global and realm scales. DNP was again the most important range size predictor in all realms. Large-range species were related to higher values of DNP (i.e., located at downstream positions in the drainage network), historical connectivity, body size, and lower values of topographic heterogeneity. Productivity and long-term climatic stability affected range size differently across realms: negatively in Tropical realms (e.g., Neotropics) and positively in Temperate realms (e.g., Nearctic). Our measure of diversification (i.e., the number of species within the species' genus) had a direct and negative effect only in the Tropical realms (Fig. 4 and *SI Appendix, Figs. S2–S7*). When drainage basin area, migratory behavior, swimming capacity, and temperature seasonality were directly related to species range size, effects were always positive. Precipitation seasonality affected geographic range size indirectly mainly through the effect of other climatic and geomorphological variables (*SI Appendix, Figs. S2–S7*).

We found that predictors were highly interrelated at the global scale, affecting indirectly the geographic range size of freshwater fishes (Fig. 2 and *SI Appendix, Table S9*). For example, DNP was positively linked to drainage basin area and precipitation seasonality, and negatively with topographic heterogeneity and aridity (Fig. 2). Higher values of historical connectivity were related to high DNP, smaller drainage basin area, and lower topographic heterogeneity. Geomorphological predictors (i.e., DNP, topographic heterogeneity, and drainage basin area) were highly interrelated with all other predictor types (i.e., species' traits, climatic, and historical variables), whereas species' traits and climatic predictors mainly linked to predictors belonging to the same type (Fig. 2). At the realm scale, we found slight variations among predictors' relationships, mainly for drainage basin area, DNP, and aridity (Fig. 4 and *SI Appendix, Figs. S2–S7*). The relationships of these predictors with temperature and precipitation seasonality varied in their effect, being positive or negative, depending on the realm (Fig. 4 and *SI Appendix, Figs. S2–S7*). In general, the effect size of the relationships between predictors was high (*SI Appendix, Table S9*), resulting in complex

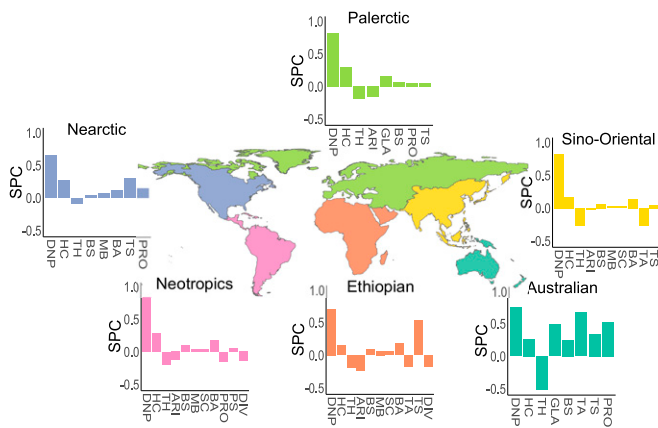


Fig. 4. SPC for each direct driver of geographic range size across the biogeographic realms proposed by Leroy et al. (42). Abbreviations for drivers are: drainage network position (DNP), historical connectivity (HC), topographic heterogeneity (TH), aridity (ARI), drainage basin area (BA), temperature anomaly (TA), glaciation history (GLA), temperature seasonality (TS), precipitation seasonality (PS), productivity (PRO), diversification (DIV), body size (BS), migratory behavior (MB), and swimming capacity (SC).

addition, seasonal flooding in lowland areas can connect previously unconnected habitats, leading to movement of organisms between locations that would not occur under base flow conditions (26). Meanwhile, changes in river slope and the direction of flow primarily determine the cost of upstream movements for strictly freshwater organisms along a river basin (27). Low river slopes in lowlands promote slow-running waters (i.e., low water velocity) characterized by wide channels and a high proportion of backwaters and pools, whereas in headwaters, streams have most often steeper slopes with torrential waters and higher portions of rapids and waterfalls (16, 28). The harsh conditions of headwaters also promote morphological and habitat specialization, resulting in the restriction of fish species distributions toward the headwaters (24). Conversely, the higher-connectivity conditions in lowlands and lower portions of rivers promote demographic connections among populations that are fundamental for species persistence and for their recovery from disturbances (29). All these factors create a hydrological connectivity gradient along the drainage network, which most probably explains the strong effect of the drainage network position on fish species ranges.

Among drainage basins, it was already found that historical connectivity has promoted fish colonization processes worldwide (19). Our measure of historical connectivity quantifies the extent of connectedness among basins during the last glacial maximum, when sea levels dropped up to 120 m and river mouths progressed through kilometers of exposed marine shelves before reaching the ocean (19, 30). This resulted in connections among previously isolated drainage basins that left an imprint on global biodiversity, where paleo-connected basins were richer and shared more species (as a result of colonization from other rivers within the same paleo-basin) than paleo-disconnected ones (19). Our findings show that such imprints on the biodiversity of river basins have been driven by the positive effect of historical connectivity in determining freshwater species' range sizes. Overall, our results suggest that lowland freshwater fish were the most efficient to expand their geographic range size, mainly because lowlands have higher levels of current and historical hydrological connectivity (Fig. 1).

Beyond the overall importance of drainage network position and historical connectivity, other factors also played a secondary role in determining freshwater fishes' range sizes. We found that topographic heterogeneity affects negatively species' range sizes. High topographic relief has long been recognized as imposing constraints on dispersal, resulting in high species turnover and

smaller range size for most animals (1), including on riverine fishes (24, 31). Furthermore, high altitudinal gradients imply less frequent drainage connections and fish species crossovers (31). Aridity was also a negative driver of species range size. In freshwater ecosystems, aridity fragments rivers' surface, dividing drainage basins in different pieces, which may result in a direct and negative effect on fish ranges by disrupting fish movements (32). Indirect effects of aridity on geographic range size may be mediated by the extrinsic effects of temperature and precipitation seasonality on aridity, which affect the water balance in riverine ecosystems, reducing basin areas and modifying the dendritic structure of river drainages (33).

Finally, species' traits related to dispersal ability (i.e., swimming capacity, migratory behavior, and body size) also affected freshwater fish range sizes, but with secondary importance. Better dispersers tend to have larger geographic ranges because they are able to sustain sink populations at large distances from source populations, whereas poor dispersers may lead to a larger proportion of potentially suitable habitats being unoccupied (9). This has been corroborated for freshwater fishes, for which greater dispersers and large-bodied species have larger geographic range sizes than poor disperser and small-bodied species (34). In addition, migratory behavior directly influences fish species range size, as reported for temperate freshwater fishes (34). Migratory behavior can also indirectly affect range size via dispersal ability and body size, because migrants tend to be better dispersers, which in turn increases range size (35), and may have larger body size (36).

To summarize, we found that the variation in geographic range sizes of freshwater fishes is jointly determined by the interaction of multiple predictors that create a complex path model, where drainage network position and historical connectivity are the most important predictors at both global and biogeographic realm scales (see *SI Appendix* for a detailed discussion about differences in minor drivers between realms). These results suggest that the geographic range size of freshwater fishes has been mainly shaped by the current and historical hydrological connectivity that determines the effort and distance of fish movements within a drainage basin, as well as the possibility of colonizing new basins during historical connections among basins resulting from sea level changes. Importantly, our results contrast with what has been observed for terrestrial and marine species for which connectivity has not been identified as a major driver of species' geographic range sizes (9, 37). It is therefore highly probable that the unique dendritic nature of river drainage basins, in which isolation can occur at much finer spatial scales than in other systems (38), generates unique dispersal processes.

The strong links that we found between range size and hydrological connectivity strengthen the vulnerability of freshwater species to fragmentation caused by damming (15, 39) and indicates that the accelerated rates in river fragmentation caused by the ongoing boom in dam construction (40) might strongly affect fish species distributions, which will likely have profound influences on fish diversity in the future.

Methods

Geographic Range Size. We compiled range maps for 9,075 species of freshwater fishes from two different sources. The IUCN Red List (<https://www.iucnredlist.org/>) provided range maps for 6,013 species worldwide, with the exception of a large portion of South America. We complemented this region covering the Amazon Basin and southern South America, using occurrence records of 3,062 species from different databases of freshwater fishes (see *SI Appendix* for further details on these datasets). To map these complementary ranges, we followed the same methodology as the IUCN, which consists of dissolving the HydroBASINS units or subbasins (41) where a species was present according to the occurrence records (*SI Appendix, Fig. S8A*). We calculated the species' range sizes as the extent of occurrence (km²) falling within the occupied subbasin areas (*SI Appendix, Fig. S8B*). We assigned each species' range to their native biogeographic realm: Neotropical, Ethiopian, Sino-Oriental, Nearctic, Palearctic, or Australian,

following Leroy et al. (42) (Fig. 4), on the basis of the midpoint of its latitudinal and longitudinal range.

Our final dataset of native ranges included 8,147 fish species, excluding island endemics and considering only strictly freshwater Actinopterygii species to ensure that all the analyzed species were restricted to freshwater environments and that their dispersion processes have been continental.

Drivers. On the basis of ecological theory and hypotheses proposed in previous studies, we developed a set of predictions regarding the potential drivers of the geographic range size of freshwater fishes (*SI Appendix, Table S1*).

Current climate. To represent current climate conditions, we measured three variables related to current climatic stability and climatic extremes. As a measure of present climatic stability, we used the average values of temperature and precipitation seasonality within the species' range (11, 12). For climatic extremes, we calculated the mean value of the Köppen aridity index (43). On a global scale, this aridity index is the best measure to describe water availability and identifies the most humid and arid regions (44). The original data on these climate variables were downloaded from WorldClim (45).

Long-term climatic changes. We measured long-term climatic changes as the mean temperature anomaly since the Last Glacial Maximum (LGM; 22 ky), encompassed by a species range. Temperature anomaly was calculated as the difference between the current mean annual temperature and mean temperature at the LGM. The current mean annual temperature was obtained from WorldClim (45), whereas the mean annual temperature at the LGM was calculated as the average of CCSM4 and MIROC-ESM (46) Paleoclimate models. Finally, we represented the LGM glaciation history by the proportion of overlapping area between species' range and the glacial extent at 21 ky before present (47).

Productivity. Our measure of within-range productivity was calculated as the mean net primary production. We obtained net primary production values from Zhao et al. (48) proposing a productivity metric that describes the growing season relationship between gross primary production and different respiration metrics.

Drainage network position. We measured the species' DNP as the average of the unique values of stream order (ref. 49; *SI Appendix, Fig. S9*) within the species range (e.g., a species occurring in stream orders 2–6 will have a DNP value of 4). The stream orders were obtained from Shen et al. (50). Stream order is a numeric measure of the river branching complexity, where increasing values describe a progressive downstream position in the dendritic structure and a lower branching (Fig. 1). As stream order decreases toward the headwaters, the dendritic branching structure becomes more complex (49). The stream order is highly related to other metrics also used to describe the species position in river networks [e.g., the "direct tributary area" used by Bertuzzo et al. (23)]. This longitudinal change in stream order also describes a gradient in the basin slope and habitats, with gentle slopes and high proportions of backwaters and pools for high stream order values, and steeper torrential waters mostly composed by rapids and waterfalls for low stream order values (16) (Fig. 1). We compared our polygon-based measure of occupied stream orders to the same measure based on occurrence records (*SI Appendix*) to control for any bias related to the potential inclusion of unoccupied streams within the polygons. These two ways of computing DNP resulted in very similar estimates ($R^2 = 0.72$).

Historical connectivity. As a measure of past connections among drainage basins, we focused on how sea-level changes reconfigured the connectivity between river systems during the LGM. Throughout the Quaternary, the Earth's climate fluctuated periodically, resulting in lower-sea-level periods (30) that allowed currently separated drainages to connect at their lower parts, making fish dispersal processes possible within these larger formed paleo-drainages (19) (Fig. 1). According to the paleo-drainages reconstruction proposed by Dias et al. (19), at the global scale, we derived a metric of historical connectivity as the number of basins in which a species currently occurs divided by the number of paleo-basins covered by that species range. This metric indicates to what extent currently occupied drainage basins were regrouped into larger connected paleo-drainages during lower-sea-level periods.

Geomorphology. We evaluated the effect of two geomorphological drivers of drainage basins on species range size: the area of the drainage basins occupied by each species and the topographic heterogeneity within their distribution range. The drainage basin area can be considered as the maximum surface extent that a freshwater fish species could potentially occupy, analogous to the continental extent applied in a similar analysis for terrestrial vertebrates (8). We measured this proxy of area availability as the mean drainage area of the basins where a species occurs. To measure topographic heterogeneity, we created a raster layer based on the variance of elevation

among each grid-cell and all other grid-cells within a 15-km buffer. High values of this measure represent high topographic heterogeneity between a grid-cell and its neighboring cells. We computed an overall topographic heterogeneity metric for each species as the mean value across all grid-cells that overlapped with the species range.

Diversification. We used the total number of species within each genus as a coarse proxy of the clade's diversification level that each species has experienced (51). Total species numbers by genus were obtained from FishBase (52).

Species traits. We used four traits related to locomotion ability, migratory behavior, energy demand, and trophic position (*SI Appendix, Table S1*) to evaluate their effect on fish range size. The maximum body length (mm) reported in FishBase (52) for each species was used as a measure of body size. The presence of migratory behavior (only potamodromous species in our case) for each species was also drawn from FishBase (52). Prey-capture ability and swimming capacity measures were calculated from morphological measurements available from Toussaint et al. (53). From this comprehensive morphological database of freshwater fishes, we used six traits (*SI Appendix, Table S10*) commonly used in the assessment of fish functional diversity (53–55). This database covered 93% ($\pm 0.03\%$) of the fish species considered here. All six traits were assigned to a species function (i.e., prey-capture ability or swimming capacity; *SI Appendix, Table S10*) and then ordered by a principal components analysis, using a regularized algorithm designed for ordination analysis that handles missing values (56). We retained the first axis of each principal components analysis (which accounted for >50% of the variance; *SI Appendix, Table S10*) to represent each species function.

All the distribution data and spatial variables mentioned were projected into the Behrmann equal-area cylindrical projection, and all rasters were rescaled to a resolution of 2.5 arc-minutes.

Data Analysis. We performed MLPMs (57) to identify the drivers of the geographic range size variation in freshwater fishes and how these drivers are related with each other. MLPMs allow moving beyond the estimation of direct effects and analyze the relative importance of different causal models, including direct and indirect paths of influence among multiple variables (57). To apply MLPMs, we used an integrative modeling approach that sequentially integrates a series of complementary procedures. We first assembled an expected path model to depict the expected relationships and interrelationships between the species range size and the multiple predictors, based on hypotheses previously proposed in the literature (*SI Appendix, Fig. S10 and Tables S1 and S11*). Next, we identified drivers of each endogenous (response) variable, using Multilevel and Generalized Multilevel Models (MLM and GMLM), in which we included genus, family, and order as random nested factors to account for the taxonomic relatedness among species. Residual spatial autocorrelation in regression models can lead to biased parameter estimates and P values. We found differences in the residuals among biogeographic realms ($P < 0.0001$), suggesting that the inclusion of realms as random effects could improve the parameter estimates of the models. Finally, we ran all the MLM and GMLM, including all possible combinations of the explanatory variables as fixed terms, based on the expected path model (*SI Appendix, Fig. S10*), including taxonomy and biogeographic realms as nested- and multilevel-random factors, respectively. All variables except migratory behavior and DNP were \log_{10} -transformed, and weak correlations ($R < 0.5$) among predictor variables were observed (*SI Appendix, Fig. S11*).

We performed multimodel inference based on information theory (58) to determine the average parameters from the MLM regressions. As a cutoff criterion to delineate a top model set, we used fitted models with $\Delta AICc \leq 2$ (58). Variance explained by each inferred model was estimated with marginal and conditional R^2 (59). Marginal R^2 (R^2_m) is concerned with variance explained by fixed terms, and conditional R^2 (R^2_c) with variance explained by both fixed and random factors.

We then combined the inferred MLMs to set the observed path model and test whether this model was consistent with our data, using the d-separation test (57). The d-separation test specifies the minimum set of independence and examines the validity of conditional independence statements that hold true among all variables in a given causal model. We tested the composite validity of all independence statements combining the P values through Fisher's C statistic and tested missing linkages, using the criterion that unlinked variables are conditionally independent (60). Hence, we obtained the residuals of the inferred models of each endogenous variable to examine relationships among those residuals and unlinked variables. For variables with no predictors (e.g., topographic heterogeneity), we used the raw values instead of the residuals (61). Because very large datasets can detect very minor residual associations between variables and lead models with very complex and nonsignificant scientific graphical relations, we only included

missing linkages of conditional statements with fixed effects sizes (R^2_m) > 0.1 and P values < 0.01 (ref. 61; *SI Appendix, Tables S2–S8*). To compare the relative strength of each causal, we calculated SPC of the causal linkages. Finally, we applied the above modeling approach for each biogeographic realm, considering only the species endemic to each realm.

Both data analyses and calculations of variables were performed in R 3.4.3 (62). For details on R packages used, see *SI Appendix, Table S12*.

- J. H. Brown, G. C. Stevens, D. M. Kaufman, The geographic range: Size, shape, boundaries, and internal structure. *Annu. Rev. Ecol. Syst.* **27**, 597–623 (1996).
- K. J. Gaston, *The Structure and Dynamics of Geographic Ranges* (Oxford University Press, 2003).
- N. J. Gotelli *et al.*, Patterns and causes of species richness: A general simulation model for macroecology. *Ecol. Lett.* **12**, 873–886 (2009).
- C. N. Jenkins, S. L. Pimm, L. N. Joppa, Global patterns of terrestrial vertebrate diversity and conservation. *Proc. Natl. Acad. Sci. U.S.A.* **110**, E2602–E2610 (2013).
- IUCN, The IUCN Red List of threatened species. (2018). <https://www.iucnredlist.org/en>. Accessed 29 January 2019.
- R. Ohlemüller *et al.*, The coincidence of climatic and species rarity: High risk to small-range species from climate change. *Biol. Lett.* **4**, 568–572 (2008).
- B. Sandel *et al.*, The influence of Late Quaternary climate-change velocity on species endemism. *Science* **334**, 660–664 (2011).
- Y. Li *et al.*, Climate and topography explain range sizes of terrestrial vertebrates. *Nat. Clim. Chang.* **6**, 498–502 (2016).
- S. E. Lester, B. I. Ruttenberg, S. D. Gaines, B. P. Kinlan, The relationship between dispersal ability and geographic range size. *Ecol. Lett.* **10**, 745–758 (2007).
- P. Calosi, D. T. Bilton, J. I. Spicer, S. C. Votier, A. Atfield, What determines a species' geographical range? Thermal biology and latitudinal range size relationships in European diving beetles (Coleoptera: Dytiscidae). *J. Anim. Ecol.* **79**, 194–204 (2010).
- F. J. S. Whitton, A. Purvis, C. D. L. Orme, M. Á. Olalla-Tárraga, Understanding global patterns in amphibian geographic range size: Does Rapoport rule? *Glob. Ecol. Biogeogr.* **21**, 179–190 (2012).
- N. Morueta-Holme *et al.*, Habitat area and climate stability determine geographical variation in plant species range sizes. *Ecol. Lett.* **16**, 1446–1454 (2013).
- E. H. C. Campbell Grant, W. H. Lowe, W. F. Fagan, Living in the branches: Population dynamics and ecological processes in dendritic networks. *Ecol. Lett.* **10**, 165–175 (2007).
- A. Rinaldo, R. Rigon, J. R. Banavar, A. Maritan, I. Rodriguez-Iturbe, Evolution and selection of river networks: Statics, dynamics, and complexity. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 2417–2424 (2014).
- W. F. Fagan, Connectivity, fragmentation, and extinction risk in dendritic metapopulations. *Ecology* **83**, 3243–3249 (2002).
- L. Benda *et al.*, The network dynamics hypothesis: How channel networks structure riverine habitats. *Bioscience* **54**, 413–427 (2004).
- L. R. Parenti, Ocean basins and the biogeography of freshwater fishes. *Aust. Syst. Bot.* **4**, 137–149 (1991).
- J. S. Albert, D. R. Schoolmaster, Jr, V. Tagliacollo, S. M. Duke-Sylvester, Barrier displacement on a neutral landscape: Toward a theory of continental biogeography. *Syst. Biol.* **66**, 167–182 (2017).
- M. S. Dias *et al.*, Global imprint of historical connectivity on freshwater fish biodiversity. *Ecol. Lett.* **17**, 1130–1140 (2014).
- J. Radinger, C. Wolter, Patterns and predictors of fish dispersal in rivers. *Fish Fish.* **15**, 456–473 (2014).
- J. D. Tonkin *et al.*, The role of dispersal in river network metacommunities: Patterns, processes, and pathways. *Freshw. Biol.* **63**, 141–163 (2018).
- P. A. Tedesco *et al.*, A global database on freshwater fish species occurrence in drainage basins. *Sci. Data* **4**, 170141 (2017).
- E. Bertuzzo *et al.*, On the geographic range of freshwater fish in river basins. *Water Resour. Res.* **45**, W11420 (2009).
- J. D. Carvajal-Quintero *et al.*, Variation in freshwater fish assemblages along a regional elevation gradient in the northern Andes, Colombia. *Ecol. Evol.* **5**, 2608–2620 (2015).
- D. Schmera *et al.*, Does isolation influence the relative role of environmental and dispersal-related processes in stream networks? An empirical test of the network position hypothesis using multiple taxa. *Freshw. Biol.* **63**, 74–85 (2018).
- A. Morán-Ordóñez *et al.*, Aquatic communities in arid landscapes: Local conditions, dispersal traits and landscape configuration determine local biodiversity. *Divers. Distrib.* **21**, 1230–1241 (2015).
- T. Datry *et al.*, Metacommunity patterns across three Neotropical catchments with varying environmental harshness. *Freshw. Biol.* **61**, 277–292 (2016).
- J. D. Allan, M. M. Castillo, *Stream Ecology: Structure and Function of Running Waters* (Springer Netherlands, ed. 2, 2007). <https://www.springer.com/gp/book/9781402055829>. Accessed 29 January 2019.
- R. K. Cowen, S. Sponaugle, Larval dispersal and marine population connectivity. *Annu. Rev. Mar. Sci.* **1**, 443–466 (2009).
- H. K. Voris, Maps of Pleistocene sea levels in Southeast Asia: Shorelines, river systems and time durations. *J. Biogeogr.* **27**, 1153–1167 (2000).
- G. R. Smith, Late cenozoic freshwater fishes of North America. *Annu. Rev. Ecol. Syst.* **12**, 163–193 (1981).
- P. J. Unmack, Biogeography of Australian freshwater fishes. *J. Biogeogr.* **28**, 1053–1089 (2001).
- R. Seager *et al.*, Projections of declining surface-water availability for the southwestern United States. *Nat. Clim. Chang.* **3**, 482–486 (2013).
- S. Blanchet *et al.*, Phenotypic and phylogenetic correlates of geographic range size in Canadian freshwater fishes. *Glob. Ecol. Biogeogr.* **22**, 1083–1094 (2013).
- M. W. Baldwin, H. Winkler, C. L. Organ, B. Helm, Wing pointedness associated with migratory distance in common-garden and comparative studies of stonechats (*Saxicola torquata*). *J. Evol. Biol.* **23**, 1050–1063 (2010).
- M. Zhao *et al.*, Time versus energy minimization migration strategy varies with body size and season in long-distance migratory shorebirds. *Mov. Ecol.* **5**, 23 (2017).
- K. R. Crooks, C. L. Burdett, D. M. Theobald, C. Rondinini, L. Boitani, Global patterns of fragmentation and connectivity of mammalian carnivore habitat. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **366**, 2642–2651 (2011).
- J. M. Hughes, D. J. Schmidt, D. S. Finn, Genes in streams: Using DNA to understand the movement of freshwater fauna and their riverine habitat. *Bioscience* **59**, 573–583 (2009).
- J. D. Carvajal-Quintero *et al.*, Damming fragments species' ranges and heightens extinction risk. *Conserv. Lett.* **10**, 708–716 (2017).
- C. Zarfl, A. E. Lumsdon, J. Berlekamp, L. Tydecks, K. Tockner, A global boom in hydropower dam construction. *Aquat. Sci.* **77**, 161–170 (2015).
- B. Lehner, G. Grill, Global river hydrography and network routing: Baseline data and new approaches to study the world's large river systems. *Hydrol. Processes* **27**, 2171–2186 (2013).
- B. Leroy *et al.*, Global biogeographical regions of freshwater fishes. [bioRxiv:10.1101/319566](https://doi.org/10.1101/319566) (11 May 2018).
- W. P. Köppen, *Grundriss der Klimakunde* (Walter de Gruyter & Co., Berlin, Leipzig, 1931).
- C. Quan, S. Han, T. Utescher, C. Zhang, Y.-S. (Christopher) Liu, Validation of temperature–precipitation based aridity index: Paleoclimatic implications. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **386**, 86–95 (2013).
- S. E. Fick, R. J. Hijmans, WorldClim 2: New 1-km spatial resolution climate surfaces for global land areas. *Int. J. Climatol.* **37**, 4302–4315 (2017).
- R. J. Hijmans, S. E. Cameron, J. L. Parra, P. G. Jones, A. Jarvis, Very high resolution interpolated climate surfaces for global land areas. *Int. J. Climatol.* **25**, 1965–1978 (2005).
- W. R. Peltier, Ice age paleotopography. *Science* **265**, 195–201 (1994).
- M. Zhao, F. A. Heinsch, R. R. Nemani, S. W. Running, Improvements of the MODIS terrestrial gross and net primary production global data set. *Remote Sens. Environ.* **95**, 164–176 (2005).
- A. N. Strahler, Dynamic basis of geomorphology. *Bull. Geol. Soc. Am.* **63**, 923–938 (1952).
- X. Shen, E. N. Anagnostou, Y. Mei, Y. Hong, A global distributed basin morphometric dataset. *Sci. Data* **4**, 160124 (2017).
- M. Verdú, Age at maturity and diversification in woody angiosperms. *Evolution* **56**, 1352–1361 (2002).
- R. Froese, D. Pauly, FishBase. <https://www.fishbase.org/>. Accessed 28 January 2019.
- A. Toussaint, N. Charpin, S. Brosse, S. Villéger, Global functional diversity of freshwater fish is concentrated in the Neotropics while functional vulnerability is widespread. *Sci. Rep.* **6**, 22125 (2016).
- S. Villéger, J. Ramos Miranda, D. Flores Hernández, D. Mouillot, Contrasting changes in taxonomic vs. functional diversity of tropical fish communities after habitat degradation. *Ecol. Appl.* **20**, 1512–1522 (2010).
- A. A. Pease, J. M. Taylor, K. O. Winemiller, R. S. King, Ecoregional, catchment, and reach-scale environmental factors shape functional-trait structure of stream fish assemblages. *Hydrobiologia* **753**, 265–283 (2015).
- J. Josse, F. Husson, missMDA: A package for handling missing values in multivariate data analysis. *J. Stat. Softw.* **70**, 1–31 (2016).
- B. Shipley, Confirmatory path analysis in a generalized multilevel context. *Ecology* **90**, 363–368 (2009).
- K. P. Burnham, D. R. Anderson, *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach* (Springer-Verlag, New York, ed. 2, 2002). <https://www.springer.com/fr/book/9780387953649>. Accessed 29 January 2019.
- S. Nakagawa, H. Schielzeth, A general and simple method for obtaining R^2 from generalized linear mixed-effects models. *Methods Ecol. Evol.* **4**, 133–142 (2013).
- B. Shipley, *Cause and Correlation in Biology: A User's Guide to Path Analysis, Structural Equations and Causal Inference with R* (Cambridge University Press, 2016).
- J. B. Grace *et al.*, Guidelines for a graph-theoretic implementation of structural equation modeling. *Ecosphere* **3**, art73 (2012).
- R Development Core Team, R: A Language and Environment for Statistical Computing (Version 3.4.3, R Found Stat Comput, Vienna, 2017).