ECOLOGY

Unexpected fish diversity gradients in the Amazon basin

Thierry Oberdorff[1]*, Murilo S. Dias[2]*, Céline Jézéquel[1], James S. Albert[3], Caroline C. Arantes[4], Rémy Bigorne[1], Fernando M. Carvajal-Valleros[5], Aaike De Wever[6,7,8], R. G. Frederico[9], Max Hidalgo[10], Bernard Hugueny[1], Fabien Leprieur[11], Mabel Maldonado[12], Javier Maldonado-Ocampo[13†], Koen Martens[6,7], Hernan Ortega[10], Jaime Sarmiento[14], Pablo A. Tedesco[1], Gislene Torrente-Vilara[15], Kirk O. Winemiller[16], Jansen Zuanon[17]

Using the most comprehensive fish occurrence database, we evaluated the importance of ecological and historical drivers in diversity patterns of subdrainage basins across the Amazon system. Linear models reveal the influence of climatic conditions, habitat size and sub-basin isolation on species diversity. Unexpectedly, the species richness model also highlighted a negative upriver-downriver gradient, contrary to predictions of increasing richness at more downriver locations along fluvial gradients. This reverse gradient may be linked to the history of the Amazon drainage network, which, after isolation as western and eastern basins throughout the Miocene, only began flowing eastward 1–9 million years (Ma) ago. Our results suggest that the main center of fish diversity was located westward, with fish dispersal progressing eastward after the basins were united and the Amazon River assumed its modern course toward the Atlantic. This dispersal process seems not yet achieved, suggesting a recent formation of the current Amazon system.

INTRODUCTION

The Amazon basin covers more than 6,000,000 km^2, produces about 16% of the world's freshwater discharge (1–3), and contains the highest freshwater biodiversity on Earth (4). This is especially true for fishes, as the 2257 species (including 1248 endemics; i.e., species found nowhere else on Earth) recognized in the Amazon basin (www.amazon-fish.com) represent ~15% of the world freshwater fishes described so far (5). The Amazonian ichthyofauna has ancient origins, and many clades of Amazonian fishes had achieved modern phenotypes by the early Neogene [~23 million years (Ma) ago] (6, 7). The vast majority of the modern fish fauna is represented by lowland adapted species, with only ~6% of the Amazonian fish species (i.e., 129 species) having a geographical distribution restricted above 300-m elevation according to our database. The distribution of Amazonian fish species is highly uneven (8–10), but diversity patterns of this mega diverse fauna and the processes generating these patterns are still incompletely understood. In particular, the drivers shaping species richness and composition gradients at the scale of the entire Amazon basin remain poorly documented, while there is evidence that the structure of Amazonian freshwater ecosystems is increasingly affected by rapid expansions in urbanization and other human economic activities (3, 11–13).

It has been well established that both ecological and historical processes contribute to generating and maintaining diversity gradients, but no consensus has yet emerged on which of these processes dominate, in part, because of strong context dependency (14). Three major hypotheses are widely invoked to explain diversity gradients (15), and these hypotheses also apply to riverine fishes [reviewed in (16)]. The species-area (or species-discharge) hypothesis refers to the positive relationship between the number of fish species and the size or total habitat volume of a river (17, 18), due to river size–dependent extinction and speciation rates and to an increase in habitat heterogeneity with the size of the river (19–21). The species-energy hypothesis predicts a positive link between species richness and the energy available within a system either by increasing population sizes and thus lowering extinction rates, by increasing metabolic rates and thus promoting higher rates of speciation, or by affecting ecophysiological limits (e.g., growth and reproduction) of species (4, 16, 17, 22). Last, the historical hypotheses explain differences in diversity gradients by combining past environmental conditions with geographic contingencies regulating dispersal possibilities and thus colonization, extinction, and speciation processes [e.g., (23, 24)]. For example, paleogeographic and paleoenvironmental conditions in Amazonia during the Miocene (between ~23 and 5 Ma ago) were very different from the present, with a large mega-wetland system occupying much of western Amazonia (6, 25) and a disconnected proto–Amazon River in the East (26), impeding fish dispersal between these two drainage systems (fig. S3).

[1]UMR EDB (Laboratoire Évolution et Diversité Biologique), CNRS 5174, IRD253, UPS; 118 route de Narbonne, F-31062 Toulouse, France. [2]Departamento de Ecologia, Universidade de Brasília, Brasilia, DF, Brazil. [3]University of Louisiana at Lafayette, Lafayette, LA 70504, USA. [4]Center for Global Change and Earth Observations, Michigan State University, East Lansing, MI 48824, USA. [5]Museo de Historia Natural Alcide d'Orbigny, Av. Potosí 1458, zona Queru Queru, Cochabamba, Estado Plurinacional de Bolivia. [6]Department of Freshwater Biology, Royal Belgian Institute of Natural Sciences, Vautierstraat 29, B-1000 Brussels, Belgium. [7]Department of Biology, University of Ghent, K.L. Ledeganckstraat 35, B-9000 Ghent, Belgium. [8]Research Institute for Nature and Forrest (INBO), Havenlaan 88 bus 73, 1000 Brussels, Belgium. [9]Universidade Federal de Minas Gerais (UFMG), Instituto de Ciências Biológicas, Av. Antônio Carlos, 6627 Belo Horizonte, MG, Brazil. [10]Departamento de Ictiología, Museo de Historia Natural, Universidad Nacional Mayor San Marcos, Lima, Perú. [11]UMR MARBEC (CNRS, IRD, IFREMER, UM), Université de Montpellier, Montpellier Cedex 5, France. [12]Unidad de Limnología y Recursos Acuáticos, Universidad Mayor de San Simón, Calle Sucre y Parque La Torre s/n, Cochabamba, Bolivia. [13]Unidad de Ecología y Sistemática (UNESIS), Laboratorio de Ictiología, Departamento de Biología, Facultad de Ciencias, Pontificia Universidad Javeriana, Bogotá, Colombia. [14]Departamento de Ictiología, Museo Nacional de Historia Natural, La Paz, Bolivia. [15]Universidade Federal de São Paulo, Instituto do Mar, Campus Baixada Santista, Rua Doutor Carvalho de Mendonça, 144, Encruzilhada, 11070-100 Santos, SP, Brazil. [16]Department of Wildlife and Fisheries Sciences, Texas A&M University, College Station, TX 77840, USA. [17]Instituto Nacional de Pesquisas da Amazônia, Coordenação de Biodiversidade, Avenida André Araújo, 2936, Petrópolis, 69067-375 Manaus, AM, Brazil.
*Corresponding author. Email: thierry.oberdorff@ird.fr (T.O.); murilosd@hotmail.com (M.S.D.)
†Deceased.
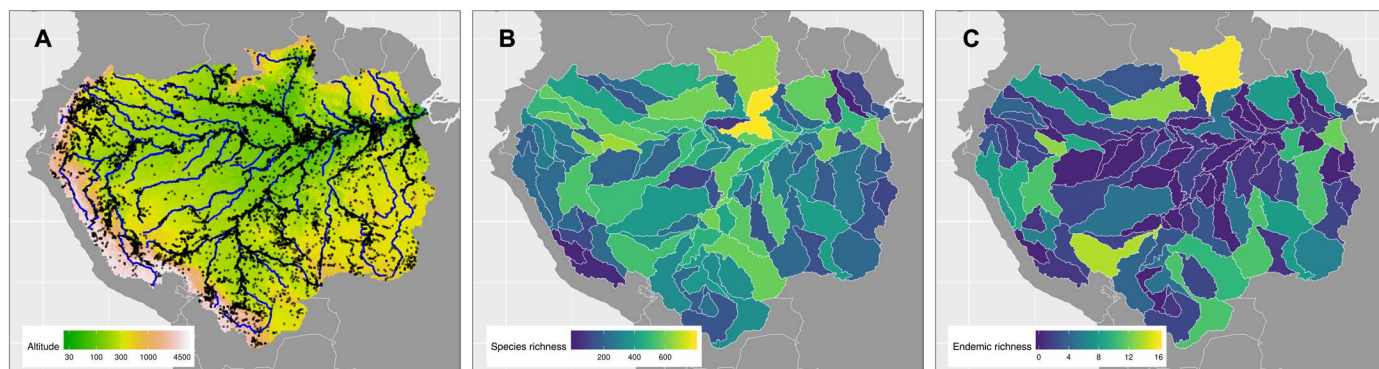
Fig. 1. Sampling sites, species richness, and endemism patterns. (**A**) Fish occurrence records available in the AmazonFish database for each subdrainage basin. Gradients in total species richness (**B**) and endemism (**C**) across the 97 subdrainage basins of the Amazon basin. The Amazon basin is flowing West-East to the Atlantic Ocean.

Using a macroecological and geospatial approach and making use of the most comprehensive fish occurrence database currently available (www.amazon-fish.com) and assembled for this study (27), we evaluate the importance of these three major hypotheses in explaining present-day fish diversity patterns in 97 subdrainage basins along the Amazon main stem and its major tributaries (Fig. 1A). Using a set of generalized linear models (GLMs) and species distribution models (SDMs), we explore a range of environmental and historical factors related to the three hypotheses for their abilities to explain richness and endemism patterns (Fig. 1), hence providing insights into the potential drivers of current Amazonian fish diversity (table S1).

RESULTS AND DISCUSSION
Species richness drivers
The total fish species richness per subdrainage basin is significantly and positively associated with increased area ($z = 9.95$, $P < 0.001$; z value represents the respective regression coefficient divided by the estimated SE in the GLM), increased current temperature (PC2_temp, positively associated with minimum, maximum, and mean temperature: $z = 3.98$, $P < 0.001$), increased stability in current temperature conditions (PC1 temp, positively associated with temperature variability: $z = -4.08$, $P < 0.001$), increased energy availability (PC2 energ, positively associated with mean net primary productivity and minimum actual evapotranspiration: $z = 2.32$, $P = 0.02$), increased distance of the subdrainage basin from the river mouth ($z = 2.58$, $P = 0.01$), and increased sampling effort ($z = 13.87$, $P < 0.001$). The model also depicts a significant negative effect of habitat harshness, mostly associated with high elevation and steep gradients in our case (see Materials and Methods) (PC1_elev, negatively correlated with all selected descriptors: $z = -2.51$, $P = 0.01$). Together, these factors explain 82% of the total variance in fish species richness (see Materials and Methods and tables S1 to S3 for variables and model description). Note that sampling effort has the largest estimated coefficient and z value in the model and thus the largest effect on species richness. We tried to control as far as possible for this bias by working at the subdrainage basin grain, but its effects remain high even at this spatial extent. However, if the main tributaries of the Amazon basin appear well surveyed, some gaps do exist in the central and peripheral parts of the Basin (see fig. S1B). These gaps represent locations that are hardly accessible to sampling either because they are highly isolated or because they are located in protected areas (i.e., indigenous territories or protected areas). The iden-

tification of undersampled sub-basins is a first step to guide increasing biodiversity knowledge in these still unknown areas. We have already initiated this process by supporting the numeric digitalization of the national freshwater fish collections from Peru and by initiating sampling campaigns in detected gaps in Colombia, Peru, and Brazil (www.amazon-fish.com). However, after excluding sampling effort as a predictor, the model still explains 51% of the total variation in richness.

Endemism drivers
Endemism richness per subdrainage is significantly associated with increased area ($z = 3.83$, $P < 0.001$), increased current energy stability (PC1_energy negatively linked with an increase in net primary productivity variability: $z = -2.04$, $P = 0.04$), increased stability in current and past climatic conditions (PC2_water, positively associated with precipitation variability: $z = -2.31$, $P = 0.02$; PC2_ClimCurrPast_diff, negatively associated with differences in maximum and mean precipitation: $z = 2.10$, $P = 0.04$; PC3_ClimCurrPast_diff, negatively associated with differences in minimum temperature: $z = 2.48$, $P = 0.01$), increased isolation (measured as the number of waterfalls in the sub-basin: $z = 2.71$, $P = 0.01$), and increased sampling effort ($z = 2.14$, $P = 0.03$). The model also depicts a significant effect of total species richness ($z = 2.29$, $P = 0.02$). The standardized effect sizes of these variables (mean estimates and z) are quite similar in magnitude. Together, these effects explain 69% of total variance in endemic species richness (see Materials and Methods and tables S1 to S3 for predictors and model description).

Area, climate, and energy hypotheses support species richness patterns
After controlling for sampling effort, results of the species richness model first support the notion that area, climate per se, and energy availability all play a significant role in explaining freshwater fish richness patterns in the Amazon basin. These findings are in good agreement with previous research demonstrating that these are major drivers of riverine fish species richness patterns worldwide at both interbasin (4, 17) and intrabasin scales (28, 29). Among these drivers, area is the most important, lending support to the area hypothesis through potentially three nonexclusive mechanisms. The first mechanism is that the probability of species extinctions increases with a reduction in river size because of a decrease in population sizes (19, 21). The second mechanism is that the probability of speciation increases with an increase of river size by exposing species to greater ecological

heterogeneity and/or geographical barriers (20). Similarly, the third mechanism is that greater ecological heterogeneity offers a larger number of available niches, favoring the coexistence of a larger number of species (21). In addition to area, results showing that energy availability and temperature were positively related to species richness are compatible with both components of the species energy hypothesis: Productivity facilitates larger population sizes, decreasing extinction rates, or allows more niche specialists to cohabit by increasing resources availability (16), and higher temperatures promote species richness by increasing metabolic rates and thus potentially speciation rates [(22), but see (30, 31) for questioning this last hypothesis] and/or by relaxing thermal constraints on species, lowering extinction rates (16). The positive effect of temperature stability on fish diversity suggests that climatically stable regions may promote lower rates of extinction through relative constancy of resources (32). Last, results showing the negative effect of habitat harshness on sub-basin richness were highly expected, as harsh habitat conditions (e.g., high elevation and steep gradients in our case) are well known to exclude fishes from successful colonization (33). The historical hypothesis appears, at first sight, not supported by our species richness model, as all our pre-selected historical predictors were nonsignificant (table S3).

Area, climate, and historical hypotheses support endemism patterns

Regarding endemism, our model first demonstrates a positive effect of area and environmental stability (both current and historical) on the extent of diversification. As previously noted, while area may increase the rate of speciation through greater habitat heterogeneity (16, 32, 34), environmental stability through time may further allow finer specializations and adaptations because of the relative constancy of resources (16, 34). Besides area and environmental stability, our model also depicts a positive effect of sub-basin isolation measured here by physical discontinuities (i.e., number of natural waterfalls). These discontinuities may have promoted speciation by maintaining divergence among isolated populations, as previously noticed in the Amazon basin (35, 36) and in other South American freshwater systems (20).

West-East unsaturation in species hypothesis

While the drivers and processes underlying richness and endemism patterns we described here are consistent with findings from previous research on freshwater fishes (20, 32, 34), our richness model also captures an unexpected slight, yet significant, linear decrease of species richness along the upriver-downriver gradient. This result suggests that our defined sub-basins are overall richer in upriver (Western) portions of the Amazon network as compared to downstream (Eastern) ones (Fig. 2). This overall atypical richness pattern is also observed for 14 of the 15 most species-rich families (~78% of the total Amazonian fish richness) and is statistically significant (or marginally significant) for 6 of them, representing ~58% of the total Amazonian fish fauna (table S4). These patterns are surprising in that they tend in the opposite direction of the usual increase in riverine fish diversity observed along the upriver-downriver gradient, regardless of the spatial grain or extent, across numerous temperate and tropical river systems worldwide [e.g., (29, 37–41)]. This positive longitudinal species richness gradient has been attributed to low colonization and high extinction rates in upriver portions of the hydrological network and to low extinction rates in downriver ones due to the increase in productivity, connectivity, and habitat diversity with river volume (see
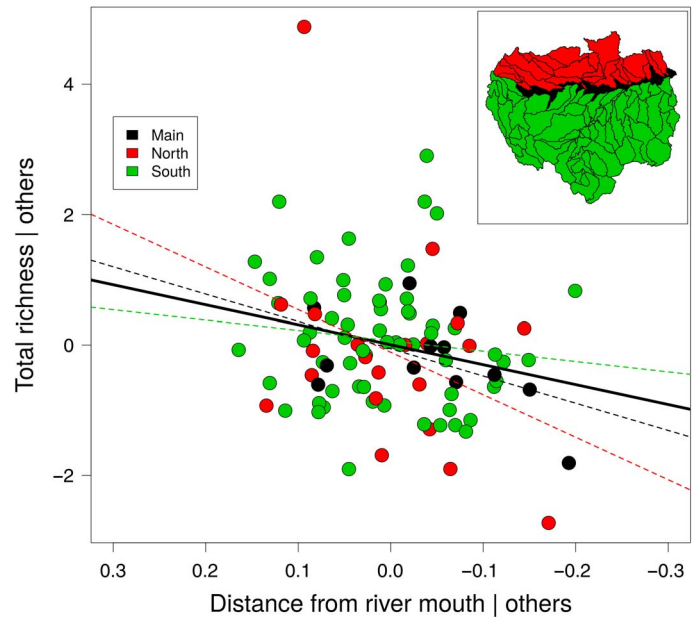
Fig. 2. Partial correlation plot (i.e., Pearson residuals) of the distance from river mouth on subdrainage basins total fish species richness after controlling for all other predictors considered in our richness model. The overall decrease in sub-basins species richness from upriver to downriver is statistically significant (solid line, table S3) (Pearson's $r = 0.24$, $P = 0.020$, $n = 97$). Note that removing the outlier sub-basin (red point on top of the graph) from the model changes neither the significance nor the trend of the relationship (Pearson's $r = 0.21$, $P = 0.040$, $n = 96$). The atypical richness pattern (i.e., a decreasing trend in species richness from upriver to downriver) is depicted for the Amazon main channel (black) and subdrainage basins located north (red) or south (green) of this mainstem and is marginally significant for basins located within the mainstem (Pearson's $r = 0.51$, $P = 0.089$, $n = 12$) or north (Pearson's $r = 0.36$, $P = 0.098$, $n = 22$) of the mainstem (red and black dashed lines), but not significant for sub-basins located south of the mainstem (Pearson's $r = 0.13$, $P = 0.320$, $n = 63$).

Materials and Methods for a more detailed explanation) (29, 33). We interpret this reverse pattern as a potentially West-East gradient of increasing "unsaturation" of subdrainage assemblage richness (i.e., subdrainages open to colonization by new species independently of any local environmental limit) (42, 43). Although the increase in riverine fish diversity along the upriver-downriver gradient has rarely been examined in systems as large as the Amazon [but see (29)], we see no valid ecological reason to expect such a reverse trend. To further test the validity of our hypothesis of a West-East unsaturation in species, we used SDMs for 1351 species (60% of all Amazonian species) to simulate sub-basin richness under the assumption of free colonization of suitable habitats by species within the Amazon basin and reran the richness model using these new SDM-derived richness data (see Materials and Methods). In doing so, we expected the inverse richness gradient noted (i.e., a decrease in species richness from upriver to downriver) to disappear or even return to the expected increase in species richness from upriver to downriver. Results corroborate our expectation, as the new model confirms the disappearance of the inverse gradient in richness and further depicts a statistically significant ($z = -2.26$, $P = 0.02$) increase in overall species richness from upriver to downriver conforming to the expected pattern (table S4). This increase in species richness from upriver to downriver is found for 14 of the 15 most species-rich families and is statistically

significant (or marginally significant) for 5 of them (table S5). The return to the expected longitudinal pattern in richness as depicted by SDMs makes the reverse trend found hard to explain without reference to historical contingencies.

The paleogeographic history of northern South America may contribute substantially to interpreting the origin of this reverse longitudinal species richness gradient. There is a general consensus that during the Early Miocene (from ~23 Ma), an inland lacustrine and marginally marine (44) Amazonian system (i.e., the Pebas system) partially flooded northwest South America until the Purus Arch (Fig. 3), isolating basins west and east of this Arch (6). During this period, a restricted eastward-flowing proto–Amazon River may have drained to the Atlantic coast (26, 45). Within this context, our identified West-East decreasing diversity gradient leads us to propose that the western part of the Amazon may be the primary geographic area of origination and colonization for the dominant component of the contemporary Amazonian fish fauna. Further, high levels of dissimilarity in the composition of fish faunas observed between the most southeastern (downriver) part of the Amazon and the rest of the basin (Fig. 3) suggest that this southeastern region is a second center of origin of the Amazonian fish fauna, even if much smaller in size and with lower species richness (Fig. 3 and table S6).

Following our results, we propose a center of origin–dispersal–adaptation model that assumes a mostly westward restricted origin of the species pool, followed by dispersal and adaptation to previously unavailable aquatic habitats generated by the fusion of western and eastern aquatic systems when the modern Amazon River and its major tributaries became established. However, this West-East dispersal phase, which is mostly restricted to lowland species (Fig. 3), seems incomplete, as depicted by the unusual negative upstream/downstream gradients in sub-basin richness (Fig. 2).

An alternative and/or complementary explanation to the origin of the West-East decreasing diversity gradient could be related to Quaternary hydroclimatic variability in Amazonia. A previous study based on absolute-dated speleothem oxygen isotope records indicates that the climate in eastern Amazonia alternated drastically between dry and wet conditions during the Last Glacial Maximum [LGM; ~21 thousand years (ka)], whereas western Amazonia has had a more stable climate over the last 250 ka (46, 47). This different history of climate evolution between western and eastern Amazonia may have maintained high diversity in the western part (48) through a decrease in extinction rates while increasing extinction rates in the eastern part of the basin (46). This mechanism may also help explain the West-East richness gradient reported for some other groups of organisms, such as trees, mammals, and birds (6, 48). While this scenario is compatible with our finding of an overall richer fish fauna in the western part, the absence of an effect of LGM climatic stability in our richness model strongly minimizes this possibility (table S2). Another potential
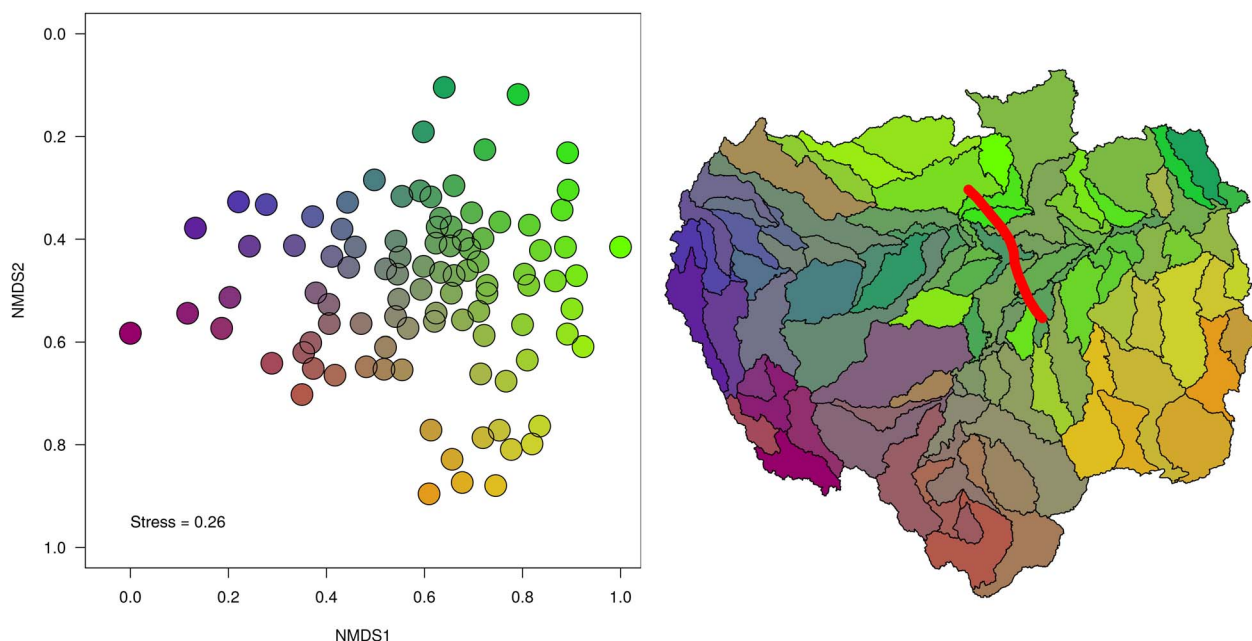
Fig. 3. Composition of freshwater fishes for the 97 Amazonian subdrainage basins inferred from nonmetric multidimensional scaling ordination using a dissimilarity matrix based on the Simpson's dissimilarity index (βsim), a measure of spatial turnover of species composition without the influence of richness gradients. Colors on the map reflect the corresponding point position along both nonmetric multidimensional scaling (NMDS) axes. The analysis depicts stronger dissimilarities in fauna's composition for the most southeastern part (yellow shades) of the Amazon and the most westerly Andean end (purple and pink shades) compared to the rest of the Basin. The most southeastern part is mainly represented by the Xingu and Tapajós rivers, both hosting a relatively high number of highland endemic species, i.e., 27 endemic species having a geographical distribution restricted above the 300-m elevation. The most westerly Andean end is mainly represented by the Ucayali and Marañón rivers in Peruvian Amazonia and hosts, as for the most southeastern part of the basin, a high number of highland endemic species, i.e., 28 endemic species having a geographical distribution restricted above the 300-m elevation. The red line in the figure at the right indicates the delimitation of the Purus Arch. According to our data, sub-basins located upstream of the Purus Arch (i.e., belonging to the historical Pebas system; fig. S3) currently cover 64% of the total surface area of the entire Amazon basin and host 86% of its total richness (with two families, 72 genera, and 709 species found exclusively in this area), while sub-basins located downstream of the Arch cover only 36% of the total surface and host 68% of the total richness of the basin (with two families, 46 genera, and 313 species found exclusively in this area).

explanation for the West-East decreasing diversity gradient could be linked to historical sea-level fluctuations within the Amazon basin. The eastern Amazon region has suffered a sea rise estimated from 50 to 100 m ~5 Ma ago for a duration of ~0.8 Ma ago (49, 50), followed by more recent Pleistocene marine incursions (<1 Ma ago) of smaller magnitudes ~25 m (50). These seawater incursions could have eliminated freshwater habitats in the low-lying areas of the lower and central Amazon, probably leading to high extinction rates of lowland freshwater fish species in these areas while at the same time favoring diversification processes in the remaining higher elevation isolated areas (51). An important outcome of these marine incursions and consequent extinctions is that the return to lower sea-level conditions may have left freshwater habitats of the lowland eastern and central Amazon basin unsaturated in species and thus potentially open to colonization. However, the absence of statistically significant effects of the two variables related to historical marine incursions in our richness model cannot make this explanation the primary one responsible for our West-East decreasing diversity gradient.

The formation date of the modern Amazon River after transition of the Pebas system to the Acre system with the origin of the transcontinental river flowing West-East to the Atlantic Ocean is controversial, differing greatly depending on authors. Some authors (6, 26, 52) proposed a starting formation date during the late Miocene (~9 to 10.5 Ma ago) or early Pliocene (~5 Ma ago) (25) following disappearance of the Pebas system, while others suggest that the Pebas system [or a complex of interconnected mega-lakes (53)] persisted until the modern Amazon River system developed ~2.5 Ma ago (early Pleistocene) (53) or even later, during the Middle Pleistocene (<1 Ma ago) (54). Our results are more consistent with the view of a "Young Amazon" ecosystem, in which the reverse longitudinal species gradient is interpreted to have originated within the last ~2.5 Ma ago. Our interpretation of a West-East colonization pattern still in process seems poorly compatible with an origin of this process in the late Miocene (~11 to 5 Ma ago), as we expect such long time periods to be sufficient to allow species to colonize most suitable habitats. Our findings rather agree with several phylogeographic studies detecting signals of easterly (35, 36) or westerly (51) trajectories of colonization for several fish taxa across the Purus Arch accompanied by demographic expansions (35) during the early and mid-Pleistocene (~2.6 to 0.7 Ma ago). Our results are also well in line with biogeographic studies reporting diversification for various avian and mammal taxa dated between ~2.6 and 0.13 Ma ago and attributed to a sequential formation of the major tributaries of the Amazon during this time step (55–60).

CONCLUSION

In summary, our results support the idea that combined influences of environmental heterogeneity, climate, and historical contingencies explain the contemporary patterns of fish diversity in the Amazon basin. Following the West-East gradient of decreasing richness, we hypothesize that fish assemblages experience a corresponding gradient in unsaturation in species, most of the lowland eastern part of the Basin, in contrast to the western one, being constrained in its richness and composition by ongoing downriver dispersal and ecological filtering processes. This West-East gradient in assemblage unsaturation is likely to have an influence on the potential for lowland species to shift their distributions in tracking changing climates either upriver, where more saturated assemblages and thus stronger competition are expected, or downriver, where the regional pool of species

appears still unbounded (i.e., downriver part of the Basin). Further, these dynamic natural dispersal processes could be durably compromised by habitat destruction linked to ongoing deforestation and expansion of plantations, particularly in the eastern region (47), and the disruption of river longitudinal connectivity generated by hydropower dams already built or planned in this highly biodiverse Basin (3, 12, 13).

MATERIALS AND METHODS
Biological data
All biological data were compiled under the AmazonFish project (www.amazon-fish.com). This still ongoing project aims to build a high-quality freshwater fish biodiversity database for the entire Amazon catchment. This was done by mobilizing and integrating all information available in published articles, books, gray literature, online databases, foreign and national museums, and universities and by checking for systematic reliability and consistency for each species recorded. At this time, the database includes ~18,000 sampled sites, 56 families, 510 genera, and 2257 valid native freshwater fish species (27). As far as we know, this database contains the most complete and up-to-date information currently available on freshwater fish species distribution for the entire Amazon drainage basin. The only other database available at this spatial extent reports species distribution at a much coarser grain and currently suffers from a substantial lack of available information. We found an almost constant high negative difference in species number between each of the 23 Amazonian units listed in (10) and what we obtained for these units using our data (mean value = −288 species, median value = −290 species).

Patterns of freshwater fish diversity were analyzed using two diversity descriptors: species richness and number of endemic species. Species richness is a measure of the total number of native species present in a subdrainage basin, whereas the number of endemic species is calculated as the sum of species present in a subdrainage basin and, as far as we know, nowhere else on Earth (5).

Subdrainage basin delineation
To harmonize, as far as possible, sampling effort, we decided to work at the subdrainage basin grain. To classify our subdrainage basins, we used the HydroBASINS framework (www.hydrosheds.org/page/hydrobasins), a subset of the HydroSHEDS database (61). We combined different HydroBASINS levels to retain only sub-basins >20,000 km^2. Some adjacent sub-basins were grouped to further optimize the sampling effort (i.e., the number of sampling sites within sub-basins). The sub-basins located in the river mainstem were delineated on the basis of the distance between two main tributaries entering the mainstem, resulting in eight sub-basins with a surface area of <20,000 km^2. This led to a total of 97 sub-basins covering the entire Amazon system (fig. S1, A and B).

Subdrainage position in the Amazon catchment
We used two variables to position each of the subdrainages within the Amazon River network: (i) the distance of each of the subdrainages to the river mouth (in kilometers) and (ii) the downriver-upriver position of the principal tributary hosting a sub-basin (fig. S1C). The first variable relates to the usual decrease in fish diversity along the downriver-upriver gradient already noticed in numerous temperate and tropical river mainstems [e.g., (37, 38)]. The second variable is also related to the downriver-upriver gradient in richness. As species richness is supposed to decrease from downriver to upriver in the river

mainstem and as the river mainstem likely acts as an immigrant source for many species that colonize tributaries, we expect downriver tributaries to support higher species than do similar-sized tributaries located upriver in the drainage network (29, 39, 40).

Current environmental factors

We divided environmental factors in accordance with the "climate/productivity," "area/environmental heterogeneity" [see (15) for a detailed description of predictions related to these hypotheses], and historical hypotheses. Data sources and definitions are presented in table S1, in addition to the brief overview below. Before the analyses, predictors were transformed, when necessary, to minimize potential effects of extreme values.

To test the climate/productivity hypothesis, we used for each sub-drainage basin the annual mean and seasonality [Coefficient of variation (CV) of intrayear monthly values] of (i) temperature, (ii) precipitation, (iii) actual evapotranspiration, (iv) potential evapotranspiration, (v) net primary productivity, (vi) solar radiation, (vii) runoff, and (viii) the lowest (or highest) value of minimum (or maximum) temperature of the coldest (or warmest) month. These variables measure the mean current climatic condition, the seasonal climatic variability, and the energy availability within each subdrainage basin. We used terrestrial temperature and productivity as surrogates for water temperature and productivity, as these terrestrial and aquatic variables are likely to covary closely (16). Principal components analyses (PCAs) were applied separately for each group of variables related to precipitation, temperature, and productivity to reduce the multidimensionality and to eliminate multicollinearity within groups. We retained the first two PCA axes as synthetic predictors for each group of variables (table S2).

To test for the "habitat size/diversity" hypothesis, we considered four synthetic variables recognized as important factors shaping fish biodiversity [table S1; see (4, 34)]: (i) the surface area of the sub-drainage basin (in square kilometers); (ii) the network density (length of the riverine network divided by the surface area of the sub-basin), a measure of habitat availability for fishes; (iii) the land cover heterogeneity (measured as the Shannon diversity index based on the proportion of land cover classes within each subdrainage basin); and (iv) the soil heterogeneity (measured as the Shannon diversity index based on the proportion of each soil type within each subdrainage basin) (table S1). We further classified sub-basins according to their main water color (black water, white water, and clear water rivers) following (2, 62) (fig. S2). Water color mostly reflects geological conditions, and white waters are known to be comparatively richer in energy availability (e.g., nutrients, zooplankton, and aquatic insect larvae) than "black" or "clear" waters (8). The three water colors were coded as categorical variables.

To test for natural barriers to colonization ("fragmentation/isolation" hypothesis), we used four variables: (i) the number of waterfalls within each sub-basin using data from (61), (ii) the sub-basin elevation (mean, maximum, minimum, range, and SD) (in meters), (iii) the proportion of the sub-basin surface with terrain slope above 15% (in square kilometers), and (iv) the proportion of the sub-basin surface above 1000 m in altitude (in square kilometers). A PCA was further performed on the last three variables (i.e., seven factors), and the first axis was retained as a synthetic predictor describing environmental harshness.

We also included as a potential predictor the number of sampling sites divided by the surface area of each sub-basin in our models. This predictor is important to control for a potential sampling effort effect in our models (fig. S1B).

Historical factors

To test the "history/dispersion" hypothesis, we considered the following predictors: (i) sub-basins belonging (1) or not (0) to the Pebas system at ~23 Ma ago (6) (fig. S3), (ii) Quaternary climate stability within the sub-basin (~21 ka), (iii) the surface area of each sub-basin under seawater (following current topography) considering a sea-level rise of 100 m (~5 Ma ago) (50), and (iv) the surface area of each sub-basin under seawater considering a sea-level rise of 25 m during recent Pleistocene marine incursions (<1 Ma ago) (fig. S4) (50). To describe climate stability during the Quaternary, we used reconstructions of mean, max, and min annual temperatures and precipitations at the LGM (21 ka) and calculated the difference between present and LGM values. The difference between the LGM and the present is one of the strongest climatic shifts in all of the Quaternary (63). We extracted the annual temperature and precipitation during the LGM from three general circulation models (GCMs), namely, CCSM (Community Climate System Model), MIROC (Model for Interdisciplinary Research on Climate), and MPI (Max Planck Institute model) (data available from www.worldclim.org/version1). For each GCM, the changes in temperature and precipitation between the present and the LGM were calculated, and the resulting values were averaged to account for variation among models. A PCA was applied to reduce the multidimensionality and to eliminate multicollinearity between these last variables. We retained the first three PCA components as synthetic predictors (table S2).

Statistical analyses

We used negative binomial and Poisson GLMs to evaluate the support for the major hypotheses by relating each diversity descriptor to our environmental and historical predictors. As each quantitative predictor displays a different measurement unit, we standardized each predictor by subtracting the mean and dividing by two times the SD to get comparable coefficients from our models. In this way, model-estimated coefficients are on the same scale and can be directly compared. Predictor significance was determined by Wald' z statistic and associated P values from GLM outputs. The z values were calculated by dividing the estimated regression coefficients by the estimated SE. The higher the absolute value of z, the stronger the effect on the dependent variable and the lower the P value. We also checked predictor significance using a likelihood ratio test (LRT) by dropping each predictor from the full model and by calculating differences in model fit based on χ^2 distributions. As the obtained P values from LRT were similar to the those from the z statistic, we reported only here the z and associated P values.

After model fitting, we checked for broad spatial autocorrelation in model residuals by computing the Moran's I statistic (and corresponding P values) using the inverse of the watercourse distances among sub-basins as weights. The Moran's I statistic varies between −1 and 1, larger values indicating strong (either positive or negative) spatial structures. As spatial autocorrelation was either weak or not significant in model residuals, we opted to maintain nonspatial GLM models (tables S3 to S5). We further calculated the pseudo-R^2 using null and residual deviances from GLM models as a measure of model fit. Multicollinearity was checked using a variance inflation factor (VIF) procedure. As expected, we noticed some multicollinearity problems for some of our explanatory variables (i.e., DistMouth, PC1_water, PC1_energ, and PC1_elev having VIF > 10). We ran our models with and without these last four variables. However, as models from both procedures gave very similar responses, we decided to preferentially present models including all original variables to not ignore the potentially

unique contribution of each variable (see table S3). All analyses and graphics were performed under the R environment (64).

To test the validity of our West-East unsaturation in species hypothesis, we used SDMs for 1351 species showing more than 10 occurrence points in our database (60% of all Amazonian species) to simulate richness in sub-basins under the assumption of free colonization of suitable habitats by species within the Amazon basin. To model the distribution of species, we used 19 bioclimatic variables related to temperature and precipitation (averaged for the period 1950–2000) from the WorldClim database (65), plus a set of biologically meaningful physical variables, i.e., elevation (GDEM ACE2), elevation range, maximum slope, stream length, and flow accumulation (66). We aggregated biological data at the grid scale corresponding to the resolution of the bioclimatic dataset (10 km). For both bioclimatic and physical variables, we selected the least correlated variables (Pearson's $r < 0.70$) and kept the most ecologically meaningful one when two variables were correlated with Pearson's $r \geq 0.70$ (67). Distributions were projected under the BIOMOD2 platform (68) using five modeling techniques [GLM, generalized additive model, generalized boosted model, multivariate adaptive regression splines, and maximum entropy (MaxEnt)]. We generated three sets of 1000 randomly selected pseudo-absences with equal weighting for presence and absence. The models were calibrated with 70% of the data selected at random, and the predictive performance of each model was evaluated on the remaining 30% using the area under the relative operating characteristic curve (AUC) and the true skill statistic (TSS). This process was repeated three times. To produce robust distribution forecasts, we applied an ensemble forecast method to combine the five modeling techniques (68). Models with TSS evaluations below 0.6 were discarded, and the current consensus distributions were obtained by averaging distributions with weights proportional to their TSS evaluation. Probability maps were transformed into maps of suitable versus nonsuitable areas by choosing the probability threshold that maximized the TSS value. We finally combined all the SDM results in a unique framework to extract a species list for each sub-basin.

To measure sub-basin compositional (dis)similarity, we used the Simpson's index of beta diversity (βsim) based on occurrence data as

$$\beta sim = \min(b, c)/[\min(b, c) + a]$$

where a is the number of species shared between two sub-basins, and b and c represent the number of species unique to each sub-basin. The βsim value ranges from 0 to 1, where 0 means pairs of sub-basins having identical taxa lists and 1 means no shared taxa between pairs of sub-basins. We further performed a nonmetric multidimensional scaling ordination, using the "metaMDS" function from the vegan package (69) under the R environment.

SUPPLEMENTARY MATERIALS
Supplementary material for this article is available at http://advances.sciencemag.org/cgi/content/full/5/9/eaav8681/DC1

Fig. S1. Names, sampling effort, and longitudinal position of the 97 sub-basins.
Fig. S2. Principal water quality of major subdrainages within the Amazon basin.
Fig. S3. Delimitation of the Pebas lake ~23 to 10 Ma ago and Purus Arch.
Fig. S4. Surfaces of the current Amazon basin covered by seawater under 25- and 100-m sea-level rises.
Table S1. Environmental predictors used to explain fish diversity and endemism in the Amazon basin.

Table S2. Correlation coefficients among input variables and components for each of the PCAs performed.
Table S3. Coefficients and SEs from the GLMs performed for total and endemic richness.
Table S4. Coefficients and SEs (log link scale) from the GLMs for the richness of each of the 15 most abundant families (raw data).
Table S5. Coefficients and SEs (log link scale) from the GLMs for species richness estimated from SDMs of each of the 15 most abundant families.
Table S6. Families, genera, and number of species found exclusively west and east of the Purus Arch.

REFERENCES AND NOTES
1. E. Salati, P. B. Vose, Amazon basin: A system in equilibrium. *Science* **225**, 129–138 (1984).
2. E. Venticinque, B. Forsberg, R. Barthem, P. Petry, L. Hess, A. Mercado, C. Cañas, M. Montoya, C. Durigan, M. Goulding, An explicit GIS-based river basin framework for aquatic ecosystem conservation in the Amazon. *Earth Syst. Sci. Data* **8**, 651–661 (2016).
3. E. M. Latrubesse, E. Y. Arima, T. Dunne, E. Park, V. R. Baker, F. M. d'Horta, C. Wight, F. Wittmann, J. Zuanon, P. A. Baker, C. C. Ribas, R. B. Norgaard, N. Filizola, A. Ansar, B. Flyvbjerg, J. C. Stevaux, Damming the rivers of the Amazon basin. *Nature* **546**, 363–369 (2017).
4. C. Tisseuil, J.-F. Cornu, O. Beauchard, S. Brosse, W. Darwall, R. Holland, B. Hugueny, P. A. Tedesco, T. Oberdorff, Global diversity patterns and cross-taxa convergence in freshwater systems. *J. Anim. Ecol.* **82**, 365–376 (2013).
5. P. A. Tedesco, O. Beauchard, R. Bigorne, S. Blanchet, L. Buisson, L. Conti, J.-F. Cornu, M. S. Dias, G. Grenouillet, B. Hugueny, C. Jézéquel, F. Leprieur, S. Brosse, T. Oberdorff, A global database on freshwater fish species occurrence in drainage basins. *Sci. Data* **4**, 170141 (2017).
6. C. Hoorn, F. P. Wesselingh, H. ter Steege, M. A. Bermudez, A. Mora, J. Sevink, I. Sanmartin, A. Sanchez-Meseguer, C. L. Anderson, J. P. Figueiredo, C. Jaramillo, D. Riff, F. R. Negri, H. Hooghiemstra, J. Lundberg, T. Stadler, T. Sarkinen, A. Antonelli, Amazonia through time: Andean uplift, climate change, landscape evolution, and biodiversity. *Science* **330**, 927–931 (2010).
7. J. G. Lundberg, M. H. Sabaj Pérez, W. M. Dahdul, O. A. Aguilera, The Amazonian neogene fish fauna, in *Amazonia: Landscape and Species Evolution*, C. Hoorn, F. P. Wesselingh, Eds. (Blackwell, 2011).
8. C. C. Fernandes, J. Podos, J. G. Lundberg, Amazonian ecology: Tributaries enhance the diversity of electric fishes. *Science* **305**, 1960–1962 (2004).
9. J. S. Albert, P. Petry, R. E. Reis, Major biogeographic and phylogenetic patterns, in *Historical Biogeography of Neotropical Freshwater Fishes*, J. S. Albert, R. E. Reis, Eds. (University of California Press, 2011), pp. 21–57.
10. F. C. P. Dagosta, M. C. C. de Pinna, The fishes of the Amazon: Distribution and biogeographical patterns, with a comprehensive list of species. *Bull. Am. Mus. Nat. Hist.* **431**, 1–163 (2019).
11. L. Castello, D. G. McGrath, L. L. Hess, M. T. Coe, P. A. Lefebvre, P. Petry, M. N. Macedo, V. F. Renó, C. C. Arantes, The vulnerability of Amazon freshwater ecosystems. *Conserv. Lett.* **6**, 217–229 (2013).
12. K. O. Winemiller, P. B. McIntyre, L. Castello, E. Fluet-Chouinard, T. Giarrizzo, S. Nam, I. G. Baird, W. Darwall, N. K. Lujan, I. Harrison, M. L. J. Stiassny, R. A. M. Silvano, D. B. Fitzgerald, F. M. Pelicice, A. A. Agostinho, L. C. Gomes, J. S. Albert, E. Baran, M. Petrere Jr., C. Zarfl, M. Mulligan, J. P. Sullivan, C. C. Arantes, L. M. Sousa, A. A. Koning, D. J. Hoeinghaus, M. Sabaj, J. G. Lundberg, J. Armbruster, M. L. Thieme, P. Petry, J. Zuanon, G. T. Vilara, J. Snoeks, C. Ou, W. Rainboth, C. S. Pavanelli, A. Akama, A. Soesbergen, L. Sáenz, Balancing hydropower and biodiversity in the Amazon, Congo, and Mekong. *Science* **351**, 128–129 (2016).
13. E. P. Anderson, C. N. Jenkins, S. Heilpern, J. A. Maldonado-Ocampo, F. M. Carvajal-Vallejos, A. C. Encalada, J. F. Rivadeneira, M. Hidalgo, C. M. Cañas, H. Ortega, N. Salcedo, M. Maldonado, P. A. Tedesco, Fragmentation of Andes-to-Amazon connectivity by hydropower dams. *Sci. Adv.* **4**, eaao1642 (2018).
14. R. E. Ricklefs, A comprehensive framework for global patterns in biodiversity. *Ecol. Lett.* **7**, 1–15 (2004).
15. R. Field, B. A. Hawkins, H. V. Cornell, D. J. Currie, J. A. F. Diniz-Filho, J.-F. Guégan, D. M. Kaufman, J. T. Kerr, G. G. Mittelbach, T. Oberdorff, E. M. O'Brien, J. R. G. Turner, Spatial species-richness gradients across scales: A meta-analysis. *J. Biogeogr.* **36**, 132–147 (2009).
16. B. Hugueny, T. Oberdorff, P. A. Tedesco, in *Community Ecology of Stream Fishes: Concepts, Approaches, and Techniques*, K. B. Gido, D. A. Jackson, Eds. (American Fisheries Society, 2010), pp. 29–62.
17. J.-F. Guégan, S. Lek, T. Oberdorff, Energy availability and habitat heterogeneity predict global riverine fish diversity. *Nature* **391**, 382–384 (1998).

18. M. A. Xenopoulos, D. M. Lodge, J. Alcamo, M. Marker, K. Schulze, D. P. van Vuuren, Scenarios of freshwater fish extinctions from climate change and water withdrawal. *Glob. Chang. Biol.* **11**, 1557–1564 (2005).

19. B. Hugueny, A. Movellan, J. Belliard, Habitat fragmentation and extinction rates within freshwater fish communities: A faunal relaxation approach. *Glob. Ecol. Biogeogr.* **20**, 449–463 (2011).

20. M. S. Dias, J.-F. Cornu, T. Oberdorff, C. A. Lasso, P. A. Tedesco, Natural fragmentation in river networks as a driver of speciation for freshwater fishes. *Ecography* **36**, 683–689 (2013).

21. M. L. Rosenzweig, *Species Diversity in Space and Time* (Cambridge Univ. Press, ed. 1, 1995).

22. A. P. Allen, J. H. Brown, J. F. Gillooly, Global biodiversity, biochemical kinetics, and the energetic-equivalence rule. *Science* **297**, 1545–1548 (2002).

23. R. Dobrovolski, A. S. Melo, F. A. S. Cassemiro, J. A. F. Diniz-Filho, Climatic history and dispersal ability explain the relative importance of turnover and nestedness components of beta diversity. *Glob. Ecol. Biogeogr.* **21**, 191–197 (2012).

24. M. S. Dias, T. Oberdorff, B. Hugueny, F. Leprieur, C. Jézéquel, J.-F. Cornu, S. Brosse, G. Grenouillet, P. A. Tedesco, Global imprint of historical connectivity on freshwater fish biodiversity. *Ecol. Lett.* **17**, 1130–1140 (2014).

25. E. M. Latrubesse, M. Cozzuol, S. A. F. da Silva-Caminha, C. A. Rigsby, M. L. Absy, C. Jaramillo, The Late Miocene paleogeography of the Amazon Basin and the evolution of the Amazon River system. *Earth Sci. Rev.* **99**, 99–124 (2010).

26. G. E. Shephard, R. D. Müller, L. Liu, M. Gurnis, Miocene drainage reversal of the Amazon River driven by plate–mantle interaction. *Nat. Geosci.* **3**, 870–875 (2010).

27. C. Jézéquel, P. A. Tedesco, J. A. Maldonado-Ocampo, S. Prada-Pedreros, H. Ortega, M. Hildalgo, K. Martens, A. De Wever, J. Zuanon, G. Torrente-Vilara, L. M. Sousa, T. Guiarrizzo, J. Valdiviezo, J. Sarmiento, M. Maldonado, F. M. Carvajal-Vallejos, T. Yunoki, C. Donascimiento, E. Agudelo, J. I. Mojica, R. R. Herrera, F. Villa, T. Oberdorff, Metadata description of the AMAZON FISH database. *Freshw. Metadata J.* **43**, 1–9 (2019).

28. D. A. Jackson, P. R. Peres-Neto, J. D. Olden, What controls who is where in freshwater fish communities the roles of biotic, abiotic, and spatial factors. *Can. J. Fish. Aquat. Sci.* **58**, 157–170 (2001).

29. R. Muneepeerakul, E. Bertuzzo, H. J. Lynch, W. F. Fagan, A. Rinaldo, I. Rodriguez-Iturbe, Neutral metacommunity models predict fish diversity patterns in Mississippi–Missouri basin. *Nature* **453**, 220–222 (2008).

30. P. V. A. Fine, Ecological and evolutionary drivers of geographic variation in species diversity. *Annu. Rev. Ecol. Evol. Syst.* **46**, 369–392 (2015).

31. D. L. Rabosky, J. Chang, P. O. Title, P. F. Cowman, L. Sallan, M. Friedman, K. Kaschner, C. Garilao, T. J. Near, M. Coll, M. E. Alfaro, An inverse latitudinal gradient in speciation rate for marine fishes. *Nature* **559**, 392–395 (2018).

32. P. J. Hanly, G. G. Mittelbach, D. W. Schemske, Speciation and the latitudinal diversity gradient: Insights from the global distribution of endemic fish. *Am. Nat.* **189**, 604–615 (2017).

33. N. P. Hitt, J. H. Roberts, Hierarchical spatial structure of stream fish colonization and extinction. *Oikos* **121**, 127–137 (2012).

34. P. A. Tedesco, F. Leprieur, B. Hugueny, S. Brosse, H. H. Dürr, O. Beauchard, F. Busson, T. Oberdorff, Patterns and processes of global riverine fish endemism. *Glob. Ecol. Biogeogr.* **21**, 977–987 (2012).

35. L. B. Beheregaray, G. M. Cooke, N. L. Chao, E. L. Landguth, Ecological speciation in the tropics: Insights from comparative genetic studies in Amazonia. *Front. Genet.* **5**, 477 (2015).

36. G. M. Cooke, N. L. Chao, L. B. Beheregaray, Marine incursions, cryptic species and ecological diversification in Amazonia: The biogeographic history of the croaker genus *Plagioscion* (Sciaenidae). *J. Biogeogr.* **39**, 724–738 (2012).

37. W. J. Matthews, *Patterns in Freshwater Fish Ecology* (Springer US, 1998).

38. C. Ibañez, J. Belliard, R. M. Hughes, P. Irz, A. Kamdem-Toham, N. Lamouroux, P. A. Tedesco, T. Oberdorff, Convergence of temperate and tropical stream fish assemblages. *Ecography* **32**, 658–670 (2009).

39. L. L. Osborne, M. J. Wiley, Influence of tributary spatial position on the structure of warmwater fish communities. *Can. J. Fish. Aquat. Sci.* **49**, 671–681 (1992).

40. T. A. Smith, C. E. Kraft, Stream fish assemblages in relation to landscape position and local habitat variables. *Trans. Am. Fish. Soc.* **134**, 430–440 (2005).

41. N. P. Hitt, P. L. Angermeier, Evidence for fish dispersal from spatial analysis of stream network topology. *J. N. Am. Benthol. Soc.* **27**, 304–320 (2008).

42. H. V. Cornell, Is regional species diversity bounded or unbounded? *Biol. Rev.* **88**, 140–165 (2012).

43. R. G. Mateo, K. Mokany, A. Guisan, Biodiversity models: What if unsaturation is the rule? *Trends Ecol. Evol.* **32**, 556–566 (2017).

44. C. Jaramillo, I. Romero, C. D'Apolito, G. Bayona, E. Duarte, S. Louwye, J. Escobar, J. Luque, J. D. Carrillo-Briceño, V. Zapata, A. Mora, S. Schouten, M. Zavada, G. Harrington, J. Ortiz, F. P. Wesselingh, Miocene flooding events of western Amazonia. *Sci. Adv.* **3**, e1601693 (2017).

45. J. Figueiredo, C. Hoorn, P. van der Ven, E. Soares, Late Miocene onset of the Amazon River and the Amazon deep-sea fan: Evidence from the Foz do Amazonas Basin: Reply. *Geology* **38**, e213 (2010).

46. H. Cheng, A. Sinha, F. W. Cruz, X. Wang, R. L. Edwards, F. M. d'Horta, C. C. Ribas, M. Vuille, L. D. Stott, A. S. Auler, Climate change patterns in Amazonia and biodiversity. *Nat. Commun.* **4**, 1411 (2013).

47. X. Wang, R. L. Edwards, A. S. Auler, H. Cheng, X. Kong, Y. Wang, F. W. Cruz, J. A. Dorale, H.-W. Chiang, Hydroclimate changes across the Amazon lowlands over the past 45,000 years. *Nature* **541**, 204–207 (2017).

48. T. F. Rangel, N. R. Edwards, P. B. Holden, J. A. F. Diniz-Filho, W. D. Gosling, M. T. P. Coelho, F. A. S. Cassemiro, C. Rahbek, R. K. Colwell, Modeling the ecology and evolution of biodiversity: Biogeographical cradles, museums, and graves. *Science* **361**, eaar5452 (2018).

49. B. U. Haq, J. Hardenbol, P. R. Vail, Chronology of fluctuating sea levels since the Triassic. *Science* **235**, 1156–1167 (1987).

50. K. G. Miller, M. A. Kominz, J. V. Browning, J. D. Wright, G. S. Mountain, M. E. Katz, P. J. Sugarman, B. S. Cramer, N. Christie-Blick, S. F. Pekar, The Phanerozoic record of global sea-level change. *Science* **310**, 1293–1298 (2005).

51. I. P. Farias, T. Hrbek, Patterns of diversification in the discus fishes (*Symphysodon* spp. Cichlidae) of the Amazon basin. *Mol. Phylogenet. Evol.* **49**, 32–43 (2008).

52. C. Hoorn, G. R. Bogotá-A, M. Romero-Baez, E. I. Lammertsma, S. G. A. Flantua, E. L. Dantas, R. Dino, D. A. do Carmo, F. Chemale Jr., The Amazon at sea: Onset and stages of the Amazon River from a marine record, with special reference to Neogene plant turnover in the drainage basin. *Global Planet. Change* **153**, 51–65 (2017).

53. K. E. Campbell Jr., C. D. Frailey, L. Romero-Pittman, The Pan-Amazonian Ucayali Peneplain, late Neogene sedimentation in Amazonia, and the birth of the modern Amazon River system. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **239**, 166–219 (2006).

54. D. F. Rossetti, M. C. L. Cohen, S. H. Tatumi, A. O. Sawakuchi, É. H. Cremon, J. C. R. Mittani, T. C. Bertani, C. J. A. S. Munita, D. R. G. Tudela, M. Yee, G. Moya, Mid-Late Pleistocene OSL chronology in western Amazonia and implications for the transcontinental Amazon pathway. *Sediment. Geol.* **330**, 1–15 (2015).

55. F. M. d'Horta, G. S. Cabanne, D. Meyer, C. Y. Miyaki, The genetic effects of Late Quaternary climatic changes over a tropical latitudinal gradient: Diversification of an Atlantic Forest passerine. *Mol. Ecol.* **20**, 1923–1935 (2011).

56. C. C. Ribas, A. Aleixo, A. C. R. Nogueira, C. Y. Miyaki, J. Cracraft, A palaeobiogeographic model for biotic diversification within Amazonia over the past three million years. *Proc. Biol. Sci.* **279**, 681–689 (2012).

57. T. Sousa-Neves, A. Aleixo, F. Sequeira, Cryptic patterns of diversification of a widespread Amazonian Woodcreeper species complex (Aves: Dendrocolaptidae) inferred from multilocus phylogenetic analysis: Implications for historical biogeography and taxonomy. *Mol. Phylogenet. Evol.* **68**, 410–424 (2013).

58. A. M. Fernandes, M. Wink, C. H. Sardelli, A. Aleixo, Multiple speciation across the Andes and throughout Amazonia: The case of the spot-backed antbird species complex (*Hylophylax naevius/Hylophylax naevioides*). *J. Biogeogr.* **41**, 1094–1104 (2014).

59. J. W. Lynch Alfaro, J. P. Boubli, F. P. Paim, C. C. Ribas, M. N. F. Silva, M. R. Messias, F. Röhe, M. P. Mercês, J. S. Silva Júnior, C. R. Silva, G. M. Pinho, G. Koshkarian, M. T. T. Nguyen, M. L. Harada, R. M. Rabelo, H. L. Queiroz, M. E. Alfaro, I. P. Farias, Biogeography of squirrel monkeys (genus *Saimiri*): South-central Amazon origin and rapid pan-Amazonian diversification of a lowland primate. *Mol. Phylogenet. Evol.* **82**, 436–454 (2015).

60. M. Ruiz-García, P. Escobar-Armel, B. de Thoisy, M. Martínez-Agüero, M. Pinedo-Castro, J. M. Shostell, Biodiversity in the Amazon: Origin hypotheses, intrinsic capacity of species colonization, and comparative phylogeography of river otters (*Lontra longicaudis* and *Pteronura brasiliensis*, Mustelidae, Carnivora) and pink river dolphin (*Inia* sp., Iniidae, Cetacea). *J. Mammal. Evol.* **25**, 213–240 (2018).

61. B. Lehner, G. Grill, Global river hydrography and network routing: Baseline data and new approaches to study the world's large river systems. *Hydrol. Process.* **27**, 2171–2186 (2013).

62. W. J. Junk, F. Wittmann, J. Schöngart, M. T. F. Piedade, A classification of the major habitats of Amazonian black-water river floodplains and a comparison with their white-water counterparts. *Wetl. Ecol. Manag.* **23**, 677–693 (2015).

63. B. Sandel, L. Arge, B. Dalsgaard, R. G. Davies, K. J. Gaston, W. J. Sutherland, J.-C. Svenning, The influence of Late Quaternary climate-change velocity on species endemism. *Science* **334**, 660–664 (2011).

64. R Core Team, *R: A Language and Environment for Statistical Computing* (R Foundation for Statistical Computing, 2017).

65. 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).

66. S. Domisch, G. Amatulli, W. Jetz, Near-global freshwater-specific environmental variables for biodiversity analyses in 1 km resolution. *Sci. Data* **2**, 150073 (2015).

67. V. Braunisch, J. Coppes, R. Arlettaz, R. Suchant, H. Schmid, K. Bollmann, Selecting from correlated climate variables: A major source of uncertainty for predicting species distributions under climate change. *Ecography* **36**, 971–983 (2013).

68. W. Thuiller, B. Lafourcade, R. Engler, M. B. Araújo, BIOMOD – a platform for ensemble forecasting of species distributions. *Ecography* **32**, 369–373 (2009).

69. J. Oksanen, F. Guillaume Blanchet, M. Friendly, R. Kindt, P. Legendre, D. McGlinn, P. R. Minchin, R. B. O'Hara, G. L. Simpson, P. Solymos, M. Henry H. Stevens, E. Szoecs, H. Wagner, vegan: Community Ecology Package (2018); https://CRAN.R-project.org/package=vegan.

Citation: T. Oberdorff, M. S. Dias, C. Jézéquel, J. S. Albert, C. C. Arantes, R. Bigorne, F. M. Carvajal-Valleros, A. De Wever, R. G. Frederico, M. Hidalgo, B. Hugueny, F. Leprieur, M. Maldonado, J. Maldonado-Ocampo, K. Martens, H. Ortega, J. Sarmiento, P. A. Tedesco, G. Torrente-Vilara, K. O. Winemiller, J. Zuanon, Unexpected fish diversity gradients in the Amazon basin. *Sci. Adv.* **5**, eaav8681 (2019).