Received: 25 July 2018 | Revised: 1 February 2019 | Accepted: 26 February 2019
DOI: 10.1111/jbi.13559

RESEARCH PAPER

WILEY Journal of Biogeography

Correlations between broad-scale taxonomic and genetic differentiations suggest a dominant imprint of historical processes on beta diversities

Marine Robuchon[1,2] | Boris Leroy[1] | Céline Jézéquel[1,3] | Bernard Hugueny[1,3]

[1]Unité Biologie des organismes et écosystèmes aquatiques (BOREA), Muséum national d'Histoire naturelle, Sorbonne Université, Université de Caen Normandie, Université des Antilles, CNRS, IRD, Paris, France

[2]Centre d'écologie et des sciences de la conservation (CESCO), Muséum national d'Histoire naturelle, CNRS, Sorbonne Université, Paris, France

[3]Laboratoire Évolution & Diversité Biologique (EDB UMR 5174), Université de Toulouse Midi-Pyrénées, CNRS, IRD, UPS, Toulouse Cedex 9, France

Correspondence
Marine Robuchon, Unité Biologie des organismes et écosystèmes aquatiques (BOREA), Muséum national d'Histoire naturelle, Sorbonne Université, Université de Caen Normandie, Université des Antilles, CNRS, IRD, Paris, France.
Email: robuchon@mnhn.fr

Funding information
LabEx TULIP; LabEx BCDiv

Editor: Dr. Ceridwen Fraser

Abstract

Aim: Dispersal limitation, environmental selection and drift are known to influence both taxonomic similarity between communities and genetic similarity between populations. However, disentangling the relative roles of these processes on spatial patterns of differentiation—whether regarding taxonomic differentiation (TD) between communities or genetic differentiation (GD) between populations—is challenging. Investigating whether spatial patterns of TD and GD are correlated (β-SGDCs) is a promising approach to address this issue. Here, we investigated β-SGDCs over broad spatial scales and 22 freshwater fish species to elucidate the processes shaping TD and GD between drainage basins.

Location: Global scope, data mainly from Europe and North America.

Taxon: Actinopterygii and Petromyzontiformes (freshwater fishes).

Methods: We used Mantel tests to investigate the raw correlation between TD and GD. We carried out multiple regressions to characterize the effects of geographic distance (proxy of dispersal limitation), environmental distance (proxy of environmental selection) and pairwise harmonic mean area between basins (proxy of drift) on TD and GD taken separately. We then analyzed the correlation between the residuals of these two regressions with Mantel tests. Finally, we investigated whether the variation in β-SGDC between species was related to species traits.

Results: We detected an overall positive β-SGDC, both for the raw and the residual correlations. This implies that, unmeasured effects, other than geographic distance, environmental distance and harmonic mean area, influence the β-SGDC observed. Values of β-SGDCs greatly varied between species, but this was not explained by any species traits.

Main conclusions: Independently of the effects of geographic distance, environmental distance and harmonic mean area, the β-SGDC we observed suggests that historical processes strongly shaped the patterns of TD and GD between basins. Consequently, TD may be an appropriate proxy to explain the influence of historical processes on GD.

KEYWORDS
β-SGDC, dispersal limitation, distance decay, drift, environmental selection, freshwater fishes, historical processes, macrogenetics, Mantel tests, multiple regression of distance matrices

1 | INTRODUCTION

Understanding the mechanisms that shape taxonomic differentiation (TD) between spatially separated communities and those influencing genetic differentiation (GD) between spatially separated populations are long-standing issues in ecology and evolution (Mayr, 1963; Preston, 1962; Whittaker, 1960, 1972; Wright, 1943, 1951). More recently, these issues have been united (Kahilainen, Puurtinen, & Kotiaho, 2014; Vellend, 2003; Vellend & Geber, 2005), emphasizing that drift, dispersal and selection influence both TD between communities and GD between populations. Drift and selection affect the species frequencies within communities and gene frequencies within populations, which tend to increase differentiation. Conversely, dispersal of individuals between populations and dispersal of species between communities will tend to decrease differentiation.

Among the factors that can affect both TD and GD, geographic distance is probably the most studied. The similarity in taxonomic composition between communities tends to decrease with the geographic distance that separates them; a pattern long-recognized in biogeography and formalized as the distance decay of similarity (Nekola & White, 1999; Soininen, McDonald, & Hillebrand, 2007). Similarly, the genetic similarity between populations tends to decrease with the geographic distance separating them; a pattern long-recognized in population genetics and referred to as isolation-by-distance (IBD; Wright, 1943; Rousset, 1997). Two main non-mutually exclusive explanations can account for the distance decay of taxonomic and genetic similarity (see Soininen et al., 2007 and Orsini, Vanoverbeke, Swillen, Mergeay, & De Meester, 2013 for reviews). First, similarity decreases with geographic distance because the dispersal of organisms (i.e. their movement from a place to another) is limited by their intrinsic dispersal ability, which hence does not counteract anymore the differentiating effect of drift. Second, similarity decays with geographic distance because of decreasing similarity in spatially correlated environmental features; in such cases, the underlying explanation is environmental selection of different species (TD) or genotypes (GD) in distinct environments. However, differentiation between sites does not always increase with geographic distance. This is the case when the migrants are competitively excluded by already established communities (e.g. Almany, 2003) or populations (e.g. Fraser, Davies, Bryant, & Waters, 2018) at high densities, therefore counteracting the homogenizing effect of dispersal whatever the intrinsic dispersal ability of the migrants. This may also happen when the environmental variation between sites displays no—or negative—spatial autocorrelation (e.g. Derry, Arnott, Shead, Hebert, & Boag, 2009).

Although spatial patterns of TD and GD can be explained by environmental selection independently of limited dispersal and vice versa, disentangling the relative strengths of these mechanisms is challenging, as communities and populations that are geographically distant are also likely to inhabit different environments. Using modern spatial statistical methods, some recent studies have nonetheless taken up the challenge. In their study of global patterns of species turnover in terrestrial vertebrates, Qian and Ricklefs (2012) have shown that both dispersal limitation and environmental selection have played important roles in determining the patterns they observed. Similarly, in their reinterpretation of 34 representative studies, Orsini et al. (2013) highlighted that patterns associated with dispersal limitation were as common as those associated with local genetic adaptation in structuring population GD in the wild. However, the relative strength of dispersal limitation and environmental selection on differentiation patterns seems to vary according to the species and the spatial scale under consideration, both for taxonomic (Astorga et al., 2012; Moritz et al., 2013) and genetic (Orsini et al., 2013) differentiations. Comparative studies of multiple taxonomic groups suggest that the influence of dispersal limitation on TD is stronger for species with low dispersal ability than for more effective dispersers, and this result holds true both at the broad scale (e.g. Qian & Ricklefs, 2012) and at the fine scale (e.g. Astorga et al., 2012). Comparative studies of multiple species have also been identified as a promising avenue to reveal the factors influencing patterns of GD (Wang, Glor, & Losos, 2013). Yet, such studies are scarce and mainly carried out at relatively small spatial scales (e.g. Fourtune, Paz-Vinas, Loot, Prunier, & Blanchet, 2016; Wang et al., 2013) despite the increasing availability of broad-scale datasets in population genetics, suggesting that it is now time to embrace macrogenetics (Blanchet, Prunier, & De Kort, 2017).

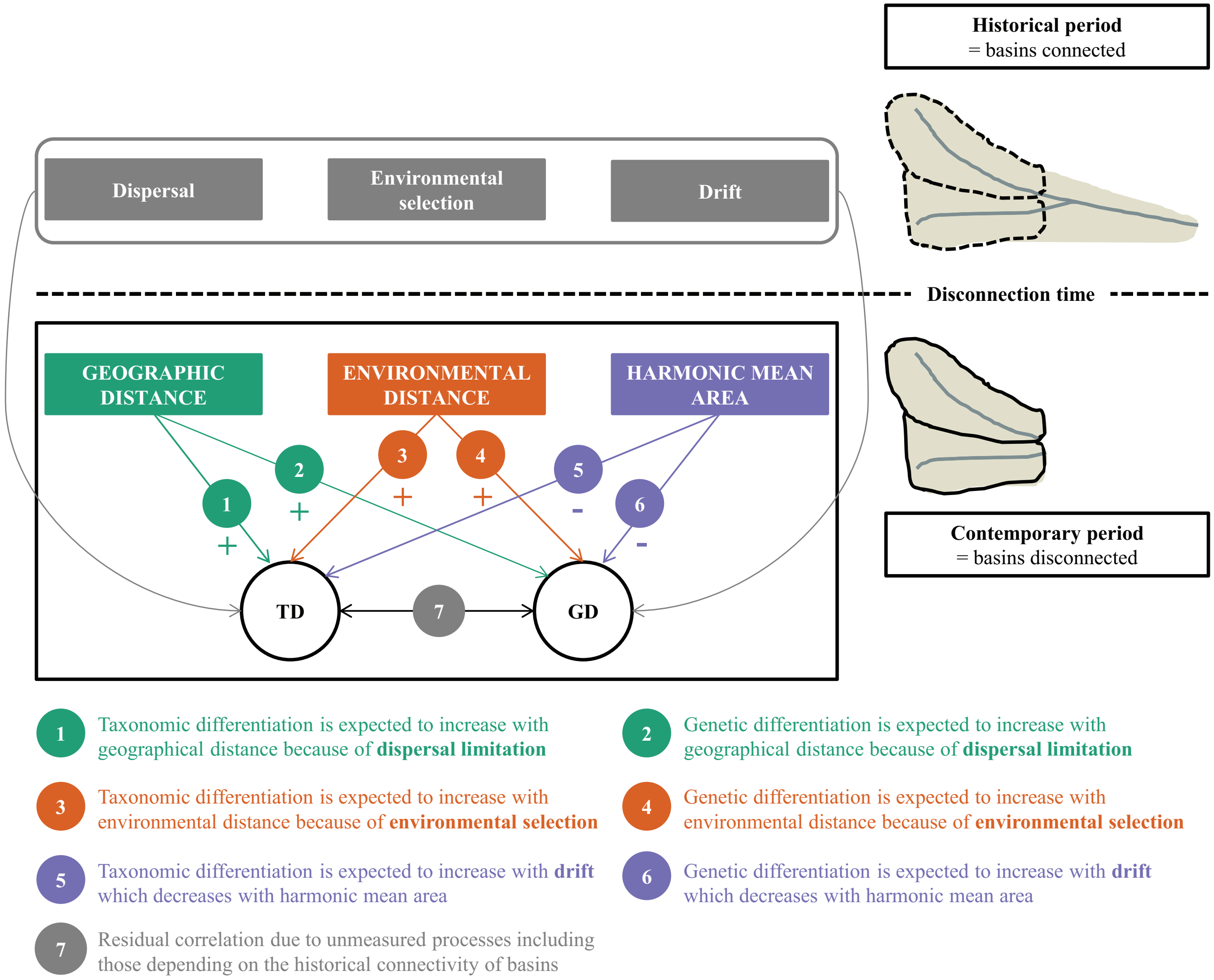
Theories in population genetics and community ecology acknowledge that dispersal, selection and drift together shape the differentiation between populations and between communities (see e.g. Vellend & Orrock, 2009 for a review of processes in both disciplines). However, the contribution of drift to TD or GD has been the subject of few empirical investigations, in contrast with dispersal limitation and environmental selection (Gilbert & Levine, 2017; Prunier, Dubut, Chikhi, & Blanchet, 2017). Genetic drift is the evolutionary process of random fluctuations in allelic frequencies occurring naturally in all populations due to their finite size, although it is stronger in small ones (Allendorf, 1986). Similarly, ecological drift corresponds to the random fluctuations in species frequencies occurring naturally in all communities due to their finite size and this too is stronger in small ones. Drift is the result of random sampling during the processes of birth, death and reproduction and ultimately leads to the loss of genetic diversity within populations, and species diversity within communities. Therefore, drift increases the differentiation between populations and between communities depending on their respective sizes: the smaller they are, the more they will differentiate. Dispersal of organisms between populations and between communities counteracts the differentiating effect of drift while environmental selection can accelerate it. Attributing the observed spatial patterns of GD or TD to dispersal limitation or to environmental selection without accounting for drift may thus be misleading.

Most of the explanatory variables used in analyses to infer the role of dispersal limitation and environmental selection on differentiation patterns are proxies representing present-day conditions. Thus, using current explanatory variables to investigate dispersal limitation and environmental selection on differentiation patterns is only relevant if the contemporary geography and environment of the study area has not changed substantially since these patterns were established (Wang et al., 2013). Otherwise, the historical processes that may have shaped the observed differentiation patterns, including those explaining colonization histories, may be overlooked. In addition, if dispersal is not continuous (as generally assumed) but rather intermittent because of landscape dynamics which have induced the appearance and disappearance of dispersal barriers, then two geographically close localities sharing a similar environment could, nevertheless, be dissimilar in terms of community composition and allele frequencies provided they have been isolated for a sufficiently long time (given their sizes) for drift to be effective. Nonetheless, examining which part of differentiation patterns is not explained by current explanatory variables may help to elucidate the influence of historical processes on those patterns.

In such complex contexts, understanding the rules that govern differentiation patterns (both taxonomic and genetic) may be enhanced by comparing taxonomic dissimilarities between communities and genetic dissimilarities between populations of a focal species (Lamy, Laroche, David, Massol, & Jarne, 2017). The few empirical studies that have examined these correlations, called beta species-genetic diversity correlations (β-SGDCs) (Kahilainen et al., 2014), revealed a majority of positive correlations (reviewed by Lamy et al., 2017). This suggests that the processes affecting taxonomic dissimilarity affect genetic dissimilarity of the species under study in the same way (Baselga, Gómez-Rodríguez, & Vogler, 2015; Baselga et al., 2013). However, because the relative strength of processes shaping patterns of TD and GD is expected to vary according to the species and the spatial scale under consideration, so is the strength of the putative resulting β-SGDCs. Among the rare studies of β-SGDCs over multiple species, some have highlighted that positive β-SGDCs were stronger for focal species with low dispersive abilities (e.g. Papadopoulou et al., 2011) while others did not find strong differences between species (e.g. Fourtune et al., 2016), and all were carried out at relatively small spatial scales. These contrasting results call for more studies of β-SGDCs for multiple species to gain a better understanding of whether and how variations in species traits lead to distinct patterns of β-SGDCs. In addition, β-SGDCs for multiple species have not been examined yet at broad spatial scales.

Freshwater fishes are a relevant model to start investigating this question for at least three reasons. First, as β-SGDCs for multiple freshwater fish species have been conducted at small scales within a single drainage basin (Fourtune et al., 2016), computing β-SGDCs for multiple freshwater fish species at broader scales will permit the comparison of β-SGDC variations across scales for the same biological model. Secondly, for strictly freshwater fishes, at present, almost no exchange of individuals occurs between drainage basins (hereafter “basins”) because they are isolated from each other by land and/or sea (although dispersal may occasionally occur between adjacent basins). Therefore, if basins are isolated from each other independently of the geographic distance between them, we can expect dispersal limitation to play a minor role on the broad-scale between-basin patterns of TD and GD. On the contrary, if basins have been isolated from each other for a long time, we can expect drift to play a significant role on the broad-scale between-basin patterns of TD and GD, as dispersal no longer counteracts the effect of drift. This is particularly true if population sizes and community sizes are small: the larger they are, the longer it will take to observe differentiation. Finally, while several recent studies have tried to disentangle the role of dispersal limitation, environmental selection and historical processes on patterns of TD between basins (Dias et al., 2014; Leprieur, Olden, Lek, & Brosse, 2009; Leprieur et al., 2011), the processes shaping GD at such broad scales have rarely been examined together. Studying β-SGDCs may therefore reveal whether the processes influencing between-basin TD, affect between-basin GD in the same way.

In this meta-analysis of 22 freshwater fish species, our main goal was to elucidate the processes underlying β-SGDCs at broad spatial scales. In particular, we wanted to test whether β-SGDCs can be explained by the parallel action of measurable processes (dispersal limitation, environmental selection and drift) on TD and GD, or are due to other, non-measurable processes. To that end, we first computed measures of GD between basins for each fish species. Then, for each species, we investigated (a) the raw correlation between TD and GD, (b) the effects of geographic distance (as a proxy of dispersal limitation), environmental distance (as a proxy of environmental selection) and pairwise harmonic mean area between basins (as a proxy of drift) on TD and GD taken separately and (c) the correlation between TD and GD independent of the effects of geographic distance, environmental distance and harmonic mean area (Figure 1). Such independent correlation integrates the effects of unmeasured processes, including the historical processes of dispersal, environmental selection and drift related to the historical connectivity of basins. This allowed us to characterize the overall (i.e. mean over the species) raw β-SGDC, the overall effects associated with each predictor for each response variable, and the overall independent β-SGDC. Finally, we tested whether the variation in β-SGDCs between species could be attributed to variations in species traits.



1 Taxonomic differentiation is expected to increase with geographical distance because of **dispersal limitation**

2 Genetic differentiation is expected to increase with geographical distance because of **dispersal limitation**

3 Taxonomic differentiation is expected to increase with environmental distance because of **environmental selection**

4 Genetic differentiation is expected to increase with environmental distance because of **environmental selection**

5 Taxonomic differentiation is expected to increase with **drift** which decreases with harmonic mean area

6 Genetic differentiation is expected to increase with **drift** which decreases with harmonic mean area

7 Residual correlation due to unmeasured processes including those depending on the historical connectivity of basins

FIGURE 1 Illustration of the different processes expected to act in parallel on between-basin taxonomic differentiation of freshwater fish communities (TD) and on between-basin genetic differentiation of freshwater fish populations (GD) at different periods and the proxies used in this study to infer those processes. The black frame encompasses the causal diagram depicting the relationships that we actually tested in this study, i.e. the effects of geographic distance (proxy of contemporary dispersal limitation), environmental distance (proxy of contemporary environmental selection) and pairwise harmonic mean area between basins (proxy of contemporary drift) on TD and GD. The double arrow between TD and GD represents the residual correlation, i.e. the correlation between TD and GD which is not explained by the parallel influence of contemporary dispersal limitation, contemporary environmental selection and contemporary drift on TD and GD. The numbers on the arrows and the signs below them correspond to our theoretical expectations regarding the different effects which are detailed in the text below the black frame. The grey arrows represent the putative effects of historical dispersal limitation, historical environmental selection and historical drift on TD and GD that we did not directly test in our statistical approach. The grey polygons on the right side represent the evolution of the basin configuration through time from one unique basin to two distinct and disconnected basins, with almost no contemporary dispersal of individuals between them (dispersal may occasionally occur between adjacent basins)

2 | MATERIALS AND METHODS

2.1 | Genetic differentiation and geographic distribution of genetic data

As we wanted to investigate GD at the broad scale, we restricted our analyses of spatial GD to species exhibiting a wide distribution. We downloaded sequence records of the mitochondrial gene cytochrome *c* oxidase I (COI) from BOLD (www.boldsystems.org) in May 2017 for the following widespread taxa: Cyprinidae, Salmonidae, *Lampetra*, *Petromyzon*, *Cobitis*, *Misgurnus*, *Barbatula*, *Silurus*, *Esox*, *Lota lota*, *Gymnocephalus* and *Sander*. We first cleaned and aligned sequences by taxon and then re-aligned them together to keep the same fragment of 501 nucleotides for the analyses. Sequence records without sufficient geographic information to assign them to a drainage basin (hereafter “basin”) or without a species name were discarded from our dataset. We defined a basin as the drainage area upstream from its mouth at the sea. According to this definition, a basin is isolated from other basins by barriers (sea or land) which are impassable for strictly freshwater fishes. The validity of species names was assessed using FishBase (Froese & Pauly, 2017), resulting in a total of 6,637 sequence records with a valid species names belonging to 716 distinct species distributed in 218 basins. We checked the environment (i.e. “freshwater”, “brackish”, “saltwater”) of each species using FishBase (Froese & Pauly, 2017), the status (i.e. “native”, “exotic” or “unknown”) of each species in each basin using the global database on freshwater fish species occurrence (Tedesco et al., 2017). We conserved only native records of strictly freshwater species, i.e. 4,894 records, corresponding to a total of 509 species in 176 basins.

We calculated between-basin GD (by pooling individuals in the drainage basins) for the polymorphic species present in at least four basins with a minimum of three individuals per basin. To calculate such GD, we converted single nucleotide polymorphisms (SNPs) of the sequence data into a table of individual genotypes and then calculated pairwise genetic distances for each pair of basins using G''_{ST} (Meirmans & Hedrick, 2011), the corrected version of Hedrick's G'_{ST} (Hedrick, 2005). G''_{ST} is independent from within population diversity and does not underestimate genetic distance when the number of populations is small (Meirmans & Hedrick, 2011). As G''_{ST} likely removes the effect of drift, we also calculated between-basin GD using F_{ST} (Nei, 1973) for comparison purpose.

All these different selection filters drastically reduced the number of species we included in analyses: we finally focused our analyses of between-basin GD on a total of 22 species in 38 basins (for more details about these basins, see Appendix S1 in Supporting Information). For instance, from all the species of *Lampetra* and *Petromyzon* we searched in BOLD, we conserved only *L. planeri* for the analyses. The number of

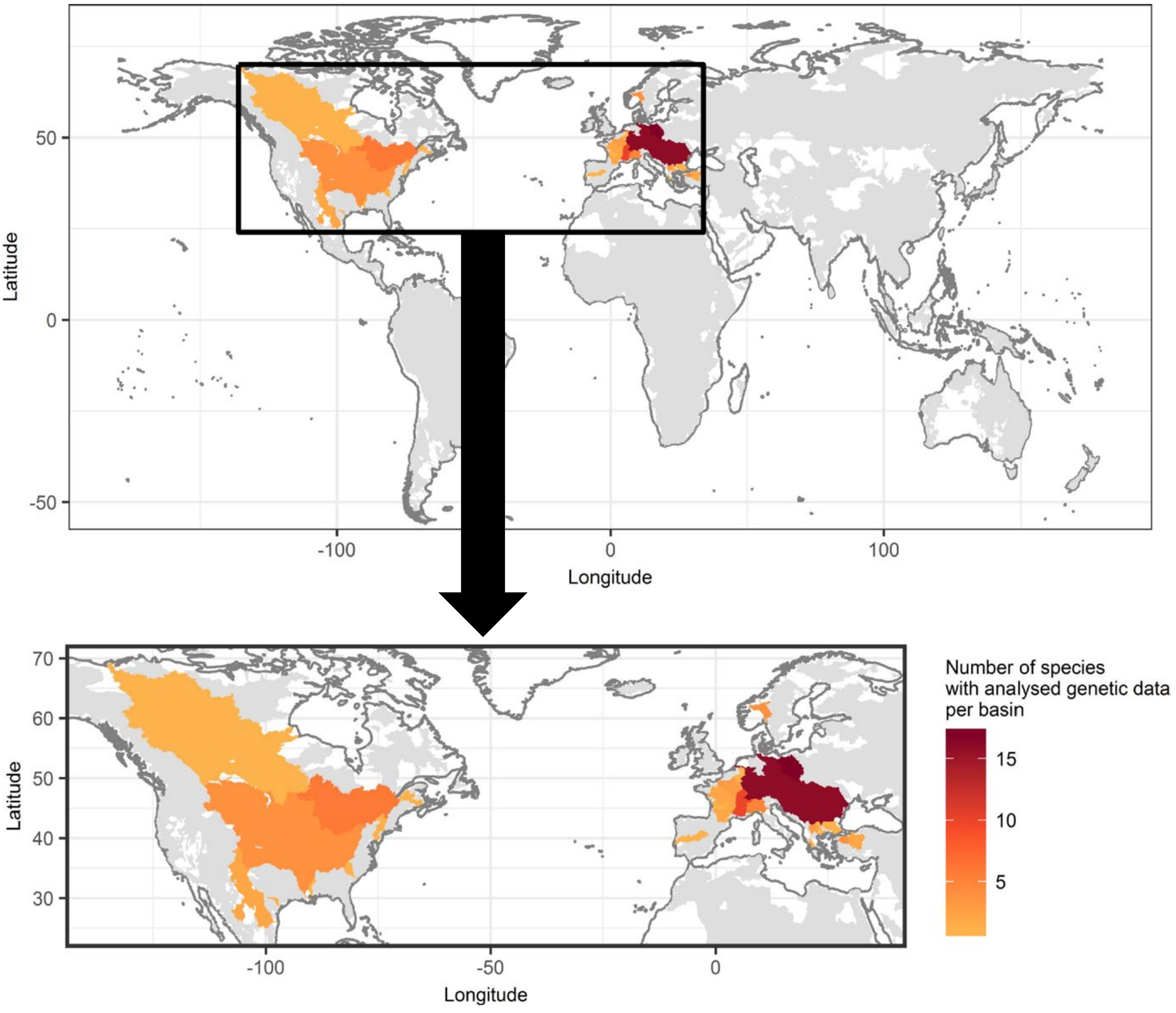


FIGURE 2 Spatial distribution of the number of species with analyzed genetic data per basin studied in our analyses at the global scale. Top: the rectangle indicates the location of the enlarged area. Bottom: enlarged area including Europe and North America. In total, our study encompasses 22 species with analyzed genetic data in 38 basins

species with analyzed genetic data per basin varies from 1 (for 20 basins) to 17 (for one basin, the basin “Odra”; Figure 2; Appendix S1). The number of basins per species varies from 4 (our minimum threshold) to 11 (for the species *Esox lucius*) covering on average 76% (standard deviation (*SD*) = 18%) of their native range, with species exhibiting on average 11.0 individuals per basin (*SD* = 11.0; Appendix S2). Most records are located in Europe and North America (Figure 2).

2.2 | Taxonomic differentiation

We calculated pairwise TD between each pair of basins based on the list of native species per basin extracted from the global database on freshwater fish species occurrence (Tedesco et al., 2017). We used Simpson dissimilarity (β_{sim})—the turnover component of Sørensen dissimilarity (Sørensen, 1948)—which is independent from total taxonomic richness (see Baselga, 2012 for review).

2.3 | Geographic distance, environmental distance and pairwise harmonic mean area

We computed between-basin geographic distance by calculating the distance between basin centroids. The average distance between basins was 5,295 km (*SD* = 3,920 km). Measurements of geographic distance were then centred and scaled before further analyses.

To characterize the environment of the basins, we used variables related to the contemporary climate known to influence freshwater fish alpha diversity patterns at the global scale (e.g. Dias et al., 2014; Oberdorff et al., 2011; Tedesco et al., 2012; Tisseuil et al., 2013) i.e. mean annual temperature and precipitation (http://www.worldclim.org/), mean annual surface runoff (http://www.grdc.sr.unh.edu/) and mean annual actual and potential evapotranspiration (https://cgiarcsi.community/data/global-aridity-and-pet-database/). The mean values of all the variables were computed for each basin. To compute pairwise environmental distances between each pair of basins, we performed a principal component analysis (PCA) on these environmental variables (previously centred and scaled) and calculated the Euclidean distance between basins based on the first three PCA axes.

Finally, we calculated the pairwise harmonic mean area between each pair of basins. Indeed, the harmonic mean population size between pairs allows the role of drift on GD (Serrouya et al., 2012) to be taken into account and can be computed using environmental proxies for local carrying capacities (Prunier et al., 2017). In our case, we used the surface area of the basin as a raw environmental proxy for local carrying capacity. The basins studied had an average surface area of 257,589 km^2 (*SD* = 584,803 km^2). Measurements of harmonic mean area were centred and scaled before further analyses.

2.4 | Statistical analyses

First, we investigated the raw correlation between TD and GD for each species individually using a Mantel correlation test with 999 permutations. To test the null hypothesis that there is no relationship between TD and GD using the information provided by all the species examined, we considered that this hypothesis has been tested independently *n* times, with *n* being the number of species included in the analyses. To combine the results from several independent tests bearing upon the same overall hypothesis, we used a modification of the Fisher's combined probability test. Under the null hypothesis that TD and GD are unrelated, observed *p*-values associated with the Mantel's test of each species are expected to be distributed according to a uniform distribution in [0, 1]. We therefore compared the observed mean *p*-value over the species to a null distribution of mean *p*-values obtained by sampling randomly 10,000 times one *p*-value by species in a uniform distribution in [0, 1]. We calculated the *p*-value associated with the combined test as the frequency at which null mean *p*-values were below the observed mean *p*-value. This procedure avoids some problems identified by the use of a chi-square test (e.g. Whitlock, 2005) which was the method initially proposed by Fisher.

To investigate the putative roles of dispersal limitation, environmental selection and drift on TD and GD, for each species we built multiple regressions of geographic distance (D_{geo}), environmental distance (D_{env}) and harmonic mean area (D_{area}) on TD and GD taken separately (Figure 1) as follows: $TD = \alpha_0 + \alpha_1{*}Dgeo + \alpha_2{*}D_{env} + \alpha_3{*}D_{area}$ and $GD = \beta_0 + \beta_1{*}Dgeo + \beta_2{*}D_{env} + \beta_3{*}D_{area}$. Our approach is very similar to multiple regressions on distance matrices (MRM; Lichstein, 2007) except in the way we calculated *p*-values associated with partial regression coefficients. As in MRM, we permuted the rows and associated columns of the response distance matrix simultaneously. We repeated this operation 1,999 times while holding the explanatory distance matrices constant to generate null distributions for partial regression coefficients. Then, in MRM, the *p*-value associated with a partial regression coefficient is calculated with a two-sided permutation test using the pseudo-t of Legendre, Lapointe, and Casgrain (1994), i.e. it tests whether the t-statistic associated with each regression coefficient is lower or higher than expected under a null distribution of pseudo-t obtained by permutation. In our approach, we hypothesized that geographic distance and environmental distance have a positive effect on both TD and GD (dispersal limitation and environmental selection enhance differentiation and are positively related to geographic and environmental distances respectively), while harmonic mean area has a negative effect on TD and GD (drift enhances differentiation and is negatively related to the harmonic mean area). To increase the statistical power of the tests, we therefore carried out one-sided tests for calculating the *p*-values associated with each partial regression coefficient. We tested the overall (i.e. over the species) effects of geographic distance, environmental distance and harmonic mean area on TD and GD by performing a combined probability test as described for the raw correlation between TD and GD, but this time considering probabilities associated with partial regression coefficients.

To examine whether TD and GD were still correlated once the effects of geographic distance and environmental distance had been taken into account, we then performed a Mantel correlation test between the residuals of the multiple regression on TD and the residuals of the multiple regression on GD with 999 permutations for each species. We tested this overall (i.e. over the species) residual correlation using a combined probability test as described for the raw correlation between TD and GD.

Finally, we examined the heterogeneity between species regarding the residual correlation between TD and GD and tested whether some species traits could explain such heterogeneity. The species traits initially considered were body length, longevity in the wild, vulnerability, used in aquaculture, use as bait, use in the aquarium, habitat, migratory behaviour and dispersal ability—which was calculated using a formula giving dispersal distance as a function of body size and caudal fin aspect ratio (Radinger & Wolter, 2014). However, as migratory behaviour was “potamodromous” for 16 out of the 22 species, “non-migratory” for one species and the information was not available for the five species left, we excluded this non-informative trait from our analyses. Species traits were directly extracted from FishBase (Froese & Pauly, 2017). All the analyses and the corresponding figures were realized with R 3.5.1 (R Core Team 2018) using the packages ‘ade4’ (Dray & Dufour, 2007), ‘ape’ (Paradis, Claude, & Strimmer, 2004), ‘betapart’ (Baselga, Orme, Villeger, De Bortoli, & Leprieur, 2018), ‘Biostrings’ (Pagès, Aboyoun, Gentleman, & DebRoy, 2018), ‘broom’ (Robinson & Hayes, 2018), ‘ggplot2’ (Wickham, 2016), ‘ggthemes’ (Arnold, 2018), ‘hierfstat’ (Goudet & Jombart, 2015), ‘maptools’ (Bivand & Lewin-Koh, 2018), ‘mmod’ (Winter, 2012), ‘RColorBrewer’ (Neuwirth, 2014), ‘rgdal’ (Bivand, Keitt, & Rowlingson, 2018), ‘rgeos’ (Bivand & Rundel, 2018), ‘reshape2’ (Wickham, 2007), ‘seriation’ (Hahsler, Buchta, & Hornik, 2018) and ‘vegan’ (Oksanen et al., 2018).

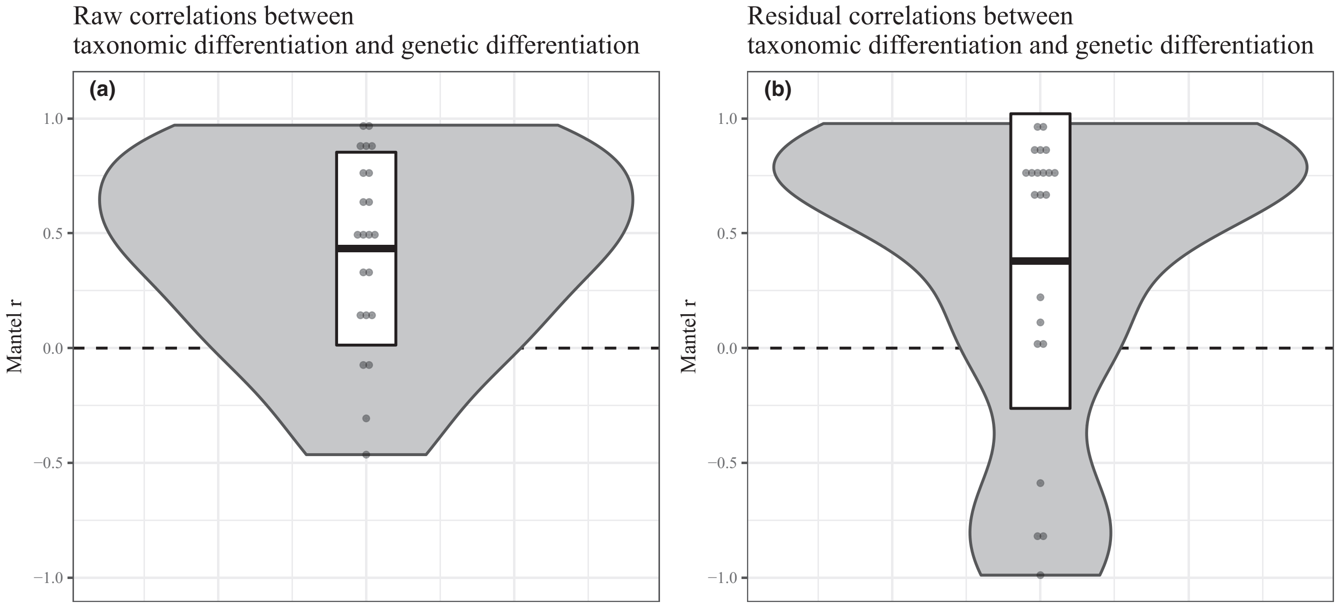
3 | RESULTS

We detected an overall positive correlation between TD and GD, significantly different from 0. This result holds true both for the raw correlation between TD and GD ($r_{\text{mean} \pm SD} = 0.433 \pm 0.420$; Figure 3a; Appendices S3 and S4) and for the residual correlation, i.e. the correlation between TD and GD independent of the effects of geographic distance, environmental distance and harmonic mean area ($r_{\text{mean} \pm SD} = 0.379 \pm 0.641$; Figure 3b; Appendices S3 and S4).

There was almost no overall effect of geographic distance on TD ($\alpha_{1\text{mean} \pm SD} = 0.032 \pm 1.051$) or on GD ($\beta_{1\text{mean} \pm SD} = -0.011 \pm 1.060$), and these overall effects did not differ significantly from 0 (Figure 4; Appendices S5–S7). In contrast, we detected an overall positive effect of environmental distance on both TD ($\alpha_{2\text{mean} \pm SD} = 0.405 \pm 0.956$) and GD ($\beta_{2\text{mean} \pm SD} = 0.229 \pm 0.865$), although it was only significantly different from 0 for TD (Figure 4; Appendices S5–S7). There was an overall negative effect of harmonic mean area on both TD ($\alpha_{3\text{mean} \pm SD} = -0.178 \pm 0.617$) and GD ($\beta_{3\text{mean} \pm SD} = -0.092 \pm 0.713$), but it was only significantly different from 0 for TD (Figure 4; Appendices S5–S7). Using F_{ST} as a measure of GD did not change the trend we observed: there was an overall negative but not significant effect of harmonic area on F_{ST} (Appendices S6 and S7).

The results of these multiple regressions as well as the residual correlations between TD and GD are summarized in Figure 5. Overall, they suggest that the effect of geographic distance, environmental distance and harmonic mean area only explain a small part of the correlation between TD and GD.

Despite this strong overall independent correlation between TD of freshwater fish communities and GD of freshwater fish populations, the correlation between TD and GD varied greatly between


Raw correlations between taxonomic differentiation and genetic differentiation

Residual correlations between taxonomic differentiation and genetic differentiation

FIGURE 3 Violin plots showing the distribution and the probability density of (a) raw correlations coefficients and (b) residual correlation coefficients between taxonomic differentiation and genetic differentiation for the 22 fish species examined at global scale. Crossbars indicate mean ± standard deviation

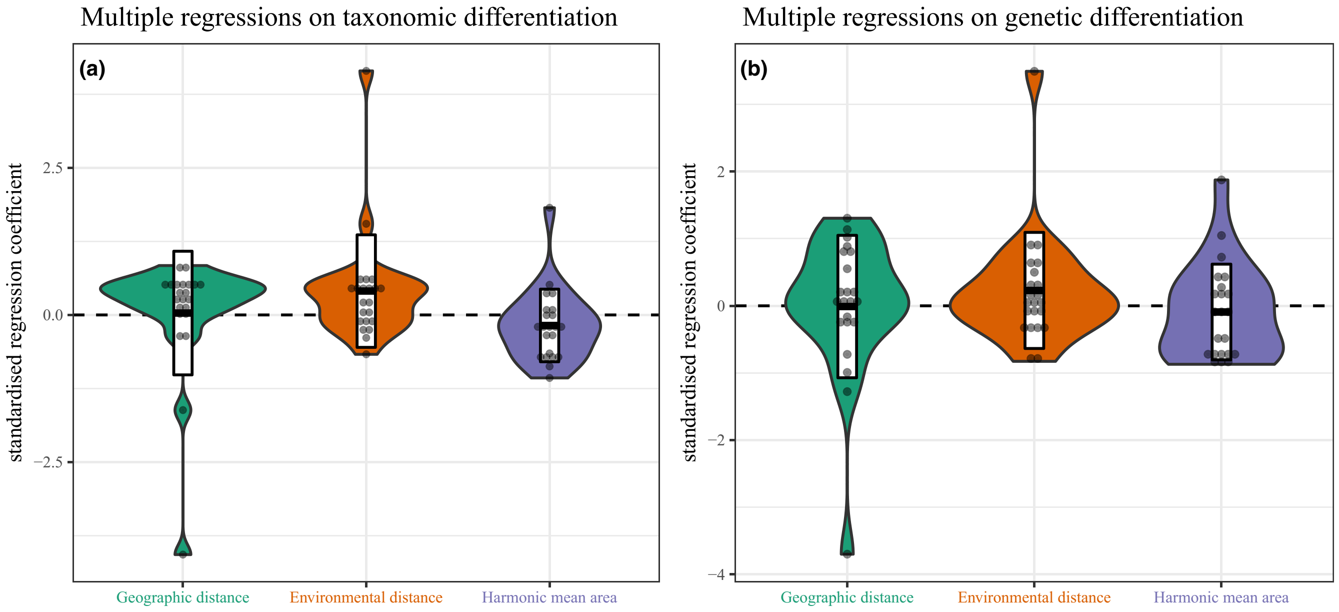


FIGURE 4 Violin plots showing the distribution and the probability density of the effect of geographic distance, environmental distance and harmonic mean area (as standardized regression coefficients) for the 22 fish species examined at global scale on (a) taxonomic differentiation and (b) genetic differentiation. Crossbars indicate mean ± standard deviation

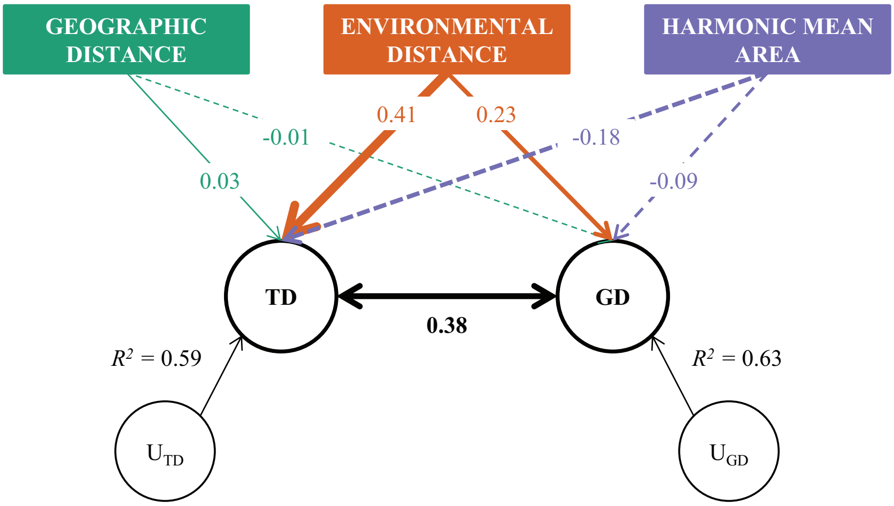


FIGURE 5 Causal diagram depicting the results of multiple regressions of geographic distance, environmental distance and harmonic mean area on taxonomic differentiation (TD) and genetic differentiation (GD) taken separately. Arrows represent putative causal effects through the standardized regression coefficients (positive values: solid arrows; negative values: dotted arrows; arrow width is proportional to coefficient value). The R^2 represents the part of variance explained by the multiple regressions on TD (left) and GD (right). U_{TD} and U_{GD} represent unspecified factors influencing TD and GD respectively. The double arrow between TD and GD corresponds to the residual correlation. All numerical values represent means for the 22 fish species examined at global scale

species (Figure 3b; Appendix S3). For some species, such as the common roach *Rutilus rutilus* (Figure 6a,b), there was a sound correlation between TD and GD, whether we considered the raw or residual correlation. In other species, such as the common chub *Squalius cephalus*, TD and GD were not correlated (Figure 6c,d). In this species, GD was better predicted by geographic distance and harmonic mean area (Appendix S5). However, none of the species traits examined explained this heterogeneity of responses between species (Appendix S8).

4 | DISCUSSION

Our analyses indicated that TD and GD are strongly correlated overall, and independently of the parallel effects of geographic distance, environmental distance and harmonic mean area on TD and GD. This suggests that the parallel influence of dispersal limitation, environmental selection and drift on TD and GD only partially explains the overall β-SGDC we observed. This independent, unexplained correlation between TD and GD thus probably results from

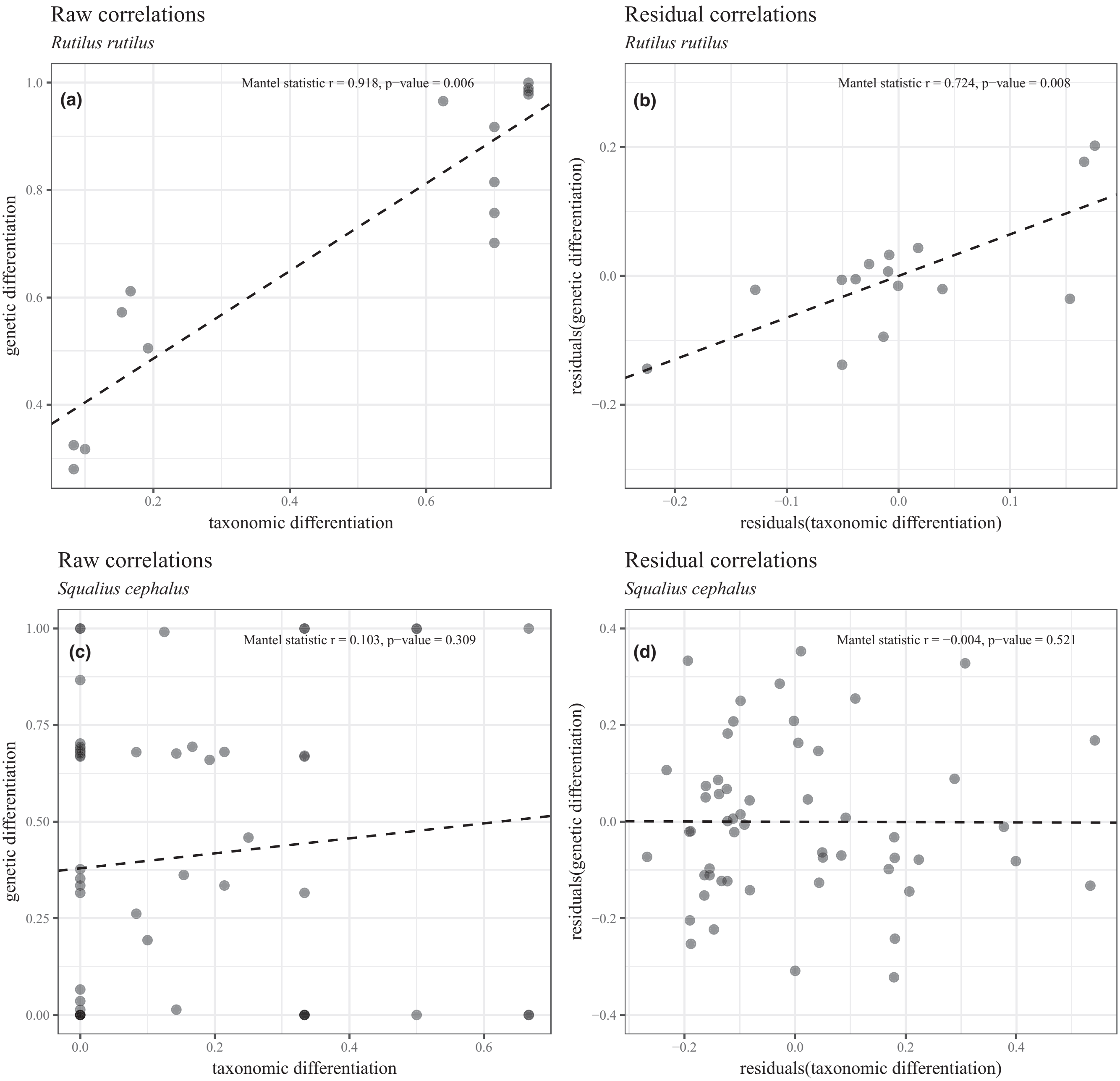


FIGURE 6 Observed (dots) and predicted (dashed line) raw values (a, c) and residual values from the multiple regressions (b, d) of genetic differentiation as a function of taxonomic differentiation for the two fish species (a, b) *Rutilus rutilus* and (c, d) *Squalius cephalus*

processes that we did not explicitly consider in our study. The most probable explanation is that the independent β-SGDC we observed is due to the parallel action of historical processes, which occurred when the basins were still connected, on TD and GD. These historical processes include all the different processes that explain the colonization history of basins (Orsini et al., 2013): (a) the dispersal of individuals between previously connected basins in interaction with environmental and/or biotic selection on their dispersal route and (b) environmental and/or biotic selection in interaction with drift within their place of arrival and/or origin that may have led to extinction. This explanation is supported by previous studies on the influence of historical processes on broad-scale TD between freshwater fish communities on the one hand and on broad-scale GD between freshwater fish populations on the other hand. Two recent studies on global freshwater fish biodiversity have shown that present-day patterns of taxonomic dissimilarity across basins are well explained by the historical connectivity of basins (Dias et al., 2014) and also historical climatic oscillations that took place during the Quaternary and the Holocene (Leprieur et al., 2011). For fish species, it is well known that historical dispersal influences genetic structure across basins (Carvalho, 1993), especially for zones that were subject to intense postglacial recolonization such as our study area. Multispecies comparisons at large spatial scales in Europe (Seifertová, Bryja, Vyskočilová, Martínková, & Šimková, 2012) and in North America

(April, Hanner, Mayden, & Bernatchez, 2013) have shown that the climatic fluctuations that caused glacial cycles during the Pleistocene had a generalized effect on the patterns of fish genetic divergence. In addition, three other, non-exclusive hypotheses may explain the independent β-SGDC we observed. First, the proxies we used to account for dispersal limitation, environmental selection and drift may not fully represent these processes. For instance, we may have overlooked some environmental variables that are important in driving environmental selection. Alternatively, we may have poorly estimated the role of drift, either because the area of the basin we used to account for drift may be too approximate to accurately represent the carrying capacities of populations and communities or because we did not consider that the influence of drift within a basin increases with its age of isolation. Secondly, the sequence data we used to calculate GD may not detect dispersal limitation and/or environmental selection properly, either because they did not cover the entire native ranges of species (see possible consequences of such sampling bias in e.g. Meirmans, 2015), or because the mitochondrial gene COI is not directly under environmental selection. Finally, the proposed frameworks to interpret SGDCs (Lamy et al., 2017; Vellend & Geber, 2005) suggest that correlations between species diversity and genetic diversity may also result from a causal action of species diversity on genetic diversity or vice versa. For instance, if the focal species (i.e. the species for which we studied genetic diversity) is a facilitator for the other component species of the community, we would expect their populations sizes to co-vary positively, resulting in a positive α-SGDC; the converse is expected if the focal species is a competitor. If the nature of the interaction between the focal species and the other species of the community is generally the same within each basin, such causal effect is not expected to influence the β-SGDC we observed. However, if the focal species is facilitator in one basin and competitor in another, we would expect the population sizes of the other component species of the community to increase in the first case and decrease in the second case. This would result in more ecological drift within the community where the focal species is competitor, and therefore more TD between basins than would be expected without this causal effect. In the end, such causal effect would tend to reduce β-SGDC because it would increase TD between basins without influencing GD.

Moreover, our results contribute new insights into the study of β-SGDCs. With the study of 22 new β-SGDCs, we increased the number of β-SGDCs (43) previously reported in the literature by more than 50% (see Lamy et al., 2017 for the most recent review). We found an average raw correlation between TD and GD of 0.433, and an independent correlation (i.e. independent from the parallel effects of geographic distance, environmental distance and harmonic mean area on TD and GD) of 0.379. Both of these values are superior to the average value of 0.221 reported by Lamy et al. (2017). They are also around three times superior to the average value of 0.139 found by Fourtune et al. (2016) in their study of four freshwater fish species at a smaller scale, within one drainage basin. This suggests that the strong correlations we found are probably more related to the size and/or the isolation of the habitats we studied (i.e. large drainage basins isolated from each other) rather than our biological model. Previous syntheses have reported that α-SGDCs (i.e. correlations between genetic diversity within populations and species diversity within communities) are stronger and more often positive in island-like habitats (Vellend & Geber, 2005; Vellend et al., 2014), in agreement with the predictions of theoretical models (Laroche, Jarne, Lamy, David, & Massol, 2015). As drainage basins are typically island-like habitats (Hugueny, 1989; Sepkoski & Rex, 1974), our findings suggest that this explanation may hold true for β-SGDCs as well. The observed differences also suggest that for a system without present-day dispersal the β-SGDC is stronger than in systems with present-day dispersal. Actually, the opposite could have been expected (i.e. weaker β-SGDCs for systems without present-day dispersal than for systems with present-day dispersal) if the genetic similarity between two previously connected populations had been erased by genetic drift occurring after their isolation without changes in species composition (i.e. no species extinction). However, our results suggest that drift did not influence significantly the patterns of GD we observed. As this result holds true both for G''_{ST} and F_{ST}, this is probably because we mainly focused on large drainage basins and consequently large population sizes.

Our results also indicate that the sign and magnitude of β-SGDCs varied greatly between the different species we examined. We provided two examples. On the one hand, we showed that the β-SGDC was strong for *Rutilus rutilus*, suggesting that historical processes largely shaped the patterns of GD observed in this species. This is consistent with a previous study indicating that *R. rutilus* is composed of two historically isolated, independently evolving sets of populations (Larmuseau, Freyhof, Volckaert, & Van Houdt, 2009). On the other hand, our results indicate that GD in *Squalius cephalus* was not correlated with TD, but was well explained by geographic distance. This suggests that dispersal limitation has a more important effect than any other processes on broad-scale GD in *S. cephalus*. This might seem surprising, as previous analyses of the mitochondrial gene cytochrome *b* in that species indicated four main lineages originating from multiple glacial refugia (Durand, Persat, & Bouvet, 1999; Seifertová et al., 2012). However, the use of a combination of both mitochondrial and nuclear DNA, Seifertová et al. (2012) also suggested that GD in *S. cephalus* resulted from a combination of different factors, i.e. postglacial colonization from different refugia or recent evolutionary processes such as drift or dispersal limitation. Our results based on the analysis of the mitochondrial gene COI strongly support the dispersal limitation hypothesis although, as we did not design the sampling, but analysed existing data, we might not have captured the whole picture. As *R. rutilus* and *S. cephalus* have similar dispersal abilities, this difference of β-SGDC between the two species does not seem to be related to their difference in dispersal abilities. Beyond these two species and this single trait, our analyses did not permit us to uncover any trait that could explain the variability of β-SGDC over the 22 species we examined. This is surprising because life-history traits of species are known to shape spatial patterns of GD (e.g. Duminil et al., 2007; Kelly & Palumbi, 2010), and therefore β-SGDC. This unexpected finding may be due to different reasons. First, the variability of β-SGDC may be due to species traits that we did not consider. Second, the β-SGDC at the specific scale was sometimes examined for a small number of basins, which may result in wide variations in the estimated correlation coefficients. Therefore, single species outcomes should be interpreted with great caution.

To conclude, this study of broad-scale β-SGDC for multiple freshwater fish species suggests that historical processes greatly contributed to the shape of present-day patterns of TD and GD, independently of contemporary processes. This implies that TD may be an appropriate proxy to explain the role of historical processes on patterns of GD, which remains a challenging issue in landscape genetics (Dyer, Nason, & Garrick, 2010). We believe that further studies of broad-scale β-SGDCs comparing organisms whose dispersal is limited within habitat patches with organisms that are able to disperse across habitat patches, will help to investigate the extent to which our findings can be generalized.

ACKNOWLEDGEMENTS

This study was supported by the French State through the Research National Agency under the LabEx ANR-10-LABX-0003-BCDiv, within the framework of the programme 'Investing for the future' (ANR-11-IDEX-0004-02). B.H. and C.J. are supported by the French Laboratory of Excellence project "TULIP" (ANR-10-LABX-41; ANR-11-IDEX-0002-02). We thank Agnès Dettaï and Gaël Denys for their valuable suggestions during the planning of this research work. We also thank Céline Bellard, Juliette Delavenne and Jean-Baptiste Mihoub for their relevant advice on figures and rewording that greatly improved the manuscript. We are grateful for the helpful comments of the associate editor and of two anonymous reviewers.

DATA AVAILABILITY

The data and the script used for the analyses are available from the Dryad Digital Repository: https://doi.org/10.5061/dryad.8vb6160.

ORCID

Marine Robuchon https://orcid.org/0000-0001-5873-2915

Boris Leroy https://orcid.org/0000-0002-7686-4302

REFERENCES

Allendorf, F. W. (1986). Genetic drift and the loss of alleles versus heterozygosity. *Zoobiology*, *5*(2), 181–190.

Almany, G. R. (2003). Priority effects in coral reef fish communities. *Ecology*, *84*(7), 1920–1935. https://doi.org/10.1890/0012-9658(2003)084[1920:PEICRF]2.0.CO;2

April, J., Hanner, R. H., Mayden, R. L., & Bernatchez, L. (2013). Metabolic rate and climatic fluctuations shape continental wide pattern of genetic divergence and biodiversity in fishes. *PLoS ONE*, *8*(7), e70296. https://doi.org/10.1371/journal.pone.0070296

Arnold, J. B. (2018). ggthemes: Extra Themes, Scales and Geoms for 'ggplot2'. R package version 4.0.1. Retrieved from https://CRAN.R-project.org/package=ggthemes.

Astorga, A., Oksanen, J., Luoto, M., Soininen, J., Virtanen, R., & Muotka, T. (2012). Distance decay of similarity in freshwater communities: Do macro-and microorganisms follow the same rules? *Global Ecology and Biogeography*, *21*(3), 365–375. https://doi.org/10.1111/j.1466-8238.2011.00681.x

Baselga, A. (2012). The relationship between species replacement, dissimilarity derived from nestedness, and nestedness. *Global Ecology and Biogeography*, *21*(12), 1223–1232. https://doi.org/10.1111/j.1466-8238.2011.00756.x

Baselga, A., Fujisawa, T., Crampton-Platt, A., Bergsten, J., Foster, P. G., Monaghan, M. T., & Vogler, A. P. (2013). Whole-community DNA barcoding reveals a spatio-temporal continuum of biodiversity at species and genetic levels. *Nature Communications*, *4*, 1892. https://doi.org/10.1038/ncomms2881

Baselga, A., Gómez-Rodríguez, C., & Vogler, A. P. (2015). Multi-hierarchical macroecology at species and genetic levels to discern neutral and non-neutral processes. *Global Ecology and Biogeography*, *24*(8), 873–882. https://doi.org/10.1111/geb.12322

Baselga, A., Orme, D., Villeger, S., De Bortoli, J., & Leprieur, F. (2018). betapart: Partitioning beta diversity into turnover and nestedness components. R package version 1.5.0. Retrieved from https://CRAN.R-project.org/package=betapart.

Bivand, R., Keitt, T., & Rowlingson, B. (2018). rgdal: Bindings for the 'geospatial' data abstraction library. R package version 1.3-4. Retrieved from https://CRAN.R-project.org/package=rgdal.

Bivand, R., & Lewin-Koh, N. (2018). maptools: Tools for reading and handling spatial objects. R package version 0.9-3. Retrieved from https://CRAN.R-project.org/package=maptools.

Bivand, R., & Rundel, C. (2018). rgeos: Interface to geometry engine - open source ('GEOS'). R package version 0.3-28. Retrieved from https://CRAN.R-project.org/package=rgeos.

Blanchet, S., Prunier, J. G., & De Kort, H. (2017). Time to go bigger: Emerging patterns in macrogenetics. *Trends in Genetics*, *33*(9), 579–580. https://doi.org/10.1016/j.tig.2017.06.007

Carvalho, G. R. (1993). Evolutionary aspects of fish distribution: Genetic variability and adaptation. *Journal of Fish Biology*, *43*(Suppl. A), 53–73. https://doi.org/10.1111/j.1095-8649.1993.tb01179.x

Derry, A. M., Arnott, S. E., Shead, J. A., Hebert, P. D., & Boag, P. T. (2009). Ecological linkages between community and genetic diversity in zooplankton among boreal shield lakes. *Ecology*, *90*(8), 2275–2286. https://doi.org/10.1890/07-1852.1

Dias, M. S., Oberdorff, T., Hugueny, B., Leprieur, F., Jézéquel, C., Cornu, J. F., … Tedesco, P. A. (2014). Global imprint of historical connectivity on freshwater fish biodiversity. *Ecology Letters*, *17*(9), 1130–1140. https://doi.org/10.1111/ele.12319

Dray, S., & Dufour, A. (2007). The ade4 package: Implementing the duality diagram for ecologists. *Journal of Statistical Software*, *22*(4), 1–20.

Duminil, J., Fineschi, S., Hampe, A., Jordano, P., Salvini, D., Vendramin, G. G., & Petit, R. J. (2007). Can population genetic structure be predicted from life-history traits? *The American Naturalist*, *169*(5), 662–672.

Durand, J. D., Persat, H., & Bouvet, Y. (1999). Phylogeography and postglacial dispersion of the chub (*Leuciscus cephalus*) in Europe. *Molecular Ecology*, *8*(6), 989–997. https://doi.org/10.1046/j.1365-294x.1999.00654.x

Dyer, R. J., Nason, J. D., & Garrick, R. C. (2010). Landscape modelling of gene flow: Improved power using conditional genetic distance derived from the topology of population networks. *Molecular Ecology*, *19*(17), 3746–3759. https://doi.org/10.1111/j.1365-294X.2010.04748.x

Fourtune, L., Paz-Vinas, I., Loot, G., Prunier, J. G., & Blanchet, S. (2016). Lessons from the fish: A multi-species analysis reveals common processes underlying similar species-genetic diversity correlations. *Freshwater Biology*, *61*(11), 1830–1845. https://doi.org/10.1111/fwb.12826

Fraser, C. I., Davies, I. D., Bryant, D., & Waters, J. M. (2018). How disturbance and dispersal influence intraspecific structure. *Journal of Ecology*, *106*, 1298–1306. https://doi.org/10.1111/1365-2745.12900

Froese, R., & Pauly, D. (2017). FishBase. World Wide Web electronic publication. Retrieved from www.fishbase.org, version (06/2017).

Gilbert, B., & Levine, J. M. (2017). Ecological drift and the distribution of species diversity. *Proceedings of the Royal Society B: Biological Sciences*, *284*(1855), 20170507. https://doi.org/10.1098/rspb.2017.0507

Goudet, J., & Jombart, T. (2015). hierfstat: Estimation and tests of hierarchical F-statistics. R package version 0.04-22. Retrieved from https://CRAN.R-project.org/package=hierfstat

Hahsler, M., Buchta, C., & Hornik, K. (2018). seriation: Infrastructure for ordering objects using seriation. R package version 1.2-3. Retrieved from https://CRAN.R-project.org/package=seriation.

Hedrick, P. W. (2005). A standardized genetic differentiation measure. *Evolution*, *59*(8), 1633–1638. https://doi.org/10.1111/j.0014-3820.2005.tb01814.x

Hugueny, B. (1989). West African rivers as biogeographic islands: Species richness of fish communities. *Oecologia*, *79*, 236–243. https://doi.org/10.1007/BF00388483

Kahilainen, A., Puurtinen, M., & Kotiaho, J. S. (2014). Conservation implications of species-genetic diversity correlations. *Global Ecology and Conservation*, *2*, 315–323. https://doi.org/10.1016/j.gecco.2014.10.013

Kelly, R. P., & Palumbi, S. R. (2010). Genetic structure among 50 species of the northeastern Pacific rocky intertidal community. *PLoS ONE*, *5*(1), e8594. https://doi.org/10.1371/journal.pone.0008594

Lamy, T., Laroche, F., David, P., Massol, F., & Jarne, P. (2017). The contribution of species-genetic diversity correlations to the understanding of community assembly rules. *Oikos*, *126*(6), 759–771. https://doi.org/10.1111/oik.03997

Larmuseau, M. H. D., Freyhof, J., Volckaert, F. A. M., & Van Houdt, J. K. J. (2009). Matrilinear phylogeography and demographical patterns of *Rutilus rutilus*: Implications for taxonomy and conservation. *Journal of Fish Biology*, *75*(2), 332–353. https://doi.org/10.1111/j.1095-8649.2009.02322.x

Laroche, F., Jarne, P., Lamy, T., David, P., & Massol, F. (2015). A neutral theory for interpreting correlations between species and genetic diversity in communities. *The American Naturalist*, *185*(1), 59–69. https://doi.org/10.1086/678990

Legendre, P., Lapointe, F.-J., & Casgrain, P. (1994). Modeling brain evolution from behavior: A permutational regression approach. *Evolution*, *48*(5), 1487–1499. https://doi.org/10.1111/j.1558-5646.1994.tb02191.x

Leprieur, F., Olden, J. D., Lek, S., & Brosse, S. (2009). Contrasting patterns and mechanisms of spatial turnover for native and exotic freshwater fish in Europe. *Journal of Biogeography*, *36*(10), 1899–1912. https://doi.org/10.1111/j.1365-2699.2009.02107.x

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*(4), 325–334. https://doi.org/10.1111/j.1461-0248.2011.01589.x

Lichstein, J. W. (2007). Multiple regression on distance matrices: A multivariate spatial analysis tool. *Plant Ecology*, *188*(2), 117–131. https://doi.org/10.1007/s11258-006-9126-3

Mayr, E. (1963). *Animal species and evolution*. Cambridge, MA: Harvard University Press. https://doi.org/10.4159/harvard.9780674865327

Meirmans, P. G. (2015). Seven common mistakes in population genetics and how to avoid them. *Molecular Ecology*, *24*, 3223–3231. https://doi.org/10.1111/mec.13243

Meirmans, P. G., & Hedrick, P. W. (2011). Assessing population structure: F_{ST} and related measures. *Molecular Ecology Resources*, *11*(1), 5–18. https://doi.org/10.1111/j.1755-0998.2010.02927.x

Moritz, C., Meynard, C. N., Devictor, V., Guizien, K., Labrune, C., Guarini, J. M., & Mouquet, N. (2013). Disentangling the role of connectivity, environmental filtering, and spatial structure on metacommunity dynamics. *Oikos*, *122*(10), 1401–1410.

Nei, M. (1973). Analysis of gene diversity in subdivided populations. *Proceedings of the National Academy of Sciences*, *70*(12), 3321–3323. https://doi.org/10.1073/pnas.70.12.3321

Nekola, J. C., & White, P. S. (1999). The distance decay of similarity in biogeography and ecology. *Journal of Biogeography*, *26*(4), 867–878. https://doi.org/10.1046/j.1365-2699.1999.00305.x

Neuwirth, E. (2014). RColorBrewer: ColorBrewer Palettes. R package version 1.1-2. Retrieved from https://CRAN.R-project.org/package=RColorBrewer.

Oberdorff, T., Tedesco, P. A., Hugueny, B., Leprieur, F., Beauchard, O., Brosse, S., & Dürr, H. H. (2011). Global and regional patterns in riverine fish species richness: A review. *International Journal of Ecology*, *12*, Article ID 967631, https://doi.org/10.1155/2011/967631

Oksanen, J., Blanchet, G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., … Wagner, H. (2018). vegan: Community ecology package. R package version 2.5-2. Retrieved from https://CRAN.R-project.org/package=vegan.

Orsini, L., Vanoverbeke, J., Swillen, I., Mergeay, J., & De Meester, L. (2013). Drivers of population genetic differentiation in the wild: Isolation by dispersal limitation, isolation by adaptation and isolation by colonization. *Molecular Ecology*, *22*(24), 5983–5999. https://doi.org/10.1111/mec.12561

Pagès, H., Aboyoun, P., Gentleman, R., & DebRoy, S. (2018). Biostrings: Efficient manipulation of biological strings. R package version 2.48.0.

Papadopoulou, A., Anastasiou, I., Spagopoulou, F., Stalimerou, M., Terzopoulou, S., Legakis, A., & Vogler, A. P. (2011). Testing the species–genetic diversity correlation in the Aegean Archipelago: Toward a haplotype-based macroecology? *The American Naturalist*, *178*(2), 241–255. https://doi.org/10.1086/660828

Paradis, E., Claude, J., & Strimmer, K. (2004). APE: Analyses of phylogenetics and evolution in R language. *Bioinformatics*, *20*, 289–290. https://doi.org/10.1093/bioinformatics/btg412

Preston, F. W. (1962). The canonical distribution of commonness and rarity: Part II. *Ecology*, *43*(3), 410–432. https://doi.org/10.2307/1933371

Prunier, J. G., Dubut, V., Chikhi, L., & Blanchet, S. (2017). Contribution of spatial heterogeneity in effective population sizes to the variance in pairwise measures of genetic differentiation. *Methods in Ecology and Evolution*, *8*(12), 1866–1877. https://doi.org/10.1111/2041-210X.12820

Qian, H., & Ricklefs, R. E. (2012). Disentangling the effects of geographic distance and environmental dissimilarity on global patterns of species turnover. *Global Ecology and Biogeography*, *21*(3), 341–351. https://doi.org/10.1111/j.1466-8238.2011.00672.x

R Core Team. (2018). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.R-project.org/.

Radinger, J., & Wolter, C. (2014). Patterns and predictors of fish dispersal in rivers. *Fish and Fisheries*, *15*(3), 456–473. https://doi.org/10.1111/faf.12028

Robinson, D., & Hayes, A. (2018). broom: Convert statistical analysis objects into tidy tibbles. R package version 0.5.0. Retrieved from https://CRAN.R-project.org/package=broom.

Rousset, F. (1997). Genetic differentiation and estimation of gene flow from F-statistics under isolation by distance. *Genetics*, *145*(4), 1219–1228.

Seifertová, M., Bryja, J., Vyskočilová, M., Martínková, N., & Šimková, A. (2012). Multiple Pleistocene refugia and post-glacial colonization in the European chub (*Squalius cephalus*) revealed by combined use of nuclear and mitochondrial markers. *Journal of Biogeography*, *39*(6), 1024–1040. https://doi.org/10.1111/j.1365-2699.2011.02661.x

Sepkoski, J. J., & Rex, M. A. (1974). Distribution of freshwater mussels: Coastal rivers as biogeographic islands. *Systematic Zoology*, *23*(2), 165–188. https://doi.org/10.2307/2412130

Serrouya, R., Paetkau, D., McLellan, B. N., Boutin, S., Campbell, M., & Jenkins, D. A. (2012). Population size and major valleys explain microsatellite variation better than taxonomic units for caribou in western Canada. *Molecular Ecology*, *21*(11), 2588–2601. https://doi.org/10.1111/j.1365-294X.2012.05570.x

Soininen, J., McDonald, R., & Hillebrand, H. (2007). The distance decay of similarity in ecological communities. *Ecography*, *30*(1), 3–12. https://doi.org/10.1111/j.0906-7590.2007.04817.x

Sørensen, T. J. (1948). A method of establishing groups of equal amplitude in plant sociology based on similarity of species content, and its application to analyses of the vegetation on Danish commons. *Kongelige Danske Videnskabernes Selskabs Biologiske Skrifter*, *5*, 1–34.

Tedesco, P. A., Beauchard, O., Bigorne, R., Blanchet, S., Buisson, L., Conti, L., … Oberdorff, T. (2017). A global database on freshwater fish species occurrence in drainage basins. *Scientific Data*, *4*, 170141. https://doi.org/10.1038/sdata.2017.141

Tedesco, P. A., Leprieur, F., Hugueny, B., Brosse, S., Dürr, H. H., Beauchard, O., … Oberdorff, T. (2012). Patterns and processes of global riverine fish endemism. *Global Ecology and Biogeography*, *21*(10), 977–987. https://doi.org/10.1111/j.1466-8238.2011.00749.x

Tisseuil, C., Cornu, J. F., Beauchard, O., Brosse, S., Darwall, W., Holland, R., … Oberdorff, T. (2013). Global diversity patterns and cross-taxa convergence in freshwater systems. *Journal of Animal Ecology*, *82*(2), 365–376. https://doi.org/10.1111/1365-2656.12018

Vellend, M. (2003). Island biogeography of genes and species. *The American Naturalist*, *162*(3), 358–365. https://doi.org/10.1086/377189

Vellend, M., & Geber, M. A. (2005). Connections between species diversity and genetic diversity. *Ecology Letters*, *8*(7), 767–781. https://doi.org/10.1111/j.1461-0248.2005.00775.x

Vellend, M., Lajoie, G., Bourret, A., Múrria, C., Kembel, S. W., & Garant, D. (2014). Drawing ecological inferences from coincident patterns of population- and community-level biodiversity. *Molecular Ecology*, *23*(12), 2890–2901. https://doi.org/10.1111/mec.12756

Vellend, M., & Orrock, J. L. (2009). Ecological and genetic models of diversity. In J. B. Losos & R. E. Ricklefs (Eds.), *The theory of Island biogeography revisited* (pp. 439–462). Princeton, NJ: Princeton University Press.

Wang, I. J., Glor, R. E., & Losos, J. B. (2013). Quantifying the roles of ecology and geography in spatial genetic divergence. *Ecology Letters*, *16*(2), 175–182. https://doi.org/10.1111/ele.12025

Whitlock, M. C. (2005). Combining probability from independent tests: The weighted Z-method is superior to Fisher's approach. *Journal of Evolutionary Biology*, *18*(5), 1368–1373. https://doi.org/10.1111/(ISSN)1420-9101

Whittaker, R. H. (1960). Vegetation of the Siskiyou mountains, Oregon and California. *Ecological Monographs*, *30*(3), 279–338. https://doi.org/10.2307/1943563

Whittaker, R. H. (1972). Evolution and measurement of species diversity. *Taxon*, 213–251. https://doi.org/10.2307/1218190

Wickham, H. (2007). Reshaping data with the reshape package. *Journal of Statistical Software*, *21*(12), 1–20.

Wickham, H. (2016). *ggplot2: Elegant graphics for data analysis*. New-York, NY: Springer-Verlag. https://doi.org/10.1007/978-3-319-24277-4

Winter, D. J. (2012). mmod: An R library for the calculation of population differentiation statistics. *Molecular Ecology Resources*, *12*(6), 1158–1160. https://doi.org/10.1111/j.1755-0998.2012.03174.x

Wright, S. (1943). Isolation by distance. *Genetics*, *28*(March), 114–138.

Wright, S. (1951). The genetical structure of natural populations. *Annals of Eugenics*, *15*(4), 323–354.

BIOSKETCH

The authors are particularly interested in documenting and explaining spatial patterns of biodiversity at large spatial scales. They try to disentangle the interplay between historical and contemporary factors (including anthropogenic pressures) in shaping patterns of alpha and beta diversities in aquatic populations and communities and value this knowledge to guide conservation strategies of aquatic ecosystems.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

How to cite this article: Robuchon M, Leroy B, Jézéquel C, Hugueny B. Correlations between broad-scale taxonomic and genetic differentiations suggest a dominant imprint of historical processes on beta diversities. *J Biogeogr.* 2019;00:1–13. https://doi.org/10.1111/jbi.13559