

RESEARCH  
PAPER



# Patterns and processes of global riverine fish endemism

Pablo A. Tedesco<sup>1\*</sup>, Fabien Leprieur<sup>2</sup>, Bernard Hugué<sup>1</sup>, Sébastien Brosse<sup>5</sup>, Hans H. Dürr<sup>3</sup>, Olivier Beauchard<sup>4</sup>, Frédéric Busson<sup>1</sup> and Thierry Oberdorff<sup>1</sup>

<sup>1</sup>UMR IRD 207, Biologie des Organismes et des Ecosystèmes Aquatiques (BOREA), Département Milieux et Peuplements Aquatiques, Muséum National d'Histoire Naturelle, 43 Rue Cuvier, 75231 Paris Cedex, France, <sup>2</sup>UMR 5119 (CNRS-IFREMER-UM2-IRD), Laboratoire Ecologie des Systèmes Marins et Côtiers, Université Montpellier 2, Place Eugène Bataillon, 34095 Montpellier Cedex 5, France, <sup>3</sup>Utrecht University, Faculty of Geosciences, PO Box 80115, 3508 TC Utrecht, The Netherlands, <sup>4</sup>University of Antwerp, Faculty of Sciences, Department of Biology, Ecosystem Management Research Group, Universiteitsplein 1, BE-2610 Antwerpen (Wilrijk), Belgium, <sup>5</sup>UMR5174 (CNRS, UPS, ENFA) Laboratoire Evolution et Diversité Biologique (EDB) – Université Paul Sabatier, 118 Route de Narbonne, F-31062 Toulouse Cedex 4, France

\*Correspondence: Pablo A. Tedesco, UMR IRD 207, Biologie des Organismes et des Ecosystèmes Aquatiques, Département Milieux et Peuplements Aquatiques, Muséum National d'Histoire Naturelle, 43 Rue Cuvier, 75231 Paris Cedex, France.  
E-mail: pablo.tedesco@mnhn.fr

## ABSTRACT

**Aim** To explore global patterns of riverine fish endemism by applying an island biogeography framework to river drainage basins and highlight evolutionary mechanisms producing two kinds of endemism: neo-endemism, arising from within-drainage cladogenetic speciation, and palaeo-endemism, arising from species range contraction or anagenetic speciation.

**Location** World-wide.

**Methods** We use a uniquely comprehensive data set of riverine fish species distributions to map global fish endemism patterns. We then use the relationships between (1) total species richness and proportions of endemic species and (2) total species richness and a measure of *in situ* (i.e. within-drainage basin) probability of speciation by cladogenesis, to identify the two distinct forms of endemism. After separating drainage basins into two different sets according to dominance of one of these two forms, we apply a model averaging procedure to highlight, for both datasets, the environmental and historical variables that better explain endemism patterns. We finally analyse the effect of biotic components related to dispersal ability on the percentages of both kinds of endemism among lineages.

**Results** Our results indicate that the two types of endemism are distributed differently across space and taxonomic lineages: (1) neo-endemism, positively related to the overall richness of the drainage basin, is essentially linked to *in situ* cladogenetic speciation and is positively related to drainage basin area, negatively related to climate variability since glacial periods and negatively related to all proxies of dispersal ability; and (2) palaeo-endemism, not directly contributing to drainage basin richness, is a pure process of extinction through range contraction and/or isolation through time and is mostly related to geographic isolation, glacial history and positively related to marine-derived origin of families.

**Main conclusions** The non-random spatial and taxonomic distribution of neo-endemism and palaeo-endemism sharply reflects the role of evolutionary processes and provides a way to identify areas of high conservation interest based on their high present and future diversification potential.

## Keywords

Dispersal ability, freshwater fish, glacial maximum, global distribution, island biogeography, isolation, neo-endemism, palaeo-endemism, speciation.

## INTRODUCTION

At large geographic scales, evolutionary and ecological hypotheses have long competed as major determinants of present-day biogeographic patterns of diversity. They have shaped the past decades of research, producing numbers of biogeographic

models (MacArthur & Wilson, 1967; Goldberg *et al.*, 2005) and explanations (Dynesius & Jansson, 2000; Jablonski *et al.*, 2006; Mittelbach *et al.*, 2007) at the interface of biogeography, ecology and evolution. In this context, endemic species have always been fascinating because they should reflect the roles of speciation, extinction and dispersal ultimately responsible for their

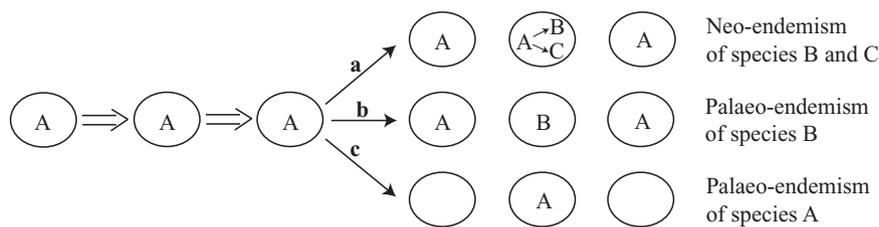
restricted geographic distribution (Jetz *et al.*, 2004; Emerson & Kolm, 2005; Rosindell & Phillimore, 2011). However, our knowledge of processes leading to endemism patterns is still largely insufficient to make a significant contribution to the debate around the driving forces of global diversity gradients (Pimm *et al.*, 1995).

Like the remote islands described in *The theory of island biogeography* (MacArthur & Wilson, 1967), drainage basins (i.e. closed systems of interconnected rivers running from their sources to the ocean) are not under equilibrium conditions for native fishes as they receive new colonists so rarely that immigration and speciation occur on similar time-scales (Hugueny *et al.*, 2010). Moreover, running waters are like ‘fish archipelagos’ (Rosenzweig, 1995), physically subdividing a species between drainage basins and limiting gene flow in ways that can promote diversification (Burrige *et al.*, 2008). Conversely, this fragmenting process also often produces small population sizes and a subsequent high extinction risk (Hugueny *et al.*, 2011). River drainage basins thus represent a specific island-like system offering an opportunity for testing ecological and evolutionary hypotheses at large spatial extents.

Our aim here is to explore global patterns of riverine fish endemism by applying an island biogeography framework to drainage basins. In drainage basins, the occurrence of endemic species (i.e. species present in a single drainage basin) can result from different configurations: the species could be (1) ‘neo-endemics’ that evolved *in situ* by cladogenesis (i.e. one species evolves into several new species) and did not disperse out of the area of origin (Fig. 1a); or (2) ‘palaeo-endemics’ that are either species that evolved *in situ* by anagenesis (i.e. one species evolves into a single new species) due to strong historical isolation (Fig. 1b, also called ‘relictualization’ by Gillespie & Roderick, 2002, or ‘endemic allospecies’ by Heaney, 2000), or relicts of previously widespread taxa, i.e. species that have experienced range contraction (Fig. 1c). On the one hand, the cladogenetic process, which generates neo-endemic species, increases both the total species richness of a drainage basin and the number and proportion of endemic species within that drainage basin. On the other hand, the anagenesis and extinction processes, which generate both kinds of palaeo-endemic species, replace

immigrant (or resident) species with endemics, and thus do not increase total species richness but do increase the number and proportion of endemic species. Accordingly, the proportion of endemics and total richness should be influenced by the type of endemism prevailing within a drainage basin (i.e. neo-endemism or palaeo-endemism). A positive relationship between the proportion of endemics and total richness is expected among drainage basins where neo-endemism dominates, while a negative relationship is expected for drainage basins where palaeo-endemism dominates. The specific case of no endemic in a given drainage basin suggests either extinction events followed by immigration of widespread species (e.g. for recent drainage basins previously covered by ice sheets; Reyjol *et al.*, 2007), or unsuitable conditions for *in situ* speciation (e.g. minimum area for speciation; Losos & Schluter, 2000).

To test this hypothesis we use a uniquely comprehensive database of riverine fish species distributions (i.e. 9750 species corresponding to 80% of all freshwater species described, and 1058 drainage basins covering more than 80% of the earth’s terrestrial surface; Blanchet *et al.*, 2010; Leprieur *et al.*, 2011). We first map global patterns of endemism at the drainage basin grain. We then formally identify the two forms of endemism, analysing the relationships between total species richness, proportions of endemics and an index of *in situ* cladogenetic speciation proposed by Coyne & Price (2000). Finally, we examine the relative roles of evolutionary and ecological processes by relating the patterns of neo-endemism and palaeo-endemism to the environmental characteristics, geographic and topographic isolation and current and past climatic conditions of drainage basins. Since the processes generating neo- and palaeo-endemic species are different, we expect that neo- and palaeo-endemic patterns will be explained by different environmental factors. According to theory, rates of neo-endemism should be positively related to drainage basin size and heterogeneity, both factors supposed to promote speciation (Losos & Schluter, 2000) and to decrease extinction (MacArthur & Wilson, 1967), and negatively related to past climatic variability, a factor supposed to increase species extinction risk (Dynesius & Jansson, 2000). We have no strong a priori expectation regarding factors acting on palaeo-endemic patterns, although past climatic variability, related to glacial



**Figure 1** The origin of drainage basin endemism types adapted from the Emerson & Kolm (2005) origin of island endemism scheme. Species A originates from the westernmost drainage basin and then colonizes the remaining ones during a time period of connection between basins. After isolation of drainage basins, there are three possible mechanisms that may result in endemism in the middle drainage basin: (a) neo-endemism of species B and C is the result of intra-drainage (i.e. *in situ*) cladogenetic speciation from the founding species A; (b) palaeo-endemism of species B is the result of inter-drainage anagenetic speciation through isolation, and represents a surviving element of the ancestral connected drainage basin (Gillespie & Roderick, 2002); and (c) palaeo-endemism of species A results from extinction of species A in the first and third drainage basins (i.e. range contraction).

history, should negatively affect palaeo-endemic richness (Reyjol *et al.*, 2007). Finally, since level of isolation is an important driver of endemism, past connections between neighbouring drainage basins should negatively affect both kinds of endemism.

The above framework considers diversity within a drainage basin as an outcome of solely environmental and geographical influences, taking a neutral view on organisms themselves. However, clade history and biological traits may also greatly contribute to explaining species distribution patterns at broad spatial scales (e.g. Knouft & Page, 2011). As a complement to our environmental approach relating patterns of neo-endemism and palaeo-endemism to drainage basin characteristics, we analysed biotic components of lineages at the family level to further explain and confirm patterns of endemism and the processes behind them. While higher speciation rates are expected in regions considered large enough for populations to be sufficiently isolated, the influence of fragmentation on speciation should also depend on the dispersal ability of organisms. For instance, as body size has been positively related to dispersal distance in active dispersers, including fish (e.g. Tales *et al.*, 2004; Jenkins *et al.*, 2007), small-bodied species may experience greater isolation and lower gene flow, and thus have a greater potential for speciation. Furthermore, the current geographic distribution range of lineages and their potentially marine-derived origin can also be viewed as proxies of dispersal ability at large temporal scales (Waters & Wallis, 2001; Lovejoy *et al.*, 2006). Because neo-endemic species are generated from *in situ* cladogenetic speciation without subsequent dispersal, we expect proportions of neo-endemism among families to be negatively related to dispersal ability. Inversely, since dispersal is part of the process producing palaeo-endemic relict species, dispersal ability should be positively related to proportions of palaeo-endemic species.

## MATERIALS AND METHODS

### Species richness and endemism data

We used a database containing fish species occurrence (comprising hagfish, lampreys and cartilaginous and bony fish) for 1058 drainage basins distributed world-wide (Blanchet *et al.*, 2010; Leprieur *et al.*, 2011). This database was compiled from an extensive literature survey (bibliographic sources including published papers, books and grey literature; references available upon request) of check lists of native freshwater fish species. Only complete species lists at the drainage basin scale were considered. Only strictly freshwater native species were included (secondary or migratory euryhaline fishes were systematically withdrawn on the basis of species descriptions provided in Fishbase; Froese & Pauly, 2010) in order to consider each drainage basin as a closed system (Hugueny *et al.*, 2010). In addition, ancient lakes (Malawi, Tana, Kivu, Tanganyika, Albert, Edward, Victoria, Turkana, Kyoga, Baikal, Titicaca and Biwa) and their fish fauna were removed from the database to avoid any bias

related to differences in speciation processes between ancient lakes and rivers (Lévêque *et al.*, 2008).

Total species richness refers to the total number of native fish species present in a drainage basin. Endemic species richness refers to the total number of endemic species present in a drainage basin. Endemic species are defined here as those inhabiting only one drainage basin (i.e. single-drainage endemics as the analogy of 'single-island endemics' often applied in insular systems); for those species, an extended literature search was conducted in order to guarantee their endemic status. Although the endemic status of organisms can have a differing biological significance depending on the geographic scale under consideration, riverine fish constitute an excellent group to unambiguously define endemism because of the insular nature of rivers for these organisms (see also Appendix S1 in Supporting Information).

### Current and historical environmental data

For each drainage basin, we compiled 11 variables to evaluate the effect of environmental and historical drainage basin conditions on levels of neo-endemism and palaeo-endemism. These variables are related to drainage basin size and environmental heterogeneity, geographic and topographic isolation, and present-day and historical climate (see Table S1 for a list of variables, corresponding references and units). A geographic information system of the global river systems and the corresponding drainage basin limits (Vörösmarty *et al.*, 2000) was used to extract absolute or mean values for each of these variables.

We collected four variables related to habitat size and heterogeneity: drainage surface area (DSA), habitat heterogeneity (HH), elevation range ( $Elev_{range}$ ) and mean annual runoff ( $RO_{ann}$ ); and four variables related to contemporary climate: temperature (T), precipitation (PR), precipitation seasonality ( $PR_{seasonality}$ ) and actual evapotranspiration ( $AET_{ann}$ ). Elevation range (i.e. the difference between maximum and minimum elevations within a given drainage basin) is commonly used as a proxy of topographic heterogeneity in large-scale ecological studies and can be considered as a rough surrogate variable reflecting historical opportunities for allopatric speciation (Jetz & Rahbek, 2002; Jetz *et al.*, 2004; Melo *et al.*, 2009; Leprieur *et al.*, 2011). Habitat heterogeneity was estimated by applying the Shannon's diversity index to proportions of biomes (i.e. vegetation types associated with regional variations in climate) within drainage basins (see Appendix S1 for further details on habitat and climatic variables).

We quantified a historically based variable related to Quaternary climate variability (i.e. Milankovitch climate oscillations), responsible for large temperature changes toward the poles. Quaternary climate variability was measured as the change in mean annual temperature between the present and the Last Glacial Maximum (LGM, c. 21 ka) estimated from six different global circulation models (GCMs) (Jansson & Davies, 2008). This measure closely relates to the amplitude of Quaternary climatic oscillations (Jansson, 2003). For each GCM, the change

in mean annual temperature between the present and the LGM was calculated and the resulting values were averaged to produce an overall anomaly variable,  $T_{\text{anomaly}}$  (Jansson & Davies, 2008).

Following Oberdorff *et al.* (1999), we also considered whether or not a drainage basin was on a land mass, a peninsula, or an island (LPI; continental mass = 0; peninsula = 1; island = 2). This can be related to geographical isolation since colonization rates in peninsulas and islands should be sufficiently low to favour the formation and maintenance of endemic species. Topographic isolation of drainage basins was estimated according to the probability of river basin connection during the Quaternary sea level fluctuations (e.g. Swartz *et al.*, 2007). At glacial periods the sea level drastically regressed, potentially connecting neighbouring drainage basins, reducing isolation and allowing for species dispersal processes. The minimum linear distance from river mouth to the bathymetry level  $-120$  m ( $\text{Dist}_{120}$ ), i.e. the sea level regression at the LGM (Lambeck *et al.*, 2002), was used as a measure of the probability of river basin connection.

### ***In situ* cladogenetic speciation index**

As an index of *in situ* speciation by cladogenesis (P), we used the proportion of endemic lineages that have diverged into two or more descendant species within a given drainage basin. This index was estimated for each drainage basin by dividing the number of genera with two or more endemic species (i.e. representing the number of lineages that have diversified *in situ* by cladogenesis) by the number of genera with one or more endemic species (i.e. representing the number of lineages that could have speciated within the drainage basin). This method was proposed by Coyne & Price (2000) and has been applied to a wide variety of organisms inhabiting different island systems (Kisel & Barraclough, 2010). Ideally, information showing that endemic species from the same genera are actually sister species (i.e. not arising from multiple colonization events) should be necessary to consider *in situ* diversification as the mechanism responsible. Unfortunately this information is not available for every eventual case of *in-situ* cladogenetic speciation. However, we systematically searched in the literature for phylogenetic information involving each genus suspected of *in situ* speciation. The search was applied on Google Scholar using 'phylogeny' and the targeted genus name as inputs. Cases of congeneric endemic species shown not to be closest relatives (i.e. sister taxa) were then reclassified as non-speciated lineages, i.e. palaeo-endemics (Kisel & Barraclough, 2010).

### **Present and historical correlates of endemism**

Based on our index of *in situ* speciation (see above), we split drainage basins into two categories (see Fig. S1): (1) drainage basins likely to have undergone *in situ* cladogenetic speciation (cladogenesis index  $P > 0$ ) and (2) other drainages (cladogenesis index  $P = 0$ ). For these two sets of drainage basins, we further built separate statistical models to assess the relative importance of each of the 11 environmental and historical variables in shaping the observed patterns of endemism. Generalized linear

models (GLMs) were fitted to proportions of endemic species (as a response variable) with 'quasibinomial' distribution errors and total richness values as weights. To determine the relative importance of the explanatory variables for each set of drainage basins, we ran models for all possible combinations of the explanatory variables and then performed model averaging based on the 'quasi-Akaike information criterion' (QAIC). As a cut-off criterion to delineate a 'top model set' providing average parameter estimates and confidence intervals, we used fitted models with  $\Delta\text{QAIC} < 2$  (Grueber *et al.*, 2011). To verify the significant effect of the explanatory variables selected by the model averaging procedure, we fitted a GLM with variables having a relative importance value of more than 0.5. Finally, to evaluate the potentially remaining spatial autocorrelation in the model residuals, we computed Moran's  $I$  indices between neighbouring drainage basins (Sokal & Oden, 1978).

In recent years, several authors have demonstrated that the number or proportion of single-island endemic species is not independent of the total richness of the island. These authors proposed null models to account for patterns of island endemism attributable to the statistical properties of random samples (e.g. Birand & Howard, 2008). To confirm that our results are not affected by this confounding factor, we applied the same statistical procedures as described above to levels of endemism generated by a null model (see Appendix S1 for a detailed description of the null model).

### **Family-level analysis**

Our taxonomic approach was based on three variables used as proxies for lineage dispersal ability: mean body size, geographic distribution range and proportion of marine or brackish species within families. Mean values of fish family body size was based on the maximum body length of species reported in FishBase (Froese & Pauly, 2010). The geographic distribution range of families was estimated as the product of latitudinal and longitudinal ranges using the median values of the drainage basins and the species occurrences from our database. The proportion of marine or brackish species (i.e. not strictly freshwater) within families was used as a proxy for a marine origin of lineages. This proportion was based on the water salinity preferences of species reported in FishBase (Froese & Pauly, 2010).

We added to the analysis the absolute mean latitude of occurrence of families and the mean surface area of the drainage basins occupied to control for their potential effects on body size (Biedermann, 2003; Blanchet *et al.*, 2010). We used generalized linear mixed models (GLMMs), with 'binomial' distribution errors and total family richness as weights, the proportions of neo-endemic and palaeo-endemic species being the dependent variables. To account for the taxonomic relatedness of families, potentially leading to pseudo-replication, taxonomic orders were used as a random factor. Prior to analysis, some variables were transformed to improve normality (see Table S1). Multicollinearity was checked using a variance inflation factor procedure.

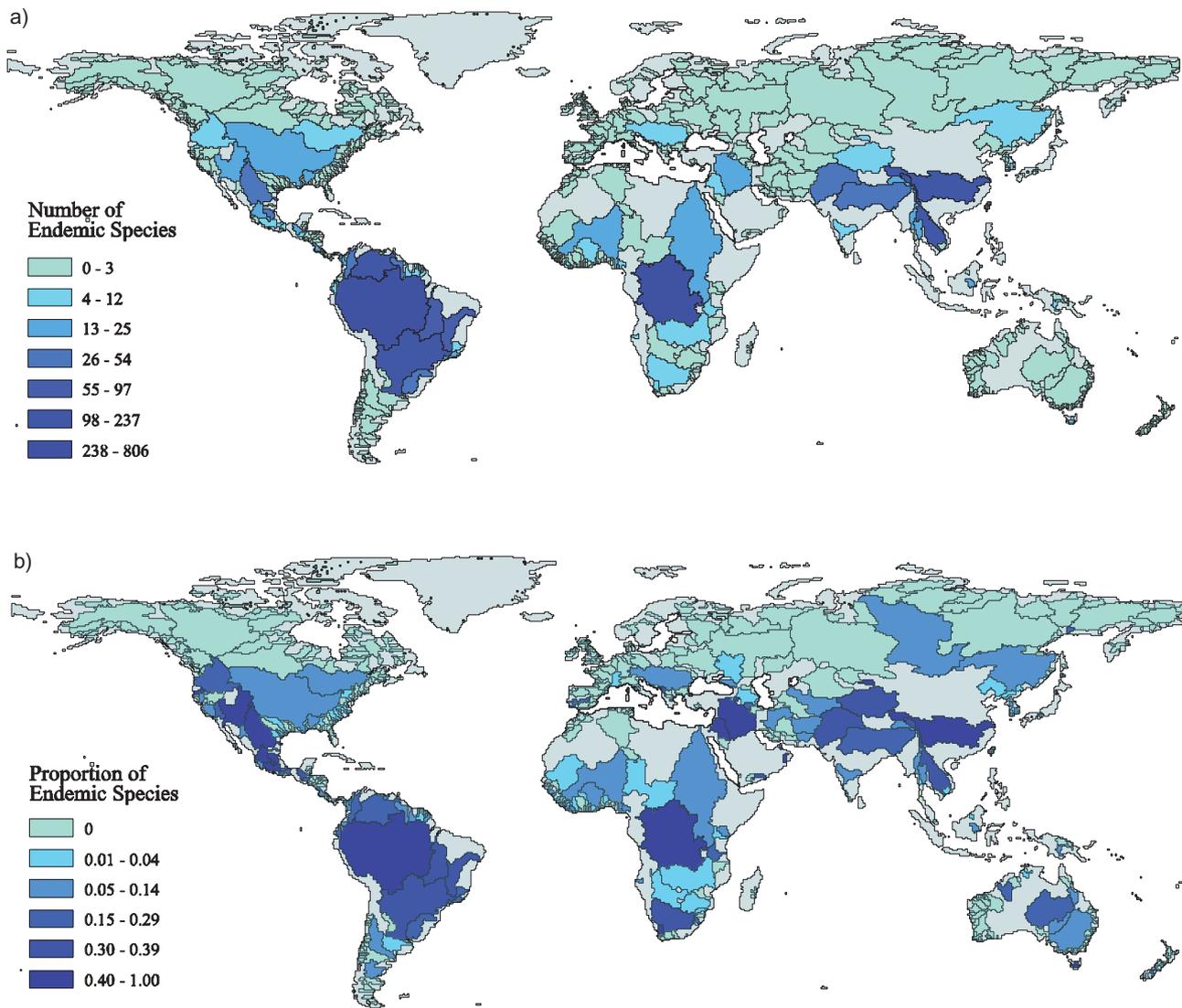


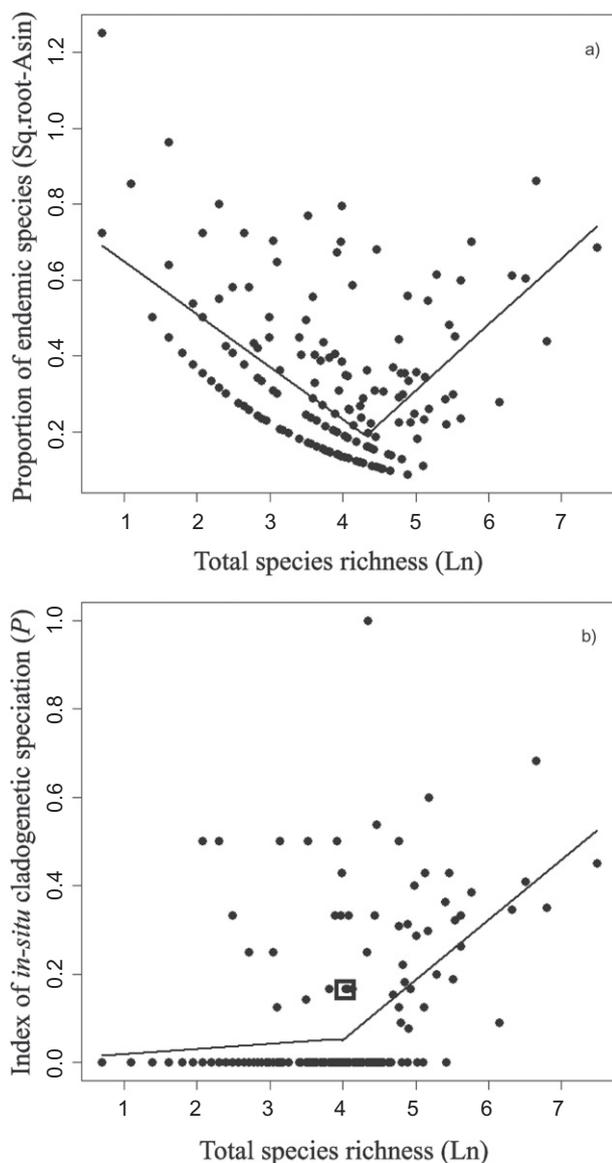
Figure 2 Global maps presenting different aspects of endemism per drainage basin: (a) numbers of endemic species; (b) proportions of endemic species.

## RESULTS

The global distribution of riverine fish endemism is summarized in Fig. 2. The largest numbers and proportions of endemic species are concentrated in large, species-rich tropical drainage basins, and to a lesser extent in large drainages from dry and temperate regions (Fig. 2a,b). More strikingly, Fig. 2(b) shows a high proportion of endemism in non-tropical regions of North and Central America, southern Africa, Australia, the Middle East and Asia. These maps also reveal the absence of endemic species in small coastal drainage basins world-wide and in most northern and southern drainage basins of temperate regions, as well as polar regions (Fig. 2a,b). Outputs from the null model (see Appendix S1) are coherent with the patterns described above, highlighting drainage basins with significantly more endemic species than expected by chance in tropical and arid regions of the world (dark blue drainage basins in Fig. S2). However, null

model outputs also highlight significantly lower levels of endemism for a large portion of Africa and all drainages from Guyana, Suriname and French Guiana (dark red drainage basins in Fig. S2).

In accordance with our palaeo-endemic/neo-endemic framework (see Fig. 1), we observe a U-shaped relationship between the proportion of endemics and total drainage basin species richness (Fig. 3a). Interpretation of this result as a differential dominance of either palaeo-endemics or neo-endemics is confirmed by the nonlinear positive relationship observed between our index of *in situ* cladogenetic speciation and overall species richness (Fig. 3b). The index of *in situ* speciation also points to the existence of a minimal surface area allowing for the presence and maintenance of neo-endemic species within drainage basins. This minimal area, about 4500 km<sup>2</sup>, corresponds to the smallest drainage basin where a non-zero index value was found (black square in Fig. 3b).



**Figure 3** (a) Proportions of endemic species per drainage basin as a function of total species richness considering drainage basins with at least one endemic species. (b) Relationship between our index of within-drainage cladogenetic speciation (Coyne & Price, 2000) and total species richness. Trend lines in both relationships result from segmented regression models with one break-point. Noticeably, the break-point estimates from both relationships do not differ significantly ( $4.31 \pm 0.24$  and  $3.98 \pm 0.54$ , respectively for A and B). The U-shaped relationship (a) is supported by a segmented regression ( $n = 181$ ,  $R^2 = 0.40$ ,  $P < 0.0001$ ), significantly performing better than a linear one ( $F_{2,180} = 46.16$ ,  $P < 0.0001$ ). The broken-stick relationship (b) is supported by a segmented regression ( $n = 181$ ,  $R^2 = 0.26$ ,  $P = 0.0001$ ), significantly performing better than a linear one ( $F_{2,180} = 9.50$ ,  $P < 0.0001$ ). The black square in panel (b) shows the position of the smallest drainage basin (approximately  $4500 \text{ km}^2$ ) with non-zero index of within-drainage cladogenetic speciation.

Results of GLM models for drainage basins dominated respectively by neo-endemics and palaeo-endemics are summarized in Table 1. For both sets of drainage basins, GLM models show higher levels of endemism in rivers that experienced stable Quaternary climatic conditions (as measured by the temperature anomalies from the LGM to the present) (Table 1). However, in drainage basins where neo-endemism dominates, higher levels of endemism are also strongly positively related to habitat size (i.e. drainage basin area) and negatively related to topographical isolation (as measured by the distance from river mouth to bathymetry level  $-120 \text{ m}$ ) and to temperature, although in a nonlinear way for the latter (Table 1). In contrast, for drainage basins where palaeo-endemism dominates, higher levels of endemism are found in geographically isolated (i.e. drainage basins situated on islands or peninsulas) and colder regions (Table 1). Results of these two models thus clearly emphasize the presence of different processes driving neo- and palaeo-endemism patterns and are corroborated by results obtained using the null model approach (see Table S2). Although using outputs from the null model revealed some differences compared with results from patterns of endemism proportions (compare Tables 1 & S2), both approaches highlighted: (1) the negative effect of Quaternary climate variability for neo- and palaeo-endemism; (2) the positive effect of habitat-related variables for neo-endemism (drainage basin area for both approaches and elevation range for the null model approach); (3) the negative effect of energy-related variables (temperature for both approaches and actual evapotranspiration for the null model approach) for palaeo-endemism; (4) the negative effect of drainage basin area for palaeo-endemism; and (5) the positive effect of geographic isolation for palaeo-endemism patterns.

Results of GLMM models confirm those from our environmental approach, showing greater proportions of neo-endemic species in lineages from lower latitudes and inhabiting larger drainage basins (Table 2). More importantly, these GLMM models also point out the significant negative influence of dispersal ability (either based on species body size, marine origin or geographic range) on proportions of neo-endemic species and a positive influence of dispersal ability (i.e. probability of marine-derived origin) on proportions of palaeo-endemic species among families (Table 2).

## DISCUSSION

### Neo- and palaeo-endemism

The terms 'neo-endemic' and 'palaeo-endemic' have usually been used to broadly identify recent and ancient endemic species, respectively (Lomolino *et al.*, 2006). Following recent works (Gillespie & Roderick, 2002; Mansion *et al.*, 2009), here we defined these terms in a more precise way to explore the origins of endemism in riverine systems. Neo-endemics, classically referring to species that have differentiated where they are found today, were here restricted to those that have been able to speciate by cladogenesis. Palaeo-endemics, classically referring

**Table 1** Results of the GLM averaging approach obtained after splitting drainage basins into two categories: (a) drainage basins likely to have undergone *in situ* speciation (*in situ* index of cladogenetic speciation  $P > 0$ ,  $n = 52$ ) versus (b) other drainages ( $P = 0$ ,  $n = 132$ ). Standardized Moran's  $I$ -values of model residuals for neighbouring (i.e. adjacent) drainage basins were  $I_{P>0} = 0.73$  (n.s.) and  $I_{P=0} = 1.83$  (n.s.), respectively for (a) and (b) (see Appendix S1 for further specifications). Pseudo  $R^2$  indicates the proportion of explained deviance.

Explanatory variables	Relative variable importance	Average standardized		GLM standardized coefficient	$t$ -value
		coefficient	Confidence interval (95%)		
(a) Drainage basins dominated by neo-endemism (i.e. with <i>in situ</i> index of speciation $> 0$ )					
DSA	1	0.724	(0.435; 1.01)	0.718	5.75***
DSA <sup>2</sup>	0.92	0.190	(-0.0226; 0.404)	0.194	2.12*
HH	0.13	0.029	(-0.0864; 0.144)		
HH <sup>2</sup>	0.04	-0.001	(-0.0102; 0.00836)		
Elev <sub>range</sub>	0.13	0.021	(-0.0687; 0.11)		
Elev <sub>range</sub> <sup>2</sup>	0.07	-0.005	(-0.0341; 0.0248)		
RO <sub>ann</sub>	0.09	0.018	(-0.0686; 0.105)		
RO <sub>ann</sub> <sup>2</sup>	0.13	0.020	(-0.0655; 0.105)		
AET <sub>ann</sub>	0.8	0.280	(-0.174; 0.734)	0.280	1.72
T	1	-0.845	(-1.37; -0.321)	-0.860	-4.04***
T <sup>2</sup>	0.96	-0.291	(-0.607; 0.0249)	-0.315	-2.18*
PR	0.09	-0.012	(-0.101; 0.0778)		
PR <sub>seasonality</sub>	0.2	0.042	(-0.113; 0.197)		
LPI	0.08	0.009	(-0.0373; 0.0544)		
Dist <sub>120</sub>	1	-0.373	(-0.651; -0.0958)	-0.365	-2.92**
T <sub>anomaly</sub>	1	-0.807	(-1.38; -0.238)	-0.779	-2.91**
Pseudo $R^2$				0.70	
(b) Drainage basins dominated by palaeo-endemism (i.e. with <i>in situ</i> index of speciation = 0)					
DSA	0.76	-0.139	(-0.381; 0.103)	-0.144	-1.4
DSA <sup>2</sup>	0.01	0.000	(-0.00206; 0.00219)		
HH	0.05	0.006	(-0.0215; 0.0332)		
HH <sup>2</sup>	0.01	0.000	(-0.00226; 0.00254)		
Elev <sub>range</sub>	0.2	0.029	(-0.0803; 0.139)		
Elev <sub>range</sub> <sup>2</sup>	0.7	-0.078	(-0.229; 0.0727)	-0.116	-1.7
RO <sub>ann</sub>	0.73	0.315	(-0.222; 0.851)	0.303	1.75
RO <sub>ann</sub> <sup>2</sup>	0.78	0.140	(-0.0702; 0.351)	0.145	1.79
AET <sub>ann</sub>	0.8	-0.294	(-0.706; 0.117)	-0.514	-3.3**
T	1	-0.393	(-0.709; -0.0765)	-0.337	-2.2*
T <sup>2</sup>	0.01	0.000	(-0.00291; 0.00308)		
PR	0.43	-0.207	(-0.742; 0.329)	-0.404	-1.7
PR <sub>seasonality</sub>	0.73	-0.183	(-0.499; 0.133)	-0.173	-1.8
LPI	1	0.306	(0.127; 0.485)	0.290	3.3**
Dist <sub>120</sub>	0.43	-0.058	(-0.229; 0.113)		
T <sub>anomaly</sub>	1	-0.513	(-0.788; -0.239)	-0.514	-3.9***
Pseudo $R^2$				0.44	

Abbreviations: DSA, drainage surface area; HH, habitat heterogeneity; Elev<sub>range</sub>, elevation range; T, temperature; PR, precipitation; PR<sub>seasonality</sub>, precipitation seasonality; AET<sub>ann</sub>, actual evapotranspiration; RO<sub>ann</sub>, annual runoff; T<sub>anomaly</sub>, temperature anomaly from the Last Glacial Maximum to the present; Dist<sub>120</sub>, distance to bathymetry level -120 m; LPI, land-peninsula-island. Significance levels: \*\*\* $P < 0.001$ ; \*\* $P < 0.01$ ; \* $P < 0.05$ ; n.s.  $P > 0.05$ .

to the narrowly endemic descendants of once widespread taxa, here also include species that have differentiated by anagenesis from an ancestral immigrant. These definitions allow the separation of drainage basins producing new species from the others. In agreement with our expectations, we found a highly significant U-shaped relationship between the proportion of endemics and overall species richness (Fig. 3a), which supports a concomitant contribution of neo- and palaeo-endemism in generating the overall pattern. In addition, the positive relation-

ship we found between the index of *in situ* cladogenetic speciation and species richness clearly supports the presence of greater proportions of neo-endemics in the richest drainage basins and palaeo-endemics in the poorest ones.

### Present and historical correlates of endemism

Overall, the global patterns in endemic riverine fish distributions show some broad similarities with patterns already

Variables	Proportions of neo-endemics		Proportions of palaeo-endemics	
	Estimate	<i>t</i> -values	Estimate	<i>t</i> -values
BS <sub>mean</sub>	-0.358	-6.694***	0.033	0.570
Marine <sub>prop</sub>	-0.973	-5.056***	0.451	2.776**
Geo <sub>range</sub>	-0.027	-2.356*	-0.013	-0.973
Lat <sub>mean</sub>	-0.018	-4.151***	0.001	0.251
DSA <sub>mean</sub>	0.414	9.892***	-0.023	-0.507

Abbreviations: BS<sub>mean</sub>, mean body size; Marine<sub>prop</sub>, proportion of marine species; Geo<sub>range</sub>, geographical range; Lat<sub>mean</sub>, mean occurrence latitude; DSA<sub>mean</sub>, mean drainage surface area.

Significance levels: \*\*\* $P < 0.001$ ; \*\* $P < 0.01$ ; \* $P < 0.05$ ; n.s.,  $P > 0.05$ .

described for other taxa (Grenyer *et al.*, 2006), with fish endemics being mostly concentrated in drainage basins from high-energy tropical regions. However, contrasting with patterns found for other taxa, high levels of fish endemism are also found in drainage basins from low-energy regions of higher latitudes. A number of recent reviews involving many taxa (e.g. Field *et al.*, 2009) have suggested that global species richness gradients most often relate to factors that can be roughly classified into four major categories: (1) habitat size and heterogeneity; (2) current climate and productivity; (3) historical conditions of climate; and (4) isolation. Patterns of riverine fish species richness are no exception, as previous studies recognized a positive role of habitat size, heterogeneity and productivity and a negative role of isolation on drainage basin species richness (Oberdorff *et al.*, 1995). Concerning endemic species, previous studies (including riverine fish studies) showed that, apart from overall species richness, a combination of factors related to contemporary environmental conditions and/or historical climate can broadly explain patterns of endemic species richness at regional and continental scales (Oberdorff *et al.*, 1999; Jetz *et al.*, 2004). Here we provide evidence that distinguishing between neo-endemism and palaeo-endemism can shed light on the processes driving riverine fish endemism. For instance, results clearly show higher levels of neo-endemism in large drainage basins having experienced stable climatic conditions through geological time (i.e. from Quaternary glaciations to the present), suggesting a strong link between neo-endemism and overall species richness patterns. In contrast, higher levels of palaeo-endemism are found in smaller and geographically isolated drainage basins situated in low-energy regions, suggesting that palaeo-endemics are clearly not a random subset of the overall species richness. Overall, these results highlight two different processes in generating endemism, as neo-endemism was mostly related to factors enhancing speciation (i.e. large drainage areas and climatic stability), while palaeo-endemism was related to factors enhancing extinction and isolation (i.e. small drainage areas and strong geographic isolation) favouring the formation of endemic species through the process of relictualization (Gillespie & Roderick, 2002).

We also noticed the presence of a threshold surface area, about 4500 km<sup>2</sup>, under which no cladogenetic *in situ* speciation seems possible within a drainage basin. We suggest that this last

**Table 2** Results of the generalized linear mixed models used to test the relationship between dispersal ability predictors, latitude (absolute values) and drainage area on the proportions of both kinds of endemic species among fish families.

pattern is due to the fact that small drainage basins, usually less topographically complex and more ephemeral than their larger counterparts, do not allow for allopatric speciation (Losos & Ricklefs, 2009). This minimal surface area falls within estimates from different island systems ranging from approximately 1 to 10,000 km<sup>2</sup> depending on the taxa considered (Losos & Schluter, 2000; Kisel & Barraclough, 2010).

### The biotic component of endemism

Proportions of neo-endemic and palaeo-endemic species also showed specific trends between families. The non-random taxonomic distribution of these two forms of endemism in relation to surrogate variables of species dispersal ability reflects the role of isolation in the evolutionary processes generating current patterns. Our three proxies of dispersal ability (i.e. body size, probability of marine origin and geographic range of families) were negatively related to proportions of neo-endemic species, suggesting antagonism between high immigration rates and the *in situ* cladogenetic speciation process. This finding is in accordance with the negative effect of the probability of river capture related to Quaternary sea level fluctuations on the spatial distribution of neo-endemism. The higher proportions of neo-endemic species found in small-bodied families also accords with the faster rates of molecular evolution associated with small-bodied species (Martin & Palumbi, 1993; Bromham, 2011).

Proportions of palaeo-endemic species were positively related to the probability of marine origin of lineages, suggesting colonization of freshwaters through marine dispersal, generating palaeo-endemic species by subsequent range contraction or anagenetic speciation. Fish transitions from marine to freshwater habitats have been previously reported and hypothesized as consequences of palaeo-environmental changes, incursions of marine waters into formerly freshwater systems and the reorientation of major river drainages (e.g. Waters & Wallis, 2001; Lovejoy *et al.*, 2006). Our results suggest that at least some of the palaeo-endemic species identified are unique freshwater representatives of marine clades.

### Limits and perspectives

The geographic distribution of relict and more derived clades has long been considered a key issue to elucidate the puzzle of

evolution and diversity (Stehli *et al.*, 1969). For instance, recent process-based models and empirical data have demonstrated that macroevolutionary source regions (i.e. centres of origin or 'cradles') must have a high proportion of young taxa. Conversely, macroevolutionary sink regions (i.e. centres of accumulation or 'museums') must be dominated by older taxa (Goldberg *et al.*, 2005; Jablonski *et al.*, 2006). According to the neo/palaeo-endemic dichotomy framework, this suggests that macroevolutionary sources should harbour higher proportions of neo-endemic species, and macroevolutionary sinks higher proportions of palaeo-endemic species, if any. However, since anagenesis also produces diversity, this hypothesis remains to be tested by disentangling anagenetic (Fig. 1b) from relict (Fig. 1c) palaeo-endemic species using distributional, phylogenetic and fossil information.

The time since isolation of each drainage basin would also be extremely helpful to disentangle the respective roles of colonization and speciation in generating current diversity patterns. Unfortunately this kind of information is not available, at least for global-scale analyses like the one intended here. More regionally focused approaches should bring interesting insights (e.g. Reyjol *et al.*, 2007; Filipe *et al.*, 2009). For instance, results from the null model approach revealed fewer endemic species than expected in three specific tropical/subtropical regions: Guyana, Suriname and French Guiana, tropical West Africa and the Okavango–Zambezi region (see Fig. S2). This finding is in line with evidence showing that river systems in these regions have experienced recent connections allowing for recent colonization (Dürr, 2003) and a subsequent decrease in levels of endemism (see Fig. S2). The biogeographic importance of past rearrangements of drainage networks, allowing dispersal between currently isolated river systems, has been pointed out in recent phylogenetic studies (e.g. Swartz *et al.*, 2007; BurrIDGE *et al.*, 2008). In agreement with these approaches, we found here a negative effect of the probability of river basin connection during the LGM sea level regression on neo-endemism (Table 1). However, other geomorphological processes (e.g. episodic swamp or tributary connections) can act on dispersal processes in freshwater-limited taxa (see BurrIDGE *et al.*, 2008), thus also potentially affecting current patterns of endemism. More precise information on past river configurations should bring new insights into the mechanistic basis leading to patterns of neo/palaeo-endemism.

### Conservation implications

Global hotspots of overall richness and endemism have been widely used to determine priority areas for conservation. However, overall richness and endemism patterns usually do not show spatial congruence either between or within taxonomic classes (e.g. Orme *et al.*, 2005). By pointing out the nonlinear relationship existing between endemism and overall richness (Fig. 3a) and by determining the factors and processes responsible for this nonlinearity, our results provide a mechanistic explanation for this lack of congruence. Results further suggest that disentangling palaeo-endemic from neo-endemic species

when mapping biodiversity hotspots may be an interesting prospect for future conservation practices (Brooks *et al.*, 2006). Indeed, priorities established on the basis of hotspots of overall richness may successfully conserve areas of high neo-endemism and consequently the main centres of evolutionary processes (Mace *et al.*, 2003). As the processes generating neo- and palaeo-endemic species in one place are not mutually exclusive, this strategy should also partly benefit palaeo-endemic species. However, when considering global conservation priorities, the choice to specifically include palaeo-endemic relict species already on the verge of natural extinction remains an open question (Erwin, 1991).

### ACKNOWLEDGEMENTS

We thank the French National Agency for Research (ANR-06-BDIV-010) and the BIOFRESH project supported by the European Commission (FP7-ENV-2008, contract no. 226874) for their financial support. We are very grateful to Bradford Hawkins, E. García-Berthou, Sean Marr and three anonymous referees for helpful comments that greatly improved the manuscript. English has been edited by P. Winterton who kindly revised the manuscript and we thank J. F. Cornu for his assistance with mapping.

### REFERENCES

- Biedermann, R. (2003) Body size and area–incidence relationships: is there a general pattern? *Global Ecology and Biogeography*, **12**, 381–387.
- Birand, A. & Howard, D.J. (2008) The relationship between proportion of endemics and species diversity on islands: expectations from a null model. *Ecography*, **31**, 286–288.
- Blanchet, S., Grenouillet, G., Beauchard, O., Tedesco, P.A., Leprieur, F., Dürr, H.H., Busson, F., Oberdorff, T. & Brosse, S. (2010) Non-native species disrupt the worldwide patterns of freshwater fish body size: implications for Bergmann's rule. *Ecology Letters*, **13**, 421–431.
- Bromham, L. (2011) The genome as a life-history character: why rate of molecular evolution varies between mammal species. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **366**, 2503–2513.
- Brooks, T.M., Mittermeier, R.A., Da Fonseca, G.A.B., Gerlach, J., Hoffmann, M., Lamoreux, J.F., Mittermeier, C.G., Pilgrim, J.D. & Rodrigues, A.S.L. (2006) Global biodiversity conservation priorities. *Science*, **313**, 58–61.
- BurrIDGE, C.P., Craw, D., Jack, D.C., King, T.M. & Waters, J.M. (2008) Does fish ecology predict dispersal across a river drainage divide? *Evolution*, **62**, 1484–1499.
- Coyne, J.A. & Price, T.D. (2000) Little evidence for sympatric speciation in island birds. *Evolution*, **54**, 2166–2171.
- Dürr, H.H. (2003) *Vers une typologie des systèmes fluviaux à l'échelle globale: quelques concepts et exemples à résolution moyenne*. PhD Thesis. Université Paris VI, Paris.
- Dynesius, M. & Jansson, R. (2000) Evolutionary consequences of changes in species' geographical distributions driven by

- Milankovitch climate oscillations. *Proceedings of the National Academy of Sciences USA*, **97**, 9115–9120.
- Emerson, B.C. & Kolm, N. (2005) Species diversity can drive speciation. *Nature*, **434**, 1015–1017.
- Erwin, T.L. (1991) An evolutionary basis for conservation strategies. *Science*, **253**, 750–752.
- Field, R., Hawkins, B.A., Cornell, H.V., Currie, D.J., Diniz-Filho, J.A.F., Guégan, J.F., Kaufman, D.M., Kerr, J.T., Mittelbach, G.G., Oberdorff, T., O'Brien, E.M. & Turner, J.R.G. (2009) Spatial species–richness gradients across scales: a meta-analysis. *Journal of Biogeography*, **36**, 132–147.
- Filipe, A.F., Araújo, M.B., Doadrio, I., Angermeier, P.L. & Collares-Pereira, M.J. (2009) Biogeography of Iberian freshwater fishes revisited: the roles of historical versus contemporary constraints. *Journal of Biogeography*, **36**, 2096–2110.
- Froese, R. & Pauly, D. (2010) FishBase. <http://www.fishbase.org>.
- Gillespie, R.G. & Roderick, G.K. (2002) Arthropods on islands: colonization, speciation, and conservation. *Annual Review of Entomology*, **47**, 595–632.
- Goldberg, E.E., Roy, K., Lande, R. & Jablonski, D. (2005) Diversity, endemism, and age distributions in macroevolutionary sources and sinks. *The American Naturalist*, **165**, 623–633.
- Grenyer, R., Orme, C.D.L., Jackson, S.F., Thomas, G.H., Davies, R.G., Davies, T.J., Jones, K.E., Olson, V.A., Ridgely, R.S., Rasmussen, P.C., Ding, T.S., Bennett, P.M., Blackburn, T.M., Gaston, K.J., Gittleman, J.L. & Owens, I.P.F. (2006) Global distribution and conservation of rare and threatened vertebrates. *Nature*, **444**, 93–96.
- Grueber, C.E., Nakagawa, S., Laws, R.J. & Jamieson, I.G. (2011) Multimodel inference in ecology and evolution: challenges and solutions. *Journal of Evolutionary Biology*, **24**, 699–711.
- Heaney, L.R. (2000) A long-term, large-scale perspective on the equilibrium model of island biogeography. *Global Ecology and Biogeography*, **9**, 59–74.
- Hugueny, B., Oberdorff, T. & Tedesco, P.A. (2010) Community ecology of river fishes: a large-scale perspective. *Community ecology of stream fishes: concepts, approaches and techniques* (ed. by K. Gido and D. Jackson), pp. 29–62. American Fisheries Society, Bethesda, MD.
- Hugueny, B., Movellan, A. & Belliard, J. (2011) Habitat fragmentation and extinction rates within freshwater fish communities: a faunal relaxation approach. *Global Ecology and Biogeography*, **20**, 449–463.
- Jablonski, D., Roy, K. & Valentine, J.W. (2006) Out of the tropics: evolutionary dynamics of the latitudinal diversity gradient. *Science*, **314**, 102–106.
- Jansson, R. (2003) Global patterns in endemism explained by past climatic change. *Proceedings of the Royal Society B: Biological Sciences*, **270**, 583–590.
- Jansson, R. & Davies, T.J. (2008) Global variation in diversification rates of flowering plants: energy vs. climate change. *Ecology Letters*, **11**, 173–183.
- Jenkins, D.G., Brescacin, C.R., Duxbury, C.V., Elliott, J.A., Evans, J.A., Grablow, K.R., Hillegass, M., Lyon, B.N., Metzger, G.A., Olandese, M.L., Pepe, D., Silvers, G.A., Suresch, H.N., Thompson, T.N., Trexler, C.M., Williams, G.E., Williams, N.C. & Williams, S.E. (2007) Does size matter for dispersal distance? *Global Ecology and Biogeography*, **16**, 415–425.
- Jetz, W. & Rahbek, C. (2002) Geographic range size and determinants of avian species richness. *Science*, **297**, 1548–1551.
- Jetz, W., Rahbek, C. & Colwell, R.K. (2004) The coincidence of rarity and richness and the potential signature of history in centres of endemism. *Ecology Letters*, **7**, 1180–1191.
- Kisel, Y. & Barraclough, T.G. (2010) Speciation has a spatial scale that depends on levels of gene flow. *The American Naturalist*, **175**, 316–334.
- Knouft, J.H. & Page, L.M. (2011) Assessment of the relationships of geographic variation in species richness to climate and landscape variables within and among lineages of North American freshwater fishes. *Journal of Biogeography*, **38**, 2259–2269.
- Lambeck, K., Yokoyama, Y. & Purcell, T. (2002) Into and out of the Last Glacial Maximum: sea-level change during Oxygen Isotope Stages 3 and 2. *Quaternary Science Reviews*, **21**, 343–360.
- Leprieur, F., Tedesco, P.A., Hugueny, B., Beauchard, O., Dürr, H.H., Brosse, S. & Oberdorff, T. (2011) Partitioning global patterns of freshwater fish beta diversity reveals contrasting signatures of past climate changes. *Ecology Letters*, **14**, 325–334.
- Lévêque, C., Oberdorff, T., Paugy, D., Stiassny, M.L.J. & Tedesco, P.A. (2008) Global diversity of fish (Pisces) in freshwater. *Hydrobiologia*, **595**, 545–567.
- Lomolino, M.V., Riddle, B.R. & Brown, J.H. (2006) *Biogeography*, 3rd edn. Sinauer Associates, Sunderland, MA.
- Losos, J.B. & Ricklefs, R.E. (2009) Adaptation and diversification on islands. *Nature*, **457**, 830–836.
- Losos, J.B. & Schluter, D. (2000) Analysis of an evolutionary species–area relationship. *Nature*, **408**, 847–850.
- Lovejoy, N.R., Albert, J.S. & Crampton, W.G.R. (2006) Miocene marine incursions and marine/freshwater transitions: evidence from Neotropical fishes. *Journal of South American Earth Sciences*, **21**, 5–13.
- MacArthur, R.H. & Wilson, E.O. (1967) *The theory of island biogeography*. Princeton University Press, Princeton, NJ.
- Mace, G.M., Gittleman, J.L. & Purvis, A. (2003) Preserving the tree of life. *Science*, **300**, 1707–1709.
- Mansion, G., Selvi, F., Guggisberg, A. & Conti, E. (2009) Origin of Mediterranean insular endemics in the Boraginales: integrative evidence from molecular dating and ancestral area reconstruction. *Journal of Biogeography*, **36**, 1282–1296.
- Martin, A.P. & Palumbi, S.R. (1993) Body size, metabolic-rate, generation time, and the molecular clock. *Proceedings of the National Academy of Sciences USA*, **90**, 4087–4091.
- Melo, A.S., Rangel, T. & Diniz, J.A.F. (2009) Environmental drivers of beta-diversity patterns in New World birds and mammals. *Ecography*, **32**, 226–236.
- Mittelbach, G.G., Schemske, D.W., Cornell, H.V. *et al.* (2007) Evolution and the latitudinal diversity gradient: speciation, extinction and biogeography. *Ecology Letters*, **10**, 315–331.

- Oberdorff, T., Guégan, J.F. & Huguény, B. (1995) Global scale patterns of fish species richness in rivers. *Ecography*, **18**, 345–352.
- Oberdorff, T., Lek, S. & Guégan, J.F. (1999) Patterns of endemism in riverine fish of the Northern Hemisphere. *Ecology Letters*, **2**, 75–81.
- Orme, C.D.L., Davies, R.G., Burgess, M., Eigenbrod, F., Pickup, N., Olson, V.A., Webster, A.J., Ding, T.S., Rasmussen, P.C., Ridgely, R.S., Stattersfield, A.J., Bennett, P.M., Blackburn, T.M., Gaston, K.J. & Owens, I.P.F. (2005) Global hotspots of species richness are not congruent with endemism or threat. *Nature*, **436**, 1016–1019.
- Pimm, S.L., Russell, G.J., Gittleman, J.L. & Brooks, T.M. (1995) The future of biodiversity. *Science*, **269**, 347–350.
- Reyjol, Y., Huguény, B., Pont, D., Bianco, P.G., Beier, U., Caiola, N., Casals, F., Cowx, I., Economou, A., Ferreira, T., Haidvogel, G., Noble, R., De Sostoa, A., Vigneron, T. & Virbickas, T. (2007) Patterns in species richness and endemism of European freshwater fish. *Global Ecology and Biogeography*, **16**, 65–75.
- Rosenzweig, M.L. (1995) *Species diversity in space and time*. Cambridge University Press, Cambridge.
- Rosindell, J. & Phillimore, A.B. (2011) A unified model of island biogeography sheds light on the zone of radiation. *Ecology Letters*, **14**, 552–560.
- Sokal, R.R. & Oden, N.L. (1978) Spatial autocorrelation in biology. 1. Methodology. *Biological Journal of the Linnean Society*, **10**, 199–228.
- Stehli, F.G., Douglas, R.G. & Newell, N.D. (1969) Generation and maintenance of gradients in taxonomic diversity. *Science*, **164**, 947–949.
- Swartz, E.R., Skelton, P.H. & Bloomer, P. (2007) Sea-level changes, river capture and the evolution of populations of the Eastern Cape and fiery redbins (*Pseudobarbus afer* and *Pseudobarbus phlegethon*, Cyprinidae) across multiple river systems in South Africa. *Journal of Biogeography*, **34**, 2086–2099.
- Tales, E., Keith, P. & Oberdorff, T. (2004) Density–range size relationships in French riverine fishes. *Oecologia*, **138**, 360–370.
- Vörösmarty, C.J., Fekete, B.M., Meybeck, M. & Lammers, R.B. (2000) Geomorphometric attributes of the global system of rivers at 30-minute spatial resolution. *Journal of Hydrology*, **237**, 17–39.
- Waters, J.M. & Wallis, G. (2001) Cladogenesis and loss of the marine life-history phase in freshwater galaxiid fishes (Osmeriformes: Galaxiidae). *Evolution*, **55**, 587–597.

## SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article:

**Appendix S1** Additional methodological and data information.  
**Figure S1** Global map distinguishing river drainage basins by their index of *in situ* speciation.

**Figure S2** Global map of the endemism levels resulting from the null model approach.

**Table S1** List of the environmental variables and their respective references.

**Table S2** Results of the model averaging and generalized linear models from the null model approach.

As a service to our authors and readers, this journal provides supporting information supplied by the authors. Such materials are peer-reviewed and may be re-organized for online delivery, but are not copy-edited or typeset. Technical support issues arising from supporting information (other than missing files) should be addressed to the authors.

## BIOSKETCH

**Pablo A. Tedesco** is a scientist at the IRD, the French institute for overseas research. His research topics include large spatial and temporal scale ecology and biogeography, with a focus on tropical and temperate freshwater fish communities.

Editor: Janne Soininen