

Article type: Original Article

Explaining global-scale diversification patterns in actinopterygian fish

Pablo A. Tedesco^{1*}, Emmanuel Paradis², Christian Lévéque³ & Bernard Hugueny³

1. UMR5174 EDB (Laboratoire Evolution et Diversité Biologique), CNRS, IRD, UPS, ENFA, Université Paul Sabatier Toulouse 3, 118 route de Narbonne, F-31062, Toulouse, France. pablo.tedesco@ird.fr; Phone : +33 5 61 55 67 47 ; Fax : +33 5 61 55 73 27
2. Institut des Sciences de l'Évolution, Université Montpellier - CNRS - IRD - EPHE, Place Eugène Bataillon – C.C. 065 - 34095 Montpellier cedex 05 – France.
emmanuel.paradis@ird.fr
3. UMR7208 BOREA (Biologie des Organismes et des Ecosystèmes Aquatiques), MNHN, IRD 207, CNRS, UPMC, UCN, UA, Muséum National d'Histoire Naturelle, 43 rue Cuvier - CP 26, 75005 Paris, France. cleveque@mnhn.fr; hugueny@mnhn.fr

* Corresponding author: Pablo A. Tedesco

Short running title: Diversification drivers in fish

Number of words in main text: 6973

Required journal pages for figures and tables: Figure 1 (one page); Figure 2 (0.5 page); Table 1 (0.4 page)

Abstract

Aim Factors that isolate populations and reduce gene flow are considered key drivers of speciation and possibly diversification. Here we analyze the diversification rates of nearly 80% of the actinopterygian fish families in relation to biological traits and habitat factors associated with isolation and fragmentation levels.

Location Global

Methods Net diversification rate for each family was estimated using the method-of-moments estimator for stem-group ages. Phylogenetic generalized least-squares analysis (PGLS), controlling for the non-independence between clades due to phylogeny, was applied with diversification rate as the response variable to test the effects of mean body size, proportions of strictly freshwater, reef-associated and migratory species and including the median latitudinal distribution and range of each family.

Results After accounting for the phylogenetic relatedness of families and for their latitudinal distribution, we found strong support in agreement with our isolation and fragmentation hypotheses: predominance of freshwater dependence, reef-association, small body size or non-migratory behavior in families is related to more rapid rates of diversification. We also found a highly significant and positive effect of latitudinal range and no clear effect of median latitude.

Main conclusions This analysis suggests that factors related to the physical fragmentation of habitats and to lower dispersal ability of species have played an important role in the diversification processes of the most diverse group of vertebrates.

Keywords: actinopterygian fish, biogeography, biological traits, diversification, fragmentation, speciation

INTRODUCTION

Explaining why some clades and regions have more species than others is one of the great challenges of evolutionary ecology, illustrated for instance by the increasing number of analyses of diversification on major groups of vertebrates (e.g. Owens *et al.*, 1999; Phillimore *et al.*, 2006; Weir & Schlüter, 2007; Vega & Wiens, 2012; Rabosky *et al.*, 2013; Rolland *et al.*, 2014). Habitat fragmentation and dispersal capacities are considered key drivers of speciation and extinction (Kisel *et al.*, 2011), the two processes ultimately responsible for differences in diversity and diversification rates between clades and regions. By limiting gene flow, the fragmentation of populations through geographical isolation is supposed to increase speciation rates. Similarly, greater species dispersal abilities should reduce isolation and speciation rates but, at the same time, increase the resilience of populations to disturbances, hence also reducing extinction rates (Riginos *et al.*, 2014). Here we explore the role of fragmentation and dispersal-related traits in driving global patterns of fish diversification rates. We analyzed the diversification rates of nearly 80% of the actinopterygian fish families accounting for phylogenetic relatedness, latitudinal distribution and several ecological and biological traits related to fragmentation and dispersal capacities supposed to influence diversification processes.

Freshwater dependence. Comparisons between land and sea have shown that life on land (including freshwater organisms) is more diverse, with ca. 86% of currently described species (Mora *et al.*, 2011), while covering only 29% of Earth's surface (Vermeij & Grosberg, 2010; Mora *et al.*, 2011). Even if the striking disparities observed between broad types of environments (i.e. terrestrial, land, aquatic, freshwater and marine) have intrigued biogeographers and ecologists for decades (Dawson & Hamner, 2008), few marine–terrestrial (or freshwater) comparative studies have been performed (Webb, 2012), and even fewer have quantitatively tested possible evolutionary and ecological causes (Dawson, 2012; Vega &

Wiens, 2012; Bloom *et al.*, 2013; Wiens, 2015a,b). The diversity-area disparity strongly increases when considering only freshwater habitats. With ca. 126 000 described animal species inhabiting freshwaters (Balian *et al.*, 2008), they account for over 10% of all animals described to date (Mora *et al.*, 2011; Wiens, 2015b) while occupying only 0.8% of the Earth's surface and 0.02% of available aquatic habitable volume (Dawson, 2012). Among aquatic organisms, fish are a good example of this 'freshwater paradox', harboring ca. 40% in freshwaters, while the remaining 60% of fish diversity inhabit marine habitats comprising >99% of available aquatic habitat (Lévéque *et al.*, 2008).

Beside the paramount difference in area or volume between marine and freshwater environments, these two habitats fundamentally differ in their degree of fragmentation (Vermeij & Grosberg, 2010; Wiens, 2015b). Marine-scape connectivity is manifested in three dimensions, as animals have several alternative paths to move from one place to another. Instead, freshwaters are usually structured as dendritic networks with a hierarchical branching finally flowing to the sea, making river drainage basins highly fragmented island-like systems (Hugueny *et al.*, 2010). These high levels of fragmentation, within and between drainage basins, are a central factor shaping evolutionary dynamics in freshwaters (e.g. Burridge *et al.*, 2008; Tedesco *et al.*, 2012; Dias *et al.*, 2013). Comparatively, marine organisms have less effective barriers to dispersal and higher levels of gene flow (e.g. Palumbi, 1994) reducing the probability of speciation events. The proportion of strictly freshwater fish species within clades should hence be positively related to diversification rates.

Reef-association. Understanding the historical forces shaping fish diversity associated to coral reefs is a long-standing question in marine biogeography. Does reef use increase the rate of diversification in fishes? A recent analysis by Cowman and Bellwood (2011) found a significant positive correlation between reef association and rates of diversification among four fish families, reflecting similar associations previously reported in the Tetraodontiformes

(Alfaro *et al.*, 2007). Coral reef ecosystems host approximately one-third of all marine fish species, although covering less than 0.1% of the ocean's surface (Rocha & Bowen, 2008). Similarly to freshwaters, coral reefs are highly fragmented habitats, where most species consist of patchily distributed populations, although connected through pelagic larval dispersal (Floeter *et al.*, 2008). A recent review shows that allopatry and parapatry are the primary modes of speciation explaining high diversity levels found in coral reef fish (Rocha & Bowen, 2008). Hard vicariant barriers, such as closures of seaways (e.g., Isthmus of Panama, Tethys) and large geographic scales (ocean-wide) are involved in allopatric speciation processes, while increasing empirical and theoretical evidence shows that parapatric speciation is a common (and probably the prevalent) mode of diversification in coral-reef fishes. This last mechanism involves speciation with limited gene flow and relates to the weak and intermittent biogeographical barriers common in oceans (e.g., sea level and climatic fluctuations). A positive relationship is then expected between diversification rates and the proportion of reef-associated species within families.

*Dispersal related-trait*s. A variety of life history traits have been discussed as potential drivers of speciation (Cardillo *et al.*, 2003; Isaac *et al.*, 2005; Phillimore *et al.*, 2006). For instance, the dominance of small-bodied species, a widely observed macroecological pattern in body size distributions, implies that small bodied organisms have experienced elevated net rates of diversification. However, despite the intuitive nature of this hypothesis, weak general support for evolutionary trends towards increased cladogenesis in smaller bodied clades has been found (e.g. Cardillo *et al.*, 2003; Isaac *et al.*, 2005). Species with low dispersal ability should experience greater isolation and lower gene flow, and thus a greater potential for local adaptation and higher rates of speciation (e.g. Riginos *et al.*, 2014). Dispersal distance has been positively related to body size in active dispersers (including fish; e.g. Radinger & Wolter, 2014) and should play a determinant role in speciation/extinction

processes shaping present diversity patterns. For both marine and freshwater environments, a negative relationship is then expected between diversification rates and body size.

Unlike body size, the role of migratory behavior on diversification has rarely been assessed, except on birds (e.g. Rolland *et al.*, 2014). Migratory behavior may either enhance or reduce opportunities for speciation, as migratory movements increase the probability of colonizing new areas leading to divergence from ancestral populations, but also increasing gene flow between populations thereby reducing genetic divergence. A recent global analysis of bird diversification (Rolland *et al.*, 2014) suggests that migratory species often diversify by generating a sedentary daughter species in addition to the ancestral migratory one, and that speciation with no character change is overall more frequent in sedentary than in migratory species. Like for birds, fish migratory behaviors (i.e., diadromous species migrating between the sea and freshwater, potamodromous species migrating within freshwaters and oceanodromous species migrating within oceans) should enhance gene flow between populations, overall reducing the probabilities for allopatric speciation to occur. A negative relationship is then expected between diversification rates and the proportion of migrating species within families. However, because migratory species may diversify by generating sedentary descendants (Rolland *et al.*, 2014), the relationship may be hump-shaped with larger diversification for families with an intermediate proportion of migratory species.

To address these hypotheses, we used a multiple regression approach to test the simultaneous effect of several factors on the net diversification rate accounting for the phylogenetic relatedness of fish families and their latitudinal distribution (median and range). We ask whether the isolation-related factors identified above have acted on fish diversification rate as hypothesized and whether a strong and coherent effect of isolation emerges from the data.

MATERIAL AND METHODS

Fish diversity, traits and latitudinal distribution

Fishbase (Froese & Pauly, 2013) provided information concerning the numbers of species within families, their biological traits and distribution over freshwater or reef habitats. We used a list of 31,252 currently valid actinopterygian fish species with information on their occurrence in freshwater or saltwater environments, reef-association, maximum adult body size, and migratory behavior (either anadromous, diadromous, catadromous, potamodromous, amphidromous or oceanodromous) to compute the proportion of strictly freshwater species, proportion of reef-associated species, mean body size, and proportion of migratory species for each family (see Fig. 1 and Appendix S1 in Supporting Information). While information on body size is available for ~87% of all fish species, information on migratory behavior is only available for ~12%. However, information on migratory behavior targets migrating species, while non-informed species can be considered, in most cases, as non-migrating. We used maximum body length data as a measure of body size based on total, standard and fork length measurements. Some variation in our data could be created if the proportions of these measurement types vary between families. However, we assume this variation to be small compared to the difference of more than four orders of magnitude in body size among the entire species pool.

The latitudinal distributions of species were taken from a global occurrence dataset of freshwater fish species (Tedesco *et al.*, 2012; Dias *et al.*, 2014) and from the Ocean Biogeographic Information System (OBIS, 2014) for marine species. Two values by family were computed from these datasets: the median latitude and the latitudinal distribution range. We included this latitudinal information to account for potential differences in tropical *vs.*

temperate diversity patterns and diversification processes. Higher speciation and lower extinction rates have been related to tropical climates, which are traditionally seen as centers of diversification (e.g. Cardillo *et al.*, 2005; Mittelbach *et al.*, 2007).

Diversification rates

We estimated the net diversification rate for each family using the method-of-moments estimator for stem-group ages (Magallón & Sanderson, 2001). We focused on stem-group ages because phylogenies most often include too few species to allow confident estimation of crown-group ages and crown divergence times cannot be obtained for monospecific clades. The definition of crown ages for fish families is not as straightforward as for well-studied groups (e.g. mammals or birds). The relationships within the vast majority of fish families are poorly known, so the definition of crown ages would be more uncertain than for stem ages. More importantly, the stem group diversification rate better represents the overall net diversification of a clade because it incorporates the entire history of the group, whereas the crown age might represent only a very recent diversification. The method-of-moments estimator requires both clade age and species richness as input, and an assumed relative extinction rate (e). We followed the methodology applied by Vega & Wiens (2012) and (Wiens, 2015b) using three different measures for the relative rates of speciation and extinction, including low (0), high (0.90) and intermediate (0.50) values to address the robustness of the results to different values of epsilon.

Family ages were estimated from dated molecular phylogenies and the fossil record. A systematic literature search was done for each clade, providing origination time estimates for 460 actinopterygian fish families (143 references; see Appendix S2). Families with age estimates uniquely based on one or more fossil records were excluded from the analyses.

When both fossil and phylogenetic information were available for a given family, phylogeny-based ages were preferred if the estimations predated the oldest fossil record age. In other words, a fossil age was used as an estimation of the origination age of a family only when stem phylogenetic information was also available, but when these phylogenetic-derived ages were younger than the oldest fossil-derived age available. In these cases the fossil age is older than the molecular estimated stem-group age, then the fossil age is presumably correct and the molecular age is not. Only 14 families were in that case, and using fossil information this way prevents an overestimation of diversification rates compared to the phylogenetic-derived rates available, making use of all available information for a given family (also note that excluding these 14 families from the analyses did not change our results; compare Table 1 and Table S3.1). For phylogenetic-derived age families, a mean value was computed when more than one phylogeny-based age was available. Families considered as non-monophyletic were also excluded based on information provided by Rabosky *et al.* (2013). These exclusions finally reduced the dataset to 377 families (see Fig. 1 and Appendix S1). To account for the uncertainty in origination age estimation, we included weights in our analyses. These weights measured the level of knowledge that we have on the origination ages, and were computed as the number of age estimates available to calculate the mean age of each family (Fig. 1). We further checked that these mean ages assigned to families were reasonably accurate by comparing the ages of sister pairs based on the phylogeny provided by Rabosky *et al.* (2013). Although this phylogenetic tree does not include all extant ray-finned fish families and our origination ages arise from many different sources, this sister pairs comparison produced a high correlation ($r = 0.88$; see Fig. S3.1).

Some authors have criticized the method-of-moments estimator because it makes the strong assumption that speciation rates have been constant through time (see Kozak & Wiens, 2016 for support arguments). Rabosky (2010) suggested that evolutionary lineages follow a

general pattern of time-dependent diversification where speciation rate is high during the early stage of a clade and then declines through time to stabilize to a value close to the extinction rate so that species richness is in an "equilibrium" state (but see Harmon & Harrison, 2015 for an opposite view). Non-constant diversification within families, as suggested by Rabosky, may create an apparent pattern of heterogeneous diversification rates among families if the ages of the families vary substantially (which is the case with our data). Under this condition, it may be possible to observe a relationship between diversification rate and a trait if this trait varies with respect to family age (e.g., a complex trait that can evolve in the oldest families, or a trait that regresses through time). To assess this potential bias we tested the relationship between the traits studied here and age among fish families (see 'Statistical analyses' section below).

Statistical analyses

The relationships between diversification rates, age, richness and our biological isolation and habitat fragmentation variables were tested using phylogenetic generalized least-squares analysis (PGLS). The PGLS approach fits a linear model controlling for the non-independence between clades due to phylogeny. We used the phylogenetic tree of actinopterygian fish (7822 species representing 420 families) provided by (Rabosky *et al.*, 2013). A family-level tree for these analyses was obtained by pruning the 7822-species tree so that each family was represented by a single terminal branch. PGLS analyses were implemented using R (R Development Core Team, 2015) with the package 'nlme' (Pinheiro *et al.*, 2016) and its 'gls' function, with the maximum-likelihood transformation of branch length optimized for the data ('method = LM'), and applying weights to families accounting for uncertainty in origination age estimations (see the 'Diversification rates' section).

Estimated values of Pagel's λ ($P\text{-}\lambda$) were used and values of κ and δ were fixed at 1. $P\text{-}\lambda$ values were used to evaluate the degree to which evolutionary relatedness of families affected ecological and biological similarity (i.e. the phylogenetic signal). A $P\text{-}\lambda$ value of 0 indicates the absence of phylogenetic signal (i.e., trait values are random with respect to phylogeny), while a value of 1 indicates a phylogenetic signal consistent with a Brownian motion evolutionary model (i.e., closely related families have more similar trait values than would be expected by chance).

PGLS models were applied with diversification rate as the response variable to test the effects of mean body size, proportions of strictly freshwater, reef-associated and migratory species and including the median latitudinal distribution and range of each family. To determine the relative importance and significance of these variables to explain diversification rates (for each assumed value $e = 0, 0.5$ and 0.9), we ran models for all possible combinations of the explanatory variables and then performed model averaging based on the 'Akaike information criterion' (AICc). As a cut-off criterion to delineate a 'top model set' providing average parameter estimates and confidence intervals, we used fitted models with $\Delta\text{AIC} < 4$ (Grueber et al., 2011). The model selection and averaging was implemented using R with the package 'MuMIn' (Bartoń, 2015) and its '*dredge*' function. Because a large proportion of monospecific families is present in our dataset ($n = 36$), we explored their influence in our results by repeating our analyses after excluding them. We checked for multicollinearity among factors using the Variance Inflation Factor (VIF) procedure, revealing no strong collinearity among predictors (maximum VIF in all models < 1.92). Log-transformation was applied to body size, and diversification rates were squared-root transformed since the latter vary between 0 and 1. In parallel to the analysis of diversification rates, we used PGLS models to evaluate the individual (single term models) and combined (complete multiple model) effects of each tested variable against family age to exclude any spurious influence of

age on the relationships between the evaluated traits and diversification rates (see ‘Diversification rates’ section).

Finally, we applied an alternative analytical method to our data, MacroCAIC (Agapow & Isaac, 2002), to assess whether our results were consistent when analyzed with a comparative analysis using independent contrasts on species richness data, avoiding all potential problems that may arise from age estimation (see ‘Diversification rates’ section). MacroCAIC requires summing of clade richness at internal nodes of a phylogeny, and computes a relative rate difference in diversification for all bifurcating nodes, which is given by $\ln(N_1/N_2)$ where the values N_1 and N_2 are the species richness of the two daughter nodes and N_1 is the species richness of the clade with the larger value. To determine the significance of the tested variables to explain diversification rates, we used the ‘*macrocaic*’ function from the R package ‘caper’ (*Orme et al.*, 2013), sequentially excluding non-significant variables from the model. We used the same phylogeny as for the PGLS analyses (using time-based or equal branch lengths did not have substantial effect on the MacroCAIC and PGLS results).

RESULTS

Phylogenetic signal

Diversification rates estimated for actinopterygian fish families under the three levels of e showed intermediate values of phylogenetic signal with confidence intervals always excluding the extreme values of 0 and 1. Whether monospecific families were considered or not, only very slight variations in lambda values were observed (Table S3.[+2](#)). When we addressed the phylogenetic signal associated with the individual explanatory variables used in the models, we found different phylogenetic signals for different variables. Freshwater and reef-associated proportions exhibited the largest lambda values, not significantly different

from 1 for freshwater proportion when excluding monospecific families, indicating that phylogenetic dependence is strong for these traits (Fig. 1; Table S3.[+2](#)). Intermediate phylogenetic dependence was observed for mean body size, median latitude and latitudinal range, while migratory proportion exhibited lower lambda values, not significantly different from 0 when excluding monospecific families (Fig. 1; Table S3.[+2](#)). When we subsequently tested for the phylogenetic signal of the covariance between diversification rate and the explanatory variables, we found intermediate values in every case, and within each variable, lambda values were always inversely related to epsilon values (Table S3.[+2](#)). Overall, these findings suggest that the PGLS approach for exploring correlates of diversification rate accounting for the phylogenetic relatedness between families is more appropriate than assuming lambda is equal to either 0 (no phylogenetic correction) or 1 (a model based on Brownian trait evolution).

Analysis of diversification rates

Multiple (complete) PGLS models of diversification rates across nearly 80% of the actinopterygian fish families explained between 29% and 44% of the total variance (pseudo- R^2) depending on whether monospecific families were considered or not and depending on the assumed epsilon value, $\epsilon = 0.9$ having the best fit in both cases (lowest AIC, see Table 1). Overall, the best PGLS models retained by the AICc selection procedure showed significant relationships between diversification rates and the tested variables. These relationships are all in agreement with our isolation and fragmentation hypotheses explaining differences in diversification rates: positive for the proportion of strictly freshwater species and the proportion of reef-associated species, and negative for body size and the proportion of migratory species (Table 1, Fig. 2). In all cases (i.e. the three epsilon values and accounting or

not for monospecific families), the proportion of strictly freshwater species, proportion of migratory species and the mean body size showed maximal relative importance values (i.e. the sum of the Akaike weights over all of the models in which the variables appear) and highly significant coefficient values (Table 1). The proportion of reef-associated species showed significant coefficient values and relative importance values of 1 only when excluding monospecific families (Table 1). The best PGLS models also retained the latitudinal range of families, showing maximal relative importance values and highly significant and positive coefficient values under all conditions (Table 1). However, median latitude had the smallest effects on diversification rates overall, showing significant and positive relationships in only two cases, $e = 0$ and 0.5 when excluding monospecific families (Table 1). The alternative comparative analysis, using MacroCAIC, confirmed these findings showing significant effects of all the tested variables excepting median latitude and reef association (see Table S3.[32](#)).

Age-trait relationships

The PGLS models evaluating the individual (single term models) and combined (full model) effects of each tested variable against family age showed a highly significant negative relationship with median latitude (Table S3.[34](#); Fig. S3.2), including or not monospecific families. All other explanatory variables showed non-significant relationships with family age, either using single term or complete models, and excluding or not monospecific families (Table S3.[34](#)). Except median latitude, these lack of relationship between traits and age overall supports our findings on the patterns of inter-family diversification, excluding potential biases related to spurious influence of age on traits.

DISCUSSION

Several studies have analyzed patterns of diversification in actinopterygian fish (e.g. Near *et al.*, 2012; Vega & Wiens, 2012; Rabosky *et al.*, 2013). Our study is however the first to highlight the importance of fragmentation and isolation attributes of species on the patterns of diversification that have shaped present diversity differences in actinopterygian fish, the most diverse group of vertebrates. Overall, our results suggest that factors related to the physical fragmentation of habitats and biological traits related to isolation have played an important role in the diversification processes of this group. After accounting for the phylogenetic relatedness and latitudinal distribution of clades, we show that the highly fragmented freshwater and coral-reef environments have promoted higher levels of diversification in fish, and that biological features related to lower dispersal ability have also enhanced diversification.

The Freshwater Fish Paradox

Freshwater habitats house a disproportional high part of the global fish diversity considering the small proportion of the earth's surface they occupy, i.e. the 'freshwater fish paradox'. Several hypotheses have been invoked to explain the higher terrestrial diversity compared to marine diversity (Vermeij & Grosberg, 2010; Wiens, 2015b). Non-marine clades would have diversified more because of higher net primary productivity, larger primary producers, greater habitat complexity, narrower ecological specialization, more effective barriers to dispersal and/or smaller geographical range sizes. However, only the two latter hypotheses, which are both related to fragmentation and isolation, can account for the 'freshwater fish paradox', the others being ultimately derived from the physical contrasts between air and water as a medium for life on land and in the sea (Vermeij & Grosberg, 2010;

Vega & Wiens, 2012). Bloom *et al.* (2013) recently observed higher speciation and net diversification rates in freshwater compared to marine lineages of the New World silverside fish clade Menidiinae. Our results confirm and extend these findings to nearly all other actinopterygian fish families, giving further support to the idea that greater number of barriers in freshwater habitats relative to marine habitats likely results in more frequent allopatric speciation events. Vega & Wiens (2012) also suggested that more effective barriers to dispersal should be responsible for the similar diversity levels found in both environments (relative to their respective areas). However, contrary to our results, these authors found no significant relationship between the net diversification rates and the proportion of saltwater species in 97 families and 22 higher clades of actinopterygian fish. Our results show higher diversification rates in freshwater fish families suggesting that the physical fragmentation of freshwater habitats is at least partially responsible for the differences in diversification rates and is the main mechanism behind the ‘freshwater fish paradox’. Since our age estimations include and are similar to the information provided by Vega & Wiens’s (2012), two factors can explain this inconsistency: the number of clades analyzed and the way freshwater species were considered. Indeed, we included nearly 80% (against 20% for Vega & Wiens) of the actinopterygian fish families in our analysis, which could greatly modify the final relationship between habitat and net diversification rate. Besides, to estimate the pre-eminence of freshwater habitat in a given family, we computed the proportions of strictly freshwater species (i.e., species occurring in freshwaters but absent from salt or brackish waters), while Vega & Wiens (2012) included all species entering freshwaters which may have weakened the signal of habitat dependency in their study. Concerning this last point, we rerun our analyses with the proportion of freshwater fish within families computed as in Vega & Wiens (2012) and found a weaker, although still significant, effect of freshwater dependency on diversification rates (see Table S3.5).

Coral reef-associated diversification in fish

Our results confirm previous comparative studies on Tetraodontiformes and other fish families that have shown that lineages occupying reefs diversify faster than non-reef fishes (Alfaro *et al.*, 2007; Cowman & Bellwood, 2011) and extend this link to almost all actinopterygian fish families as a general pattern. The pattern of higher rates of diversification for reef-associated families that we have detected here does appear to be obscured when including monospecific families. This result is mainly driven by five monospecific clades whose species inhabit reef habitats: Enoplosidae, Triodontidae, Zanclidae, Menidae and Rachycentridae, although this last clade is known to occur in a large variety of habitats (Froese & Pauly, 2013). The first four clades have known fossil congeners suggesting that speciation events have been balanced by extinctions in these families, resulting in an apparent diversification rate of zero. Furthermore, under higher relative extinction rates (i.e. $e = 0.9$ when estimating diversification rates) our results show a nearly significant effect of reef-association event when including monospecific families.

Diversification in coral reef-associated families is driven in part by ecological opportunities provided by the unique and complex reef habitat itself. Indeed, recent analyses have related diversification patterns in coral reef fish clades to functional aspects and ecological novelty (e.g. Price *et al.*, 2011). However, major paleoclimatic events over the geological times of reef formation and evolution are also likely to have increased diversification rates in reef clades by fragmenting reef habitats and their populations (e.g. Alfaro *et al.*, 2007). For instance, empirical evidence suggests that fluctuations in the effectiveness of three physical ‘soft’ barriers provide a mechanism for much of the recent diversification of reef fishes in the Atlantic (Floeter *et al.*, 2008). Just as in freshwater

systems, assembled and then fragmented by changing sea-levels and river captures, the fluctuating permeability of oceanic barriers to dispersal have promoted a ‘dispersal-isolation’ model of diversification in coral reefs (Cowman & Bellwood, 2013).

Dispersal-related biological factors

Our results showed the positive influence of two biological traits on diversification rates of fish families: small body size and non-migratory behavior. These traits are supposed to act negatively on dispersal capabilities, causing smaller and more strongly fragmented ranges, which in turn should facilitate reproductive isolation and thus favor speciation. However, weak general support has been found for evolutionary trends towards increased cladogenesis in small bodied species (e.g. Cardillo *et al.*, 2003; Isaac *et al.*, 2005). Our findings, along with positive evidence found for some groups (e.g. Gittleman & Purvis, 1998; Gardezi & da Silva, 1999; Wollenberg *et al.*, 2011), clearly support the hypothesis linking body size to net diversification rate. These opposite findings may be explained by methodological differences in comparing diversification (e.g., comparing sister clades or using clade ages) and differences in the taxonomic resolution at which diversification is observed (i.e., from genera to phyla). Furthermore, body size *per se* is unlikely to be directly related to evolutionary rates. Rather, body size can be correlated, sometimes strongly, with other traits more directly related to the mechanisms involved in evolutionary rates, dispersal capabilities being one of them, and these relationships may vary between taxa. Small bodied species should have elevated rates of diversification for several reasons. Greater rates of molecular evolution, metabolic rates, intrinsic rates of population increase, effective population sizes and shorter generation times (and its correlates, life span and age at first reproduction) have been related to body size (e.g. Martin & Palumbi, 1993; Gillooly *et al.*, 2005), suggesting that smaller organisms evolve

faster. However, all these biological traits related to body size may interact differently within different taxa, producing divergent body size-diversification rate patterns.

High annual dispersal has been identified as a significant predictor of high rates of diversification for birds (Phillimore *et al.*, 2006) and recent findings suggest that the evolution of seasonal migration in birds has facilitated diversification through the divergence of migratory subpopulations that become sedentary (Rolland *et al.*, 2014). Our results suggest an opposite mechanism acting on fish at the family level, where migratory dispersal negatively affects diversification rates by reducing opportunities for speciation. Although we regrouped five different migratory behaviors (anadromous, diadromous, catadromous, potamodromous, amphidromous and oceanodromous) to create one single variable, we also found a negative relationship with diversification rates when analyzing separately these migratory behaviors (results not shown).

Latitudinal distribution and range

A number of evolutionary hypotheses explaining the latitudinal diversity gradient assume that net diversification rates are higher in the tropics either because of increased speciation rates (i.e., tropics as a ‘cradle’) or decreased extinction rates (i.e., tropics as a ‘museum’) (Mittelbach *et al.*, 2007). Our results show a positive, although not significant, effect of median latitudinal distribution on diversification rates of actinopterygian fish at the family level, suggesting that tropical clades did not diversify faster. The significant negative relationship found between origination time and median latitude suggests that families centered in tropical waters are older than temperate ones. Together, these findings give support to the ‘museum’ hypothesis, with new species rising at similar rates at different latitudes, but tropical latitudes accumulating them for longer geological periods. Although a

number of paleontological studies support the hypothesis that net diversification rates are higher at lower latitudes (reviewed in Mittelbach *et al.*, 2007), more recent evidence based on phylogenetic analyses has provided inconsistent results (for example in birds, Cardillo *et al.*, 2005; Weir & Schluter, 2007). For instance, Wiens *et al.* (2009) found similar diversification rates in temperate and tropical clades of Old World frogs, and also found that tropical clades were older, supporting the time-based hypothesis for higher tropical diversification rates. It is likely that latitudinal differences in diversity have been generated by different combinations of ecological and evolutionary forces for different groups, and that different methodologies, taxonomic resolutions and distribution data sources may lead to contrasting results.

Concerning actinopterygian fish, the evidence presented here also suggests that clades that have been able to colonize different climatic zones (e.g. tropical and extra-tropical) have diversified faster, highlighting a positive link between colonizing new areas and producing new species. In agreement with our results, Owens *et al.* (1999) and Cardillo *et al.* (2003) found that diversification rates among bird and mammal clades were positively correlated with the total geographical area occupied, which may be attributed to increased opportunities for allopatric speciation provided by a greater area (Kisel *et al.*, 2011). However, this last result must be regarded with caution because, regardless of diversification rates, families with more species should be expected to have broader latitudinal ranges (all other things being equal).

Concluding remarks

Our analysis involved nearly 80% of the actinopterygian fish families and all available information at such taxonomic extent. The findings suggest a positive influence of the fragmentation and isolation characteristics of habitats and species on the diversification rates

of ray-finned fish at the family level. The potentially positive effects of isolation on both speciation and extinction rates might have muted the signal of any habitat fragmentation variable or dispersal-related trait on diversification rates. For instance when the size of isolated populations is too small, extinction rate may be high enough to balance speciation rate. However, here we find strong evidence for a positive role of geographic isolation on speciation. These findings suggest that, at natural levels, the physical fragmentation of habitats and isolation of populations have positive effects on speciation rather than extinction rates, resulting in enhanced diversification. The high levels of fragmentation inherent to freshwater environments have promoted a ‘freshwater fish paradox’ that may be extended to other freshwater taxa. Also, the future availability of more complete species distribution and phylogenetic data, and other biological traits (e.g. sexual selection or ecological specialization, both traits supposed to enhance speciation; Turelli *et al.*, 2001) will certainly improve our understanding of diversification processes of the most diversified clade of vertebrates.

Acknowledgements

We thank the French National Agency for Research (ANR-09-PEXT-008) for its financial support. We are very grateful to the Fishbase and OBIS consortia and especially to Nicolas Bailly, Frédéric Busson and Ward Appeltans for providing data on biotic and abiotic aspects for fish species and families. We also thank J.J. Wiens, F. Condamine, one anonymous referee and the associate editor I. Sanmartín, for their very helpful comments that greatly improved the manuscript. This is publication ISEM 2016-205.

References

- Agapow, P.-M. & Isaac, N.J.B. (2002) MacroCAIC: revealing correlates of species richness by comparative analysis. *Diversity and Distributions*, **8**, 41–43.
- Alfaro, M.E., Santini, F. & Brock, C.D. (2007) Do reefs drive diversification in marine teleosts? Evidence from the pufferfishes and their allies (Order Tetraodontiformes). *Evolution*, **61**, 2104–2126.
- Balian, E.V., Segers, H., Lévéque, C. & Martens, K. (2008) The freshwater animal diversity assessment : an overview of the results. *Hydrobiologia*, **595**, 627–637.
- Bartoń, K. (2015) MuMln: Multi-Model Inference, R package version 1.15.1. <https://cran.r-project.org/web/packages/MuMln/index.html>.
- Betancur-R, R., Wiley, E., Bailly, N., Miya, M., Lecointre, G. & Orti, G. (2014) Phylogenetic classification of bony fishes --Version 3. http://www.deepfin.org/Classification_v3.htm.
- Bloom, D.D., Weir, J.T., Piller, K.R. & Lovejoy, N.R. (2013) Do freshwater fishes diversify faster than marine fishes? a test using state-dependent diversification analyses and molecular phylogenetics of new world silversides (atherinopsidae). *Evolution*, **67**, 2040–2057.
- 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.
- Cardillo, M., Huxtable, J.S. & Bromham, L. (2003) Geographic range size, life history and rates of diversification in Australian mammals. *Journal of Evolutionary Biology*, **16**, 282–288.
- Cardillo, M., Orme, C.D.L. & Owens, I.P.F. (2005) Testing for latitudinal bias in diversification rates: an example using new world birds. *Ecology*, **86**, 2278–2287.
- Cowman, P.F. & Bellwood, D.R. (2011) Coral reefs as drivers of cladogenesis: expanding coral reefs, cryptic extinction events, and the development of biodiversity hotspots. *Journal of Evolutionary Biology*, **24**, 2543–2562.
- Cowman, P.F. & Bellwood, D.R. (2013) Vicariance across major marine biogeographic barriers: temporal concordance and the relative intensity of hard versus soft barriers. *Proceedings of the Royal Society B: Biological Sciences*, **280**, 20131541.
- Dawson, M.N. (2012) research letter: Species richness, habitable volume, and species densities in freshwater, the sea, and on land. *Frontiers of Biogeography*, **4**.
- Dawson, M.N. & Hamner, W.M. (2008) A biophysical perspective on dispersal and the geography of evolution in marine and terrestrial systems. *Journal of The Royal Society Interface*, **5**, 135–150.
- Dias, M.S., Cornu, J.-F., Oberdorff, T., Lasso, C.A. & Tedesco, P.A. (2013) Natural fragmentation in river networks as a driver of speciation for freshwater fishes. *Ecography*, **36**, 683–689.
- Dias, M.S., Oberdorff, T., Hugueny, B., Leprieur, F., Jézéquel, C., Cornu, J.-F., Brosse, S., Grenouillet, G. & Tedesco, P.A. (2014) Global imprint of historical connectivity on freshwater fish biodiversity. *Ecology Letters*, **17**, 1130–1140.
- Floeter, S.R., Rocha, L.A., Robertson, D.R., Joyeux, J.C., Smith-Vaniz, W.F., Wirtz, P., Edwards, A.J., Barreiros, J.P., Ferreira, C.E.L., Gasparini, J.L., Brito, A., Falcón, J.M., Bowen, B.W. & Bernardi, G. (2008) Atlantic reef fish biogeography and evolution. *Journal of Biogeography*, **35**, 22–47.
- Froese, R. & Pauly, D. (2013) Fishbase. World Wide Web electronic publication. www.fishbase.org.
- Gardezi, T. & da Silva, J. (1999) Diversity in relation to body size in mammals: a comparative study. *American Naturalist*, **153**, 110–123.
- Gillooly, J.F., Allen, A.P., West, G.B. & Brown, J.H. (2005) The rate of DNA evolution: Effects of body size and temperature on the molecular clock. *Proceedings of the National Academy of Sciences of the United States of America*, **102**, 140–145.
- Gittleman, J.L. & Purvis, A. (1998) Body size and species-richness in carnivores and primates. *Proceedings of the Royal Society B: Biological Sciences*, **265**, 113–119.
- Harmon, L.J. & Harrison, S. (2015) Species diversity is dynamic and unbounded at local and continental scales. *The American Naturalist*, **185**, 584–593.
- 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.

- Isaac, N.J.B., Jones, K.E., Gittleman, J.L. & Purvis, A. (2005) Correlates of species richness in mammals: body size, life history, and ecology. *The American naturalist*, **165**, 600–607.
- Kisel, Y., McInnes, L., Toomey, N.H. & Orme, C.D.L. (2011) How diversification rates and diversity limits combine to create large-scale species-area relationships. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **366**, 2514–2525.
- Kozak, K.H. & Wiens, J.J. (2016) Testing the relationships between diversification, species richness, and trait evolution. *Systematic Biology*, syw029.
- Lévéque, C., Oberdorff, T., Paugy, D., Stiassny, M.L.J. & Tedesco, P.A. (2008) Global diversity of fish (Pisces) in freshwater. *Hydrobiologia*, **595**, 545–567.
- Magallón, S. & Sanderson, M.J. (2001) Absolute diversification rates in angiosperm clades. *Evolution*, **55**, 1762–1780.
- 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.
- Mittelbach, G.G., Schemske, D.W., Cornell, H.V., Allen, A.P., Brown, J.M., Bush, M.B., Harrison, S.P., Hurlbert, A.H., Knowlton, N., Lessios, H.A., McCain, C.M., McCune, A.R., McDade, L.A., McPeek, M.A., Near, T.J., Price, T.D., Ricklefs, R.E., Roy, K., Sax, D.F., Schlüter, D., Sobel, J.M. & Turelli, M. (2007) Evolution and the latitudinal diversity gradient: speciation, extinction and biogeography. *Ecology Letters*, **10**, 315–331.
- Mora, C., Tittensor, D.P., Adl, S., Simpson, A.G.B. & Worm, B. (2011) How many species are there on Earth and in the ocean? *PLoS biology*, **9**, e1001127.
- Near, T.J., Eytan, R.I., Dornburg, A., Kuhn, K.L., Moore, J.A., Davis, M.P., Wainwright, P.C., Friedman, M. & Smith, W.L. (2012) Resolution of ray-finned fish phylogeny and timing of diversification. *Proceedings of the National Academy of Sciences*, **109**, 13698–13703.
- OBIS (2014) Data from the Ocean Biogeographic Information System. Intergovernmental Oceanographic Commission of UNESCO. Web. <http://www.iobis.org.>
- Orme, C.D.L., Freckleton, R., Thomas, G., Petzoldt, T., Fritz, S., Isaac, N.J.B. & Pearse, W. (2013) *caper: Comparative Analyses of Phylogenetics and Evolution in R*, R package version 0.5.2. <http://CRAN.R-project.org/package=caper>.
- Owens, I.P.F., Bennett, P.M. & Harvey, P.H. (1999) Species richness among birds: body size, life history, sexual selection or ecology? *Proceedings of the Royal Society B: Biological Sciences*, **266**, 933–939.
- Palumbi, S.R. (1994) Genetic divergence, reproductive isolation, and marine speciation. *Annual Review of Ecology and Systematics*, **25**, 547–572.
- Phillimore, A.B., Freckleton, R.P., Orme, C.D.L. & Owens, I.P.F. (2006) Ecology predicts large-scale patterns of phylogenetic diversification in birds. *The American naturalist*, **168**, 220–229.
- Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D. & R Core Team (2016) *nlme: Linear and Nonlinear Mixed Effects Models*, R package version 3.1-128. <https://cran.r-project.org/web/packages/nlme/index.html>.
- Price, S.A., Holzman, R., Near, T.J. & Wainwright, P.C. (2011) Coral reefs promote the evolution of morphological diversity and ecological novelty in labrid fishes. *Ecology Letters*, **14**, 462–469.
- R Development Core Team (2015) *R: A language and environment for statistical computing*, R Foundation for Statistical Computing, Vienna, Austria, <http://www.R-project.org/>.
- Rabosky, D.L., Santini, F., Eastman, J., Smith, S.A., Sidlauskas, B., Chang, J. & Alfaro, M.E. (2013) Rates of speciation and morphological evolution are correlated across the largest vertebrate radiation. *Nature Communications*, **4**, 1958.
- Radinger, J. & Wolter, C. (2014) Patterns and predictors of fish dispersal in rivers. *Fish and Fisheries*, **15**, 456–473.
- Riginos, C., Buckley, Y.M., Blomberg, S.P. & Treml, E.A. (2014) Dispersal capacity predicts both population genetic structure and species richness in reef fishes. *The American Naturalist*, **184**, 52–64.
- Rocha, L.A. & Bowen, B.W. (2008) Speciation in coral-reef fishes. *Journal of Fish Biology*, **72**, 1101–1121.

- Rolland, J., Jiguet, F., Jönsson, K.A., Condamine, F.L. & Morlon, H. (2014) Settling down of seasonal migrants promotes bird diversification. *Proceedings of the Royal Society B: Biological Sciences*, **281**, 20140473.
- Tedesco, P.A., Leprieur, F., Hugueny, B., Brosse, S., Dürr, H.H., Beauchard, O., Busson, F. & Oberdorff, T. (2012) Patterns and processes of global riverine fish endemism. *Global Ecology and Biogeography*, **21**, 977–987.
- Turelli, M., Barton, N.H. & Coyne, J.A. (2001) Theory and speciation. *Trends in Ecology & Evolution*, **16**, 330–343.
- Vega, G.C. & Wiens, J.J. (2012) Why are there so few fish in the sea? *Proceedings of the Royal Society B: Biological Sciences*, **279**, 2323–2329.
- Vermeij, G.J. & Grosberg, R.K. (2010) The great divergence: when did diversity on land exceed that in the sea? *Integrative and Comparative Biology*, **50**, 675–682.
- Webb, T.J. (2012) Marine and terrestrial ecology: unifying concepts, revealing differences. *Trends in Ecology & Evolution*, **27**, 535–541.
- Weir, J.T. & Schlüter, D. (2007) The latitudinal gradient in recent speciation and extinction rates of birds and mammals. *Science*, **315**, 1574–1576.
- Wiens, J.J. (2015a) Explaining large-scale patterns of vertebrate diversity. *Biology Letters*, **11**, 20150506.
- Wiens, J.J. (2015b) Faster diversification on land than sea helps explain global biodiversity patterns among habitats and animal phyla. *Ecology Letters*, **18**, 1234–1241.
- Wiens, J.J., Sukumaran, J., Pyron, R.A. & Brown, R.M. (2009) Evolutionary and biogeographic origins of high tropical diversity in old world frogs (Ranidae). *Evolution*, **63**, 1217–1231.
- Wollenberg, K.C., Vieites, D.R., Glaw, F. & Vences, M. (2011) Speciation in little: the role of range and body size in the diversification of Malagasy mantellid frogs. *BMC Evolutionary Biology*, **11**, 217.

SUPPORTING INFORMATION

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

Appendix S1. Complete actinopterygian fish family dataset.

Appendix S2. Mean age values (Myr) found for actinopterygian fish families in the literature and fossil databases

Appendix S3. Supplementary tables and figures

Biosketch

Pablo A. Tedesco is a research 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. All authors conceived and designed the study. PAT compiled data, performed modeling work, analyzed output data, and wrote the first draft of the manuscript, and all authors contributed substantially to revisions.

Editor: Isabel Sanmartín

Table 1: Results from model averaging and variable selection procedure with PGLS models of diversification rates as a function of the explanatory variables. The values given in the table are the relative importance of the variables, their estimated coefficients and corresponding *p*-values, for all three levels of relative extinction rate e and when including or excluding monospecific families. Pseudo-R², AIC and Lambda values of the complete PGLS models are also given.

Variables	Monospecific families included								
	$e = 0$			$e = 0.5$			$e = 0.9$		
	Importance	Coefficient	<i>p</i> -value	Importance	Coefficient	<i>p</i> -value	Importance	Coefficient	<i>p</i> -value
Mean body size	1.00	-0.0222	0.0000 0	1.00	-0.0214	0.0000 0	1.00	-0.0183	0.00000
Freshwater proportion	1.00	0.0607	0.0000 1	1.00	0.0614	0.0000 0	1.00	0.0596	0.00000
Reef-associated proportion	0.28	0.0054	0.7366 0	0.31	0.0096	0.5080 0	0.60	0.0184	0.09270
Migratory proportion	1.00	-0.0616	0.0000 3	1.00	-0.0592	0.0000 1	1.00	-0.0480	0.00001
Median Latitude (Abs)	0.63	0.0005	0.0736 0	0.49	0.0003	0.1560 0	0.29	0.0002	0.41633
Latitudinal Range	1.00	0.0015	0.0000 0	1.00	0.0015	0.0000 0	1.00	0.0013	0.00000
Full model parameters									
pseudo-R ²		0.36			0.39			0.44	
AIC		-935.53			-1005.34			-1185.99	
Pagel's lambda		0.67			0.68			0.68	
Monospecific families excluded									
Variables	$e = 0$			$e = 0.5$			$e = 0.9$		
	Importance	Coefficient	<i>p</i> -value	Importance	Coefficient	<i>p</i> -value	Importance	Coefficient	<i>p</i> -value
	1.00	-0.0186	0.0000 0	1.00	-0.0184	0.0000 0	1.00	-0.0166	0.00000
Freshwater proportion	1.00	0.0587	0.0000 0	1.00	0.0611	0.0000 0	1.00	0.0614	0.00000
Reef-associated proportion	1.00	0.0439	0.0014 0	1.00	0.0422	0.0017 5	1.00	0.0397	0.00046
Migratory proportion	1.00	-0.0559	0.0002 0	1.00	-0.0569	0.0000 6	1.00	-0.0502	0.00006
Median Latitude (Abs)	1.00	0.0006	0.0092 0	0.79	0.0005	0.0279 2	0.39	0.0002	0.26685
Latitudinal Range	1.00	0.0011	0.0000 0	1.00	0.0012	0.0000 0	1.00	0.0012	0.00000
Full model parameters									

pseudo-R ²	0.29	0.32	0.38
AIC	-1037.19	-1061.01	-1144.70
Page's lambda	0.87	0.84	0.74

Figure 1: Summary of the phylogeny, diversification rates (with $e = 0.9$), all tested variables and quality levels of origination age estimations. The phylogeny shows the evolutionary relationships for 377 families and corresponds to Rabosky *et al.* (2013) tree pruned to family level. Displayed fish orders correspond to the classification given by Betancur-R *et al.* (2014). Colors from red to blue correspond to values of each variable, from low to high.

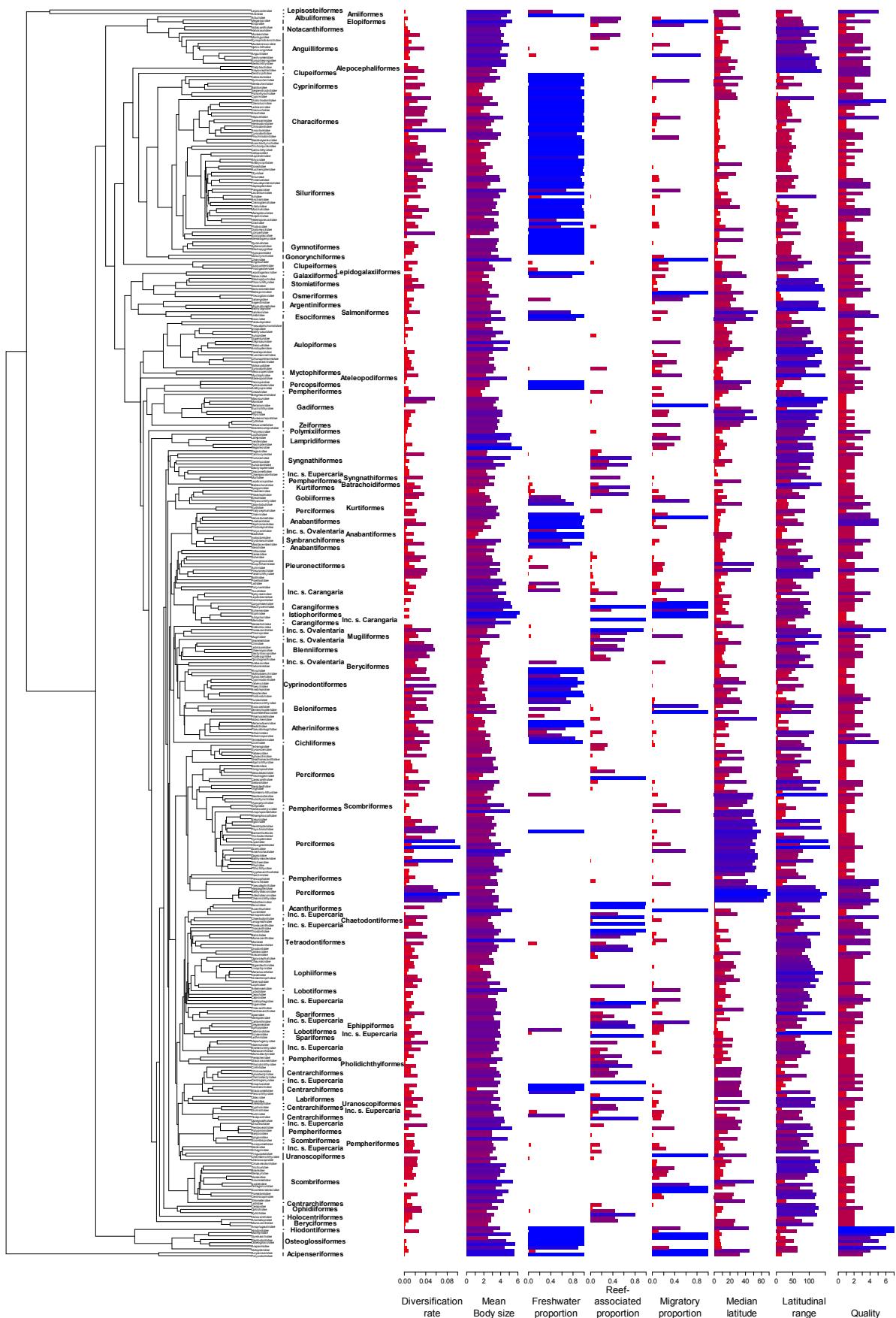


Figure 2: Partial-regression plots from PGLS models showing partial effects between all the tested variables and the diversification rates (with parameter $e = 0.9$) when including or excluding monospecific families.

