

# Lessons from the fish: a multi-species analysis reveals common processes underlying similar species-genetic diversity correlations

LISA FOURTUNE\*, IVAN PAZ-VINAS<sup>†</sup>, GÉRALDINE LOOT<sup>\*‡</sup>, JÉRÔME G. PRUNIER\* AND SIMON BLANCHET<sup>\*§</sup>

\*Centre National de la Recherche Scientifique (CNRS), Université Paul Sabatier (UPS), UMR 5321 (Station d'Écologie Théorique et Expérimentale), Moulis, France

<sup>†</sup>Aix-Marseille Université, CNRS, IRD, Avignon Université, UMR 7263 (IMBE), Équipe EGE, Centre Saint-Charles, Case 36, Marseille, France

<sup>‡</sup>Université de Toulouse, UPS, UMR 5174 (Laboratoire Évolution & Diversité Biologique, EDB), Toulouse, France

<sup>§</sup>CNRS, UPS, École Nationale de Formation Agronomique (ENFA), UMR 5174 EDB, Toulouse, France

## SUMMARY

1. Species-genetic diversity correlations (SGDCs) have been investigated over a large spectra of organisms, which has greatly improved our understanding of parallel processes potentially driving both species and genetic diversity. However, there are still few studies comparing SGDCs (and underlying processes) for multiple species sampled over a single landscape.
2. Here, focusing on freshwater fish sampled across a large river basin (the Garonne-Dordogne river basin, France), we combined a multi-species approach and causal analyses to (i) assess and compare both  $\alpha$ -SGDCs and  $\beta$ -SGDCs among species, and (ii) infer processes underlying  $\alpha$ -SGDCs and  $\beta$ -SGDCs. Genetic, intraspecific diversity was assessed for four sympatric fish (*Barbatula barbatula*, *Gobio occitaniae*, *Phoxinus phoxinus* and *Squalius cephalus*) using microsatellite markers. Species diversity was quantified as species richness using electric fishing, and environmental conditions were thoroughly described for 81 sites.
3. We found significant and moderate positive  $\alpha$ -SGDCs for all four fish species, whereas  $\beta$ -SGDCs were weaker in strength and positively significant for two of the four species. Causal analyses identified two common variables (geographical isolation and area of available habitats) underlying the  $\alpha$ -SGDC relationships. Although weak, we found that  $\beta$ -SGDC correlations related to a direct relationship between taxonomic and genomic differentiation, and to the common influence of the abiotic environment acting as a filter on both species and alleles.
4. Our study shows that similar ecological and evolutionary processes related to environmental filtering, migration, drift and colonisation history act for explaining both species and genetic diversity of fish communities.

*Keywords:* biodiversity, community, genetics, population, rivers

## Introduction

It has long been theoretically acknowledged that parallel processes (i.e. speciation/mutation, ecological/genetic drift, dispersal/gene flow, environmental filtering/natural selection) can shape spatial patterns of species diversity and intraspecific genetic diversity (Antonovics,

1976). However, until recently, empirical studies focusing on co-variation in species and intraspecific genetic diversity were rare, which left unanswered the question of whether or not parallel processes can actually shape spatial patterns of species and genetic diversity. To empirically address this issue, Vellend (2003) introduced the Species-Genetic Diversity Correlation concept

---

Correspondence: Lisa Fourtune & Simon Blanchet, Station d'Écologie Théorique et Expérimentale, UMR 5321, 2 route du CNRS, 09200 Moulis, France. E-mails: lisa.fourtune@gmail.com & simon.blanchet@sete.cnrs.fr

(SGDC), which quantifies the congruency in the distributions of species and genetic diversity. Since then, the empirical assessment of SGDCs has flourished (reviewed in Vellend *et al.*, 2014). These studies are of critical interest as they (i) test the theoretical statement that similar processes drive patterns of biodiversity at different levels, and (ii) indirectly test for the practical possibility of using one level of diversity as a surrogate for the other for setting conservation plans (He *et al.*, 2008). The general expectation is that species and genetic diversity should co-vary positively, hence producing positive SGDCs (Vellend, 2005). Although many SGDC studies confirm this expectation (Vellend *et al.*, 2014), there are still strong discrepancies in the strength and interpretation of SGDCs. Vellend *et al.* (2014) emphasise that understanding the bases of these variations should now be a research priority.

SGDC implies a correlative approach, which makes difficult explaining why SGDCs are positive, negative or null, given that 'correlation does not imply causation'. Several factors can lead to negative or null SGDCs between species and genetic diversity patterns, including opposing evolutionary forces (Derry *et al.*, 2009), factors acting at different temporal or spatial scales (Taberlet *et al.*, 2012), or different responses to static or changing environmental conditions (Puşcaş, Taberlet & Choler, 2008). Moreover, under the niche variation hypothesis (Van Valen, 1965), an increase in species diversity within a community may reduce the genetic diversity of some species through increased interspecific competition or reduction of the average intraspecific niche breadth (Xu *et al.*, 2016). Contrastingly, three non-exclusive hypotheses have been proposed to explain positive SGDCs (Vellend & Geber, 2005). First, species and genetic diversity responding similarly to environmental drivers could show a positive SGDC, such as geographical isolation acting on dispersal and gene flow or habitat area influencing ecological and genetic drift. Second, genetic diversity of one species may directly increase surrounding community species diversity (Vellend & Geber, 2005), by promoting community-level stability and reducing extinction risk (Saccheri *et al.*, 1998; Frankham, 2015). Third, genetic diversity might increase due to species diversity if increased species diversity generates diversifying selection on non-neutral genetic diversity (Vellend & Geber, 2005). Deciphering the relative (or combined) role of each three hypotheses from empirical data is still extremely challenging (Vellend *et al.*, 2014).

All hypotheses described above have been developed by considering correlations between the  $\alpha$  component of diversity (Loreau, 2000), i.e. between indices of

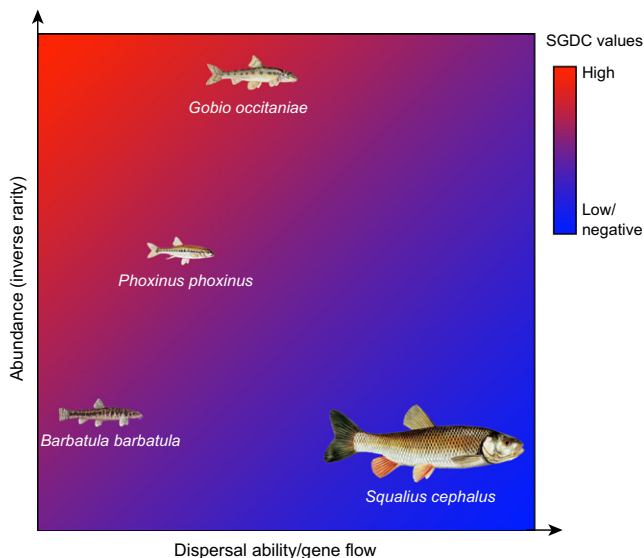
within-sites diversity such as species richness and allelic richness ( $\alpha$ -SGDC). However, recent studies also indicate that quantifying between-site diversity (i.e.  $\beta$ -diversity) is important in assessing SGDCs (e.g. Odat, Jetschke & Hellwig, 2004; Sei, Lang & Berg, 2009; Struebig *et al.*, 2011; Blum *et al.*, 2012).  $\beta$ -diversity measures differences among communities or populations across spatial or temporal scales, thus providing a complementary description of diversity and a more complete understanding of the ecological and evolutionary processes shaping it (Sei *et al.*, 2009; Sexton, Hangartner & Hoffmann, 2014).

Vellend (2005) theoretically demonstrated that SGDCs are strongly affected by the abundance of the species from which intraspecific genetic diversity is measured (hereafter, the 'target species'), with rarer species producing weaker SGDCs and *vice versa*. In the same vein, Laroche *et al.* (2015) demonstrated that the mutation-to-gene flow ratio of the target species also strongly affects SGDCs. More specifically, positive SGDCs were theoretically obtained when mutation rate was weak relative to gene flow, whereas SGDCs can be both positive and negative when mutation rate is not negligible (Laroche *et al.*, 2015). Although theoretical works suggest that the strength and sign of SGDCs may vary depending on the target species, most studies assess SGDCs by quantifying genetic diversity from a single species (e.g. Puşcaş *et al.*, 2008; Evanno *et al.*, 2009; Blum *et al.*, 2012). However, the peculiarities of each species can provide useful information on the evolutionary and ecological processes shaping diversity, given that life-history traits of species are partaking in shaping spatial patterns of genetic diversity and differentiation (e.g. Duminil *et al.*, 2007; Kelly & Palumbi, 2010; Mazé-Guilmo *et al.*, 2016). Studies focusing on comparisons of SGDCs between several species remain scarce (but see Struebig *et al.*, 2011; Taberlet *et al.*, 2012; Lamy *et al.*, 2013; Múrria *et al.*, 2015), although they generally provide key information on the processes underlying SGDCs.

Understanding the processes underlying SGDCs might be particularly tricky in spatially structured ecosystems (Blum *et al.*, 2012; Altermatt, 2013). This is the case for dendritic river networks (DRNs) in which the dispersal of individuals is highly constrained by the network spatial arrangement (Campbell Grant, Lowe & Fagan, 2007; Paz-Vinas & Blanchet, 2015). This is especially true for highly water-dependent organisms such as freshwater fish (Paz-Vinas *et al.*, 2015). Moreover, DRNs are highly heterogeneous landscapes, environmentally structured along upstream-downstream gradients (Vannote *et al.*, 1980). Critical environmental components, such as river

width and oxygen concentration, vary along stream gradients and are expected to impact the dynamics of species adaptation and selection, thereby having significant impact on SGDCs. These characteristics make the analysis of SGDCs in DRNs challenging, but also highly exciting.

Here, we measured freshwater fish species diversity and intraspecific genetic diversity for four freshwater fish species across the entirety of a river drainage to test (i) if, as expected, the strength and sign of SGDCs vary among the four target species, and (ii) if processes underlying SGDCs are similar among the four target species in a complex DRN. We focused on a set of four contrasting species (*Barbatula barbatula*, *Gobio occitaniae*, *Phoxinus phoxinus* and *Squalius cephalus*) varying in key biological traits, including rarity and dispersal ability. We predict that SGDCs should be weaker in the less abundant and more vagile species (*S. cephalus*), whereas SGDCs should be stronger for the more abundant and less vagile species (*G. occitaniae*) (see Fig. 1 for specific predictions). Regarding processes underlying SGDCs,



**Fig. 1** Theoretical biplot showing how the strength of Species-genetic diversity correlations (SGDCs) is expected to vary according to the rarity (*y*-axis) and the gene flow (*x*-axis) of the target species (i.e. the one used to quantify genetic diversity) and following theoretical works by Vellend (2005) and Laroche *et al.* (2015). We placed on the biplot a picture of each species, which corresponds (based on scientific and our own field knowledge) to their respective level of rarity and dispersal ability (see the main text for details). This allows us to predict that the strength of SGDCs should be different among these four species, and should follow this ranking (from high to low SGDCs):  $SGDCS_{Gobio\ occitaniae} > SGDCS_{Phoxinus\ phoxinus} > SGDCS_{Barbatula\ barbatula} > SGDCS_{Squalius\ cephalus}$ .

we expect processes related to the colonisation history of species and populations to have strong and common influences on both genetic and species diversity (Blanchet *et al.*, 2014; Paz-Vinas *et al.*, 2015), and hence on SGDCs. On the contrary, local abiotic parameters such as the level of oxygen concentration or water temperature may have stronger effects on species richness than on genetic diversity, since these parameters have been shown to strongly drive the spatial distribution of freshwater species (Buisson, Blanc & Grenouillet, 2008; Blanchet *et al.*, 2014). To test these predictions, we first assessed and compared  $\alpha$ -SGDCs and  $\beta$ -SGDCs among the four species at 81 sites covering an entire river drainage. Then, to disentangle the hypotheses proposed by Vellend & Geber (2005), we compiled a detailed environmental and geographical database describing each sampling site, and we applied path analyses (Tenenhaus *et al.*, 2005) for all four target species on both  $\alpha$ -SGDCs and  $\beta$ -SGDCs. Path analysis is to our knowledge the most appropriate statistical tool to unravel complex relationships within a set of variables derived from empirical data (Wright, 1921). In a companion article, Seymour *et al.* (2016) provides complementary findings of patterns and processes of SGDC in DRNs, but focusing on a taxonomic group (invertebrates) whose dispersal is not restricted to water corridors.

## Methods

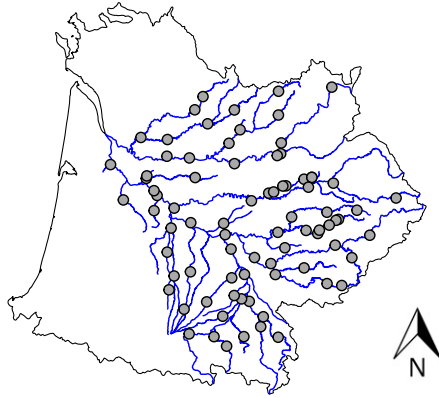
### Study area

We focused on the Garonne-Dordogne river drainage (South-Western France) that covers an area of 79 800 km<sup>2</sup>. We selected 81 sampling sites evenly scattered across the whole river basin according to the following criteria: (i) sites should be accessible for electric fishing, (ii) they should host as much of the four target species used for genetic analyses as possible, (iii) taxonomic data should be available and (iv) sites should cover most of the area of the river basin, and hence most of the environmental variability existing along the upstream–downstream river gradient (see Fig. 2a and Table S1 in Supporting Information).

### Species diversity

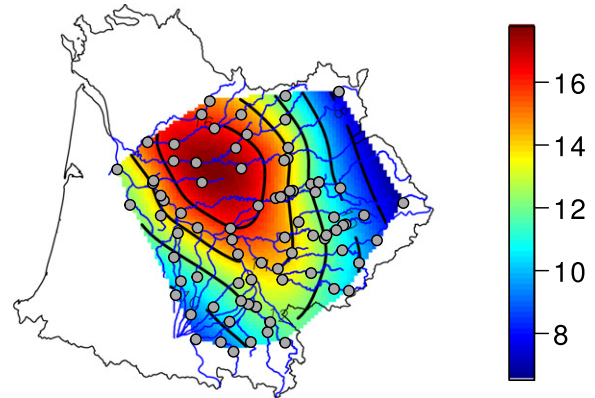
We collected data on the occurrence of freshwater fish species for all 81 sampling sites using a database provided by the 'Office National de l'Eau et des Milieux Aquatiques' (ONEMA; French freshwater agency). The ONEMA yearly monitors fish assemblages in more than

(a) Sampled sites

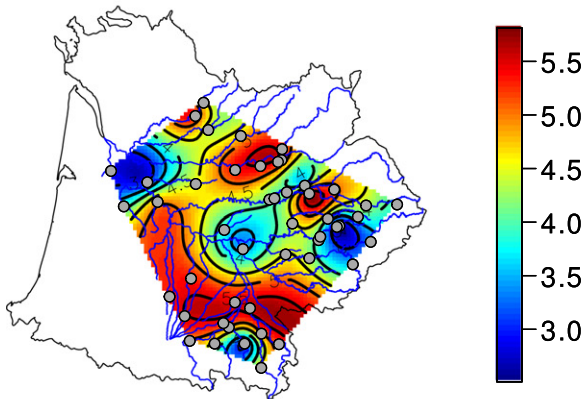


0 25 50 100 150 200 Kilometers

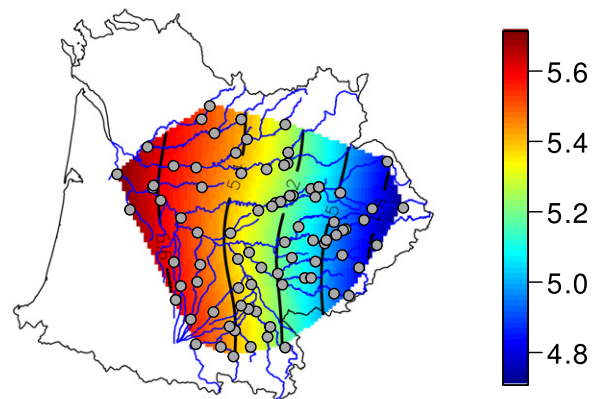
(b) Species richness



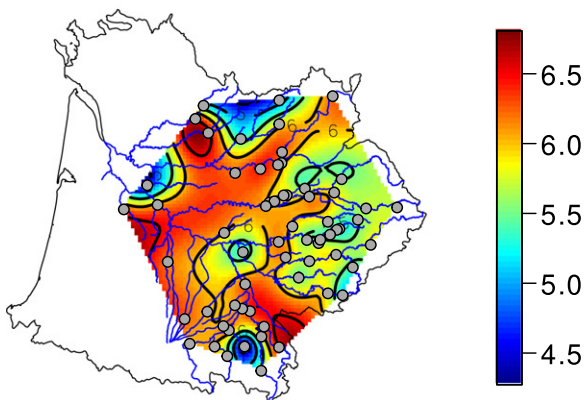
(c) *Barbatula barbatula*



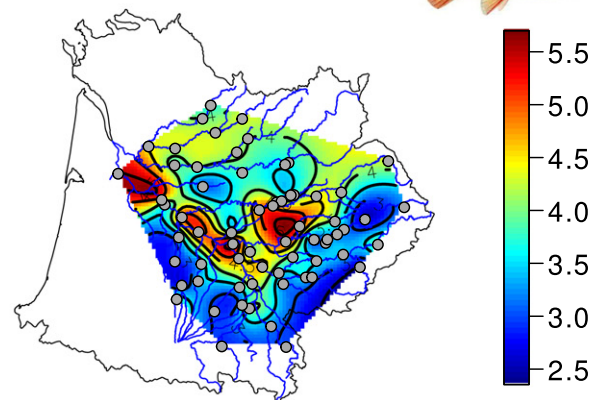
(d) *Gobio occitaniae*



(e) *Phoxinus phoxinus*



(f) *Squalius cephalus*



**Fig. 2** Maps representing the spatial distribution of interpolated species richness (a) and allelic richness of *Barbatula barbatula* (b), *Gobio occitaniae* (c), *Phoxinus phoxinus* (d) and *Squalius cephalus* (e). Red colour represents high richness and blue colour represents low richness. Each grey dot is a sampling site in which the data (species or allelic richness) were available. Interpolated values of  $\beta$ -diversity indices (i.e. true diversity and Jost's  $D$ ) could not be computed because  $\beta$ -diversity indices take the form of pairwise matrices.



1500 sampling sites in the Garonne-Dordogne river catchment (Poulet, Beaulaton & Dembski, 2011), feeding a database gathering presence/absence data for all fish species found in the river basin. These data were collected during electric-fishing campaigns from 1975 to 2011. To describe the taxonomic diversity of each sampling site, we considered data from every yearly sample by site to make the species list as exhaustive as possible. This led to a regional pool of 51 species (see Table S2 for details). It is noteworthy that this species list included both native and non-native species (17 non-native species). We ran additional analyses without the non-native species, which led to very similar results (see Table S3 for results without the non-native species).

Species  $\alpha$ -diversity was quantified as the species richness (i.e. number of species per site), and  $\beta$ -diversity was quantified as the pairwise community dissimilarity using the Jost's index of 'true diversity' (Jost, 2006), so as to make it directly comparable to genetic  $\beta$ -diversity (see below). This metric measures the community variation among pairs of sites and ranges from 1 (identical communities) to 2 (completely distinct communities) (Jost, 2006). The values were computed using the R package 'simba' (Jurasiński & Retzer, 2012).

### Genetic diversity

Intraspecific genetic diversity was estimated from four cypriniform fish: *Gobio occitaniae*, *Phoxinus phoxinus* and *Squalius cephalus* (Cyprinidae) and *Barbatula barbatula* (Nemacheilidae). We chose species that are of limited interest for anglers, so as to limit the possibility for past stocking events and uninformed translocations between river drainages. We performed preliminary 'outlier population' analyses in our genetic datasets through Factorial Correspondence Analysis using the GENETIX software (Belkhir *et al.*, 1996) and ensured the absence of recent stocking events by asking local angling associations, hence further reducing this risk. Although all these species are mainly insectivorous, they strongly differ in their foraging mode, with *S. cephalus* and *P. phoxinus* feeding in the water column, whereas *B. barbatula* and *G. occitaniae* feed preferentially on the bottom. Moreover, these species vary in their level of habitat specialisation, with *G. occitaniae* being the most generalist (i.e. it is found almost everywhere in the river basin and in many habitat types) and abundant species, whereas *B. barbatula* is a specialist species living in very specific habitats (mainly riffles in the midstream sections of rivers) at relatively low abundance. *S. cephalus* and *P. phoxinus* are intermediate species; the former is

primarily found in downstream sections at relatively low densities, whereas the latter is found in upstream sections at relatively high densities (Keith *et al.*, 2011). Additionally, all four fish species strikingly vary in their mean body length, which is often related to dispersal ability in fish (Radinger & Wolter, 2014): the largest is *S. cephalus* (300–500 mm) followed by *G. occitaniae* (120–150 mm), *B. barbatula* (100–120 mm) and *P. phoxinus* (80–90 mm) (Keith *et al.*, 2011). Due to its body shape, *B. barbatula* is considered a poor swimmer, and thus disperser, whereas *S. cephalus* is expected to have the higher dispersal ability because of its large streamline body length. *Gobio occitaniae* and *P. phoxinus* have intermediate dispersal ability. Variability observed for dispersal ability and rarity in these four species covered a non-negligible (although not complete) part of the trait space of the 51 species found in the basin. We summarised in Fig. 1 the rarity (based on abundance and level of habitat specialisation) and dispersal ability of each species, which led to theoretical predictions regarding the strength of SGDCs expected for each species.

Specimens were sampled once across each of the 81 sites using electric fishing between the summer of 2010 and 2011. Average site area was 500–1000 m<sup>2</sup> to ensure the sampling included the full range of local habitat heterogeneity. We sought to capture up to 25 individuals per species per site, although (i) not all species were found in all sites and (ii) not all sites provided 25 individuals due to low-site abundances. Sites in which less than 10 individuals were successfully genotyped were removed from the database, leading to 42 sites for *B. barbatula*, 74 sites for *G. occitaniae*, 54 sites for *P. phoxinus* and 60 sites for *S. cephalus* (Fig. 2), with a total of 5405 individual genotypes (see Table 1 for details). For each individual, a small piece of pelvic fin was collected and preserved in 70% ethanol. DNA was extracted using a salt extraction protocol (Aljanabi & Martinez, 1997) and individuals were genotyped for eight to ten microsatellite loci [*B. barbatula* ( $n = 9$ ); *G. occitaniae* ( $n = 8$ ); *P. phoxinus* ( $n = 10$ ); *S. cephalus* ( $n = 10$ )]. We used 5–20 ng of genomic DNA and QIAGEN<sup>®</sup> Multiplex PCR Kits (Qiagen, Valencia) to perform PCR amplifications. We provide more details on loci, primer concentrations, PCR conditions and multiplex recipes in Table S4. The genotyping was conducted on an ABI PRISM<sup>™</sup> 3730 Automated Capillary Sequencer (Applied Biosystems, Foster City), and the scoring of allele sizes was done using GENEMAPPER<sup>®</sup> v.4.0 (Applied Biosystems).

We determined the occurrence of null alleles and potential scoring errors with the program

**Table 1** (a) Number of sites sampled, total number of species over all sites and mean and range of species richness and true diversity; (b) number of sites and individuals sampled for all four species and mean and range of allelic richness and Jost's *D*.

(a) Species diversity	Number of sites	Total number of species	Species richness	True diversity
Freshwater fish	81	51	15.100 [4; 31]	1.302 [1; 2]
(b) Genetic diversity	Number of sites	Number of individuals	Allelic richness	Jost's <i>D</i>
<i>Barbatula barbatula</i>	42	911	4.427 [2.639; 5.595]	0.383 [−0.010; 0.880]
<i>Gobio occitaniae</i>	74	1770	5.288 [4.047; 6.240]	0.200 [−0.031; 0.627]
<i>Phoxinus phoxinus</i>	54	1324	5.830 [4.258; 6.631]	0.267 [−0.016; 0.781]
<i>Squalius cephalus</i>	60	1400	3.821 [2.477; 5.638]	0.164 [−0.023; 0.526]

MICROCHECKER 2.3 (Van Oosterhout *et al.*, 2004). We tested for departures from Hardy–Weinberg (HW) equilibrium among loci within sites using the R package 'adegenet' v1.4-2 (Jombart, 2008). The program GENEPOP v4.0 (Rousset, 2008) was used to assess linkage disequilibrium (LD) among loci within sites.

Genetic  $\alpha$ -diversity was measured as mean allelic richness using ADZE v1.0 (Szpiech, Jakobsson & Rosenberg, 2008). Beta-diversity was quantified as the pairwise genetic differentiation (between sites) using Jost's *D* (Jost, 2008) using the R package 'mmod' (Winter, 2012). Jost's *D* values range from 0 (or a slightly negative value) when there is all alleles present in both populations (no differentiation) to 1 when there is no alleles in common between populations (strong differentiation). Compared to classical  $F_{ST}$  or  $G_{ST}$  estimates, Jost's *D* is preferred as it is independent from the number of alleles sampled in the population, which allows for a better comparison of inter-specific genetic differentiation (Jost, 2008).

Finally, spatial distributions of species and genetic  $\alpha$ -diversity across the Garonne-Dordogne were described using krigged maps of species and allelic richness using the R package 'fields' (Nychka, Furrer & Sain, 2015).

#### Environmental and geographical data

Environmental variables were collected for each sampling site, including elevation and river width obtained from the French Theoretical Hydrological Network ('Réseau Hydrologique Théorique français', RHT; Pella *et al.*, 2012), along with water temperature (°C), oxygen concentration ( $\text{mg L}^{-1}$ ) and saturation (%), suspended matter ( $\text{mg L}^{-1}$ ), pH and conductivity ( $\text{mS cm}^{-1}$ ) obtained from the database of the Water Information System of the Adour Garonne basin (SIEAG, 'Système d'Information sur l'Eau du Bassin Adour Garonne'; <http://adour-garonne.eaufrance.fr>). We selected environmental variables that were related to key life-history traits of fish species or integrative of many ecosystemic

processes. The SIEAG database gathers chemical characteristics of surface water measured several times a year at numerous sites in the river catchment. Only sites where data were available for March, May, July, September and November of 2011 were selected from the SIEAG database. The mean of these five values was calculated to inform the chemical quality of the sites. When sampling location did not overlap perfectly with SIEAG data, the nearest SIEAG site, along the same river and within a 5 km radius, was used. Three sampling sites did not have proximal SIEAG sites to gather reliable information and were hence removed from subsequent analyses. The betweenness centrality value of the first node (i.e. river confluence) upstream of each site was estimated using NetworkX (Hagberg, Schult & Swart, 2008). Betweenness centrality is an index of river connectivity that quantifies the positional importance of a node within a network (Freeman, 1977; Estrada & Bodin, 2008). Topological distance from the outlet (i.e. distance along the river network between a site and the basin outlet) and topological distance between each pair of sites (i.e. distance along the river network between two sites) were computed using QuantumGIS software (QGIS; Quantum GIS Development Team 2015).

#### Statistical analysis

To quantify and compare the relationships between species and genetic  $\alpha$ -diversity indices (i.e.  $\alpha$ -SGDCs), we computed Spearman's rank correlations for each species independently. To quantify and compare the relationships between species and genetic  $\beta$ -diversity indices (i.e.  $\beta$ -SGDCs), and because these data are pairwise matrices, we used Mantel tests with 10 000 permutations.

Partial least squares path modelling (PLS-PM; Tenenhaus *et al.*, 2005) was used to unravel the relationships between environmental and geographical variables and  $\alpha$ -diversity indices. This method allows simultaneously testing the significance, sign and strength of multiple

paths (i.e. relationships, see Tenenhaus *et al.*, 2005) defined within a dataset. For each path, a coefficient is computed to provide the strength and sign of the relationship, as well as a confidence interval using a bootstrap method. We designed a model containing the paths needed to test the hypotheses formulated by Vellend & Geber (2005). As stated above, the first hypothesis stipulates that genetic and species diversity are correlated because of similar effects of environment features. To test this hypothesis, environmental variables were grouped into five latent variables (i.e. variables representing concepts that cannot be measured but are built from several measured variables, see Tenenhaus *et al.*, 2005). Each latent variable corresponded to local characteristics that might influence both levels of  $\alpha$ -diversity: Isolation (altitude and distance from the outlet), Connectivity (betweenness), Area (river width), Oxygen (oxygen concentration and saturation) and Water composition (suspended matter, pH, water temperature and conductivity) (see Table 2 for general expectations and processes linking these variables to genetic and species diversity). We constructed a model in which these five environmental variables were directly linked to both levels of  $\alpha$ -diversity (see Fig. 5 for an illustration). As some of these variables are expected to covary spatially, we included paths between them when needed (see Table S5 for the values of correlation between environmental latent variables). The second and third hypotheses stipulate that one level of diversity influences the other. We tested these two hypotheses by adding a path between allelic richness and species richness. Because of PLS-PM methodological

limitations, we could not determine the direction of the path (i.e. from allelic richness to species richness, hypothesis 2; or from species richness to allelic richness, hypothesis 3), it was thus impossible to statistically decipher between these two hypotheses and they were hence grouped into a single hypothesis linking genetic and species diversity, irrespectively of the direction of the link. We used the allelic richness of the four species as a response variable in four independent models.

We conducted a similar analysis on species and genetic  $\beta$ -diversity data. The design of the models was the same, but species and allelic richness were replaced by Jost's  $D$  and true diversity, respectively. Environmental data were converted into pairwise environmental differences between sites, with the exception of the difference in altitude which was computed as the cumulative difference in altitude between sites along the river network (e.g. the total vertical drop to cover from one site to another). The definitions of the latent variables were unchanged with the exception of Isolation, which was defined as the cumulative difference in altitude and the topological distance between sites. The values of correlation between the environmental latent variables are given in Table S5. As the PLS-PM approach does not account for the non-independence of pairwise data, we implemented a specific bootstrap procedure, which corrects for the non-independence of pairwise data, as executed in traditional Mantel tests (Legendre & Fortin, 2010). Jost's  $D$  of the four species was implemented in four independent models to detect similarity and contrast between species. These analyses were conducted using the R package 'plspm' (Sanchez, Trinchera & Russolillo, 2015).

**Table 2** General predictions (and underlying processes) regarding the relationships between each latent variable and genetic and species diversity, respectively. These predictions arise from general and local knowledge on the influence of each variable on genetic (Paz-Vinas *et al.*, 2015) and species (Buisson *et al.*, 2008; Blanchet *et al.*, 2014) diversity.

Latent variables	Expected influence on genetic diversity	Expected influence on species diversity
Isolation	Isolated populations should have been colonised later and should harbour lower genetic diversity	Isolated communities should have been colonised later and should be taxonomically depauperated
Connectivity	Connectivity should increase gene flow and hence genetic diversity (while reducing genetic differentiation)	Connectivity should increase dispersal among communities, which should lead to richer communities with less differentiation
Area	Larger areas could translate into higher effective population size and hence genetic diversity	Due to the species-area relationship, higher species diversity is expected in sites with larger areas
Oxygen	Oxygen is a strong limiting factors for many fish species, low oxygen concentrations could translate into low effective population sizes and hence low levels of genetic diversity	Oxygen is strongly driving the spatial distribution of fish species
Water composition	Pollution can have strong effects on the effective population size (and hence genetic diversity), most notably in poorly tolerant species such as <i>Phoxinus phoxinus</i>	Pollution can have strong effects on the persistence of some species, polluted sites have generally lower species richness

Using the values of path coefficients computed from each of these models, we investigated how the magnitudes of the mechanisms driving  $\alpha$ -SGDCs and  $\beta$ -SGDCs (i.e. the effect sizes, ES; Nakagawa & Cuthill, 2007) differed between species. We then used a meta-analytic approach to test which of the mechanisms stood out as common for all species. Path coefficients values were treated as correlation coefficients from which we computed the Fisher's Z effect size (and its associated variance) for each coefficient and each species (see Rosenberg, Adams & Gurevitch, 2000; Nakagawa & Cuthill, 2007 for formulas). For each path of the models shown in Figs 5 and 6 and for each SGDC type independently ( $\alpha$ -SGDC and  $\beta$ -SGDC), we then computed the cumulative effect size ( $\bar{E}$ ) and the 95% confidence interval for all pooled species. This procedure allows accounting for the differences in sampling sizes between species when estimating a mean effect size (see Rosenberg *et al.*, 2000), and to test if a path was significant over all species. Finally, we evaluated if the set of effect sizes used to calculate  $\bar{E}$  for each path was homogeneous among species or not. The total heterogeneity of a sample ( $Q_t$ ) was calculated (Rosenberg *et al.*, 2000), and its significance was tested using chi-square statistics.

## Results

### Genetic diversity

We found significant deviations from HW, after Bonferroni corrections, for a few locus/population pairs for each species. However, these departures were not consistently clustered among loci or populations for any species (Table S6). We also detected significant LD between a few pairs of loci after Bonferroni corrections for some populations for each species, but no clear patterns appeared for any species across populations (Table S6). We did not find evidence for scoring errors due to stuttering or large allele dropouts in our datasets. Given the small extent and random nature of these deviations (HW and LD), and given the size of the databases, we retained all loci for the subsequent analyses.

### Species-genetic diversity correlations

Mean allelic richness ranged from  $3.82 \pm 0.77$  SD (*S. cephalus*) to  $5.83 \pm 0.48$  SD (*P. phoxinus*). The spatial distribution of allelic richness strikingly varied among species (Fig. 2). For instance, in *G. occitaniae*, allelic richness was higher downstream, whereas in *S. cephalus*, the highest values of allelic richness were found on main

river stretches, near confluence zones. No clear pattern was detected for *B. barbatula* and *P. phoxinus* (Fig. 2). The mean pairwise Jost's *D* ranged from  $0.16 \pm 0.09$  SD (*S. cephalus*) to  $0.38 \pm 0.17$  SD (*B. barbatula*). Additional genetic  $\alpha$ -diversity and  $\beta$ -diversity indices for each species are provided in Table 1. Mean species richness was  $12.77 \pm 4.91$  SD species and mean true diversity was  $1.28 \pm 0.12$  SD (Table 1).

We found significant and positive  $\alpha$ -SGDCs for each species (Fig. 3). However, correlation coefficients were relatively weak, and contrary to our expectations (Fig. 1), similar among species, ranging from 0.29 to 0.38 (Fig. 3). Similar trends were detected for  $\beta$ -diversity since correlation coefficients between true diversity and Jost's *D* ( $\beta$ -SGDCs) were weak and of similar strength for all species, although significant and positive  $\beta$ -SGDCs were observed for *B. barbatula* and *G. occitaniae* but not for *P. phoxinus* and *S. cephalus* (Fig. 4).

### Meta-analytic approach on partial least squares path modelling

For all species, two variables significantly affected species and genetic  $\alpha$ -diversity. First, the isolation of a site (i.e. altitude and distance from the outlet) was negatively related to both levels of  $\alpha$ -diversity (Fig. 5; Table 3). The magnitudes of isolation effects on species  $\alpha$ -diversity and genetic  $\alpha$ -diversity were significantly different from zero (Table 3). The effect of isolation on genetic  $\alpha$ -diversity was significantly heterogeneous among species ( $Q_t = 19.86$ , d.f. = 3,  $P = 0.001$ ), mainly because the effect of isolation on genetic  $\alpha$ -diversity was stronger in *G. occitaniae* (ES =  $-0.91$ ) than in any other species (ESs  $\leq -0.29$ , see Table S7 for details). Second, the cumulative effect of available area on species  $\alpha$ -diversity and genetic  $\alpha$ -diversity species was significantly positive for all species (Table 3) and homogeneous among species ( $Q_t = 2.93$ , d.f. = 3,  $P = 0.40$  and  $Q_t = 6.67$ , d.f. = 3,  $P = 0.08$ , respectively). We found neither evidence for additional environmental effects on  $\alpha$ -diversity, nor significant relationships between species richness and allelic richness (Table 3, Fig. 5).

Considering  $\beta$ -diversity, we found significant effect sizes between Jost's *D* and true diversity for all species (Table 3). However, we were not able to determine the direction of the arrow (i.e. from Jost's *D* to true diversity or from true diversity to Jost's *D*) because of PLS-PM methodological limitations. This effect was heterogeneous among species ( $Q_t = 10.25$ , d.f. = 3,  $P = 0.02$ ), with higher values for *B. barbatula* and *G. occitaniae* (ESs = 0.11 and 0.12, respectively) than for *P. phoxinus*



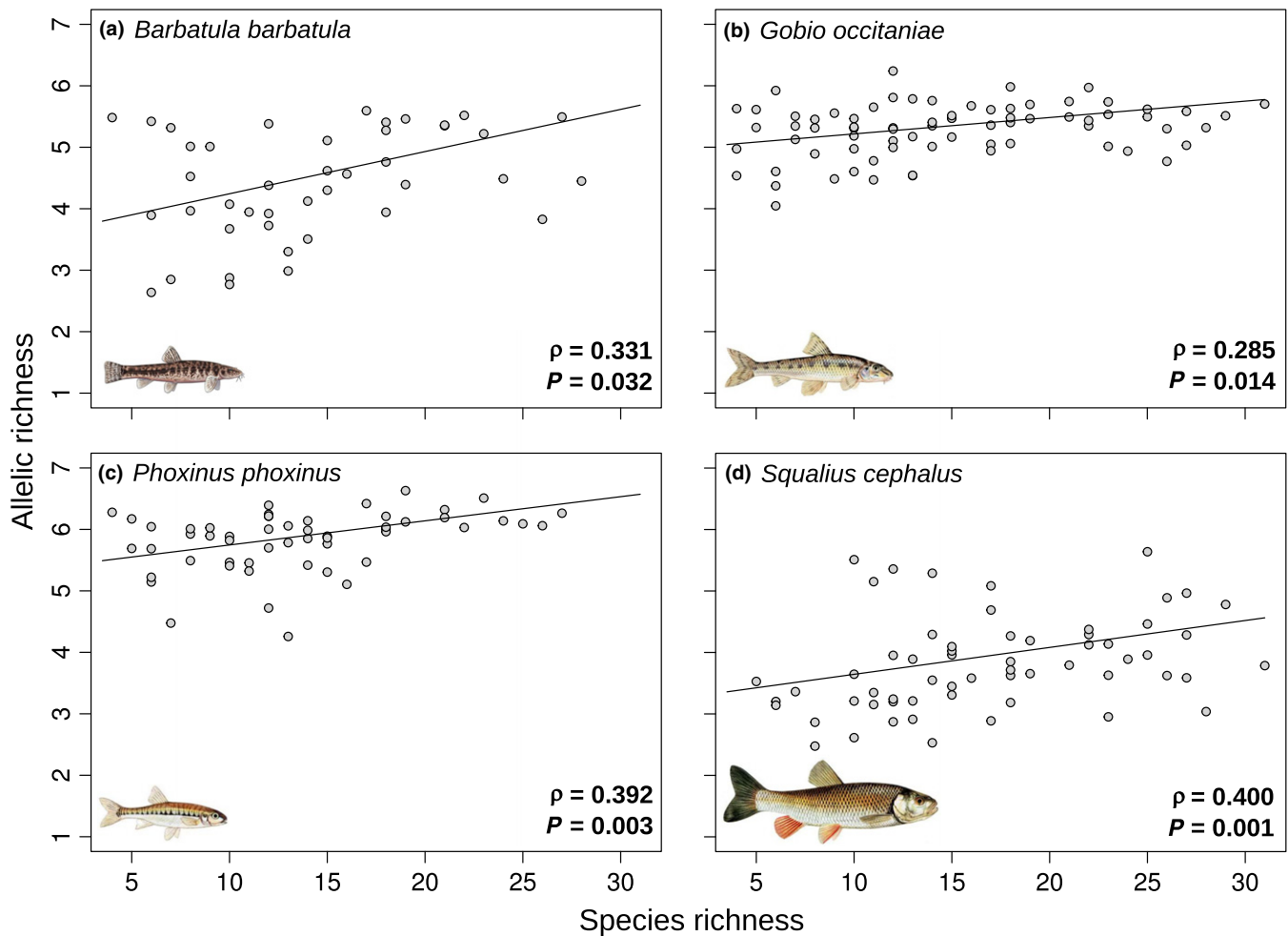


Fig. 3 Allelic richness (genetic  $\alpha$ -diversity) of *Barbatula barbatula* (a), *Gobio occitaniae* (b), *Phoxinus phoxinus* (c) and *Squalius cephalus* (d) plotted against species richness (species  $\alpha$ -diversity) with Spearman's rho and associated  $P$ -values.

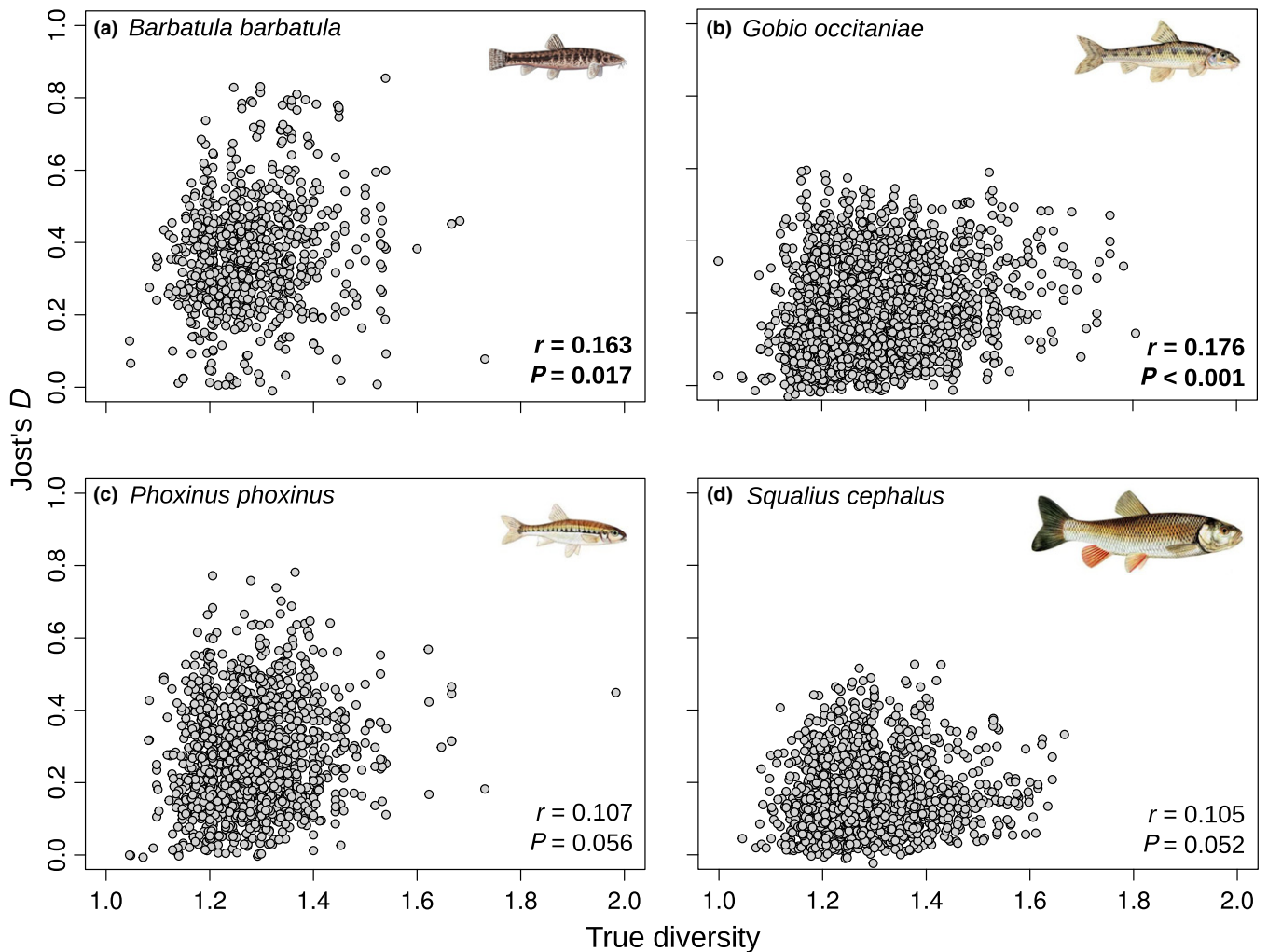
and *S. cephalus* (ESs = 0.02 and 0.05, respectively). Further, the difference in water composition between sites was significantly and positively related to both levels of  $\beta$ -diversity for all species (Table 3, Fig. 6). The isolation and area of sites were drivers of genetic  $\beta$ -diversity (Table 3), although their effects were heterogeneous among species (both  $P < 0.001$ ). Finally, difference in water oxygenation had a negative effect on species  $\beta$ -diversity (Table 3); however, this effect was heterogeneous among species ( $Q_t = 15.15$ , d.f. = 3,  $P = 0.002$ ).

## Discussion

### $\alpha$ -SGDCs in riverine fish species

Contrary to our theoretical expectations (Fig. 1),  $\alpha$ -SGDCs were significantly positive and similar in statistical strength for all four target species. Most previous

empirical studies involving multi-species approaches have highlighted the importance of high inter-specific variability in sign and strength of  $\alpha$ -SGDCs among species (Struebig *et al.*, 2011; Taberlet *et al.*, 2012 but see Lamy *et al.*, 2013). Additionally, high inter-specific variability in  $\alpha$ -SGDCs has recently been theoretically explained (Laroche *et al.*, 2015). However, our results are consistent with previous findings showing that similar positive  $\alpha$ -SGDCs are detected when species and genetic  $\alpha$ -diversity are extracted from ecologically similar taxonomic groups (He & Lamont, 2010; Seymour *et al.*, 2016). For instance, in a companion paper involving invertebrates across a large river basin, Seymour *et al.* (2016) found that  $\alpha$ -SGDC was stronger when intraspecific genetic  $\alpha$ -diversity of *Gammarus* sp. was compared to species  $\alpha$ -diversity of Amphipoda than to other invertebrate taxa, such as Ephemeroptera, Plecoptera or Trichoptera. This suggests that the four

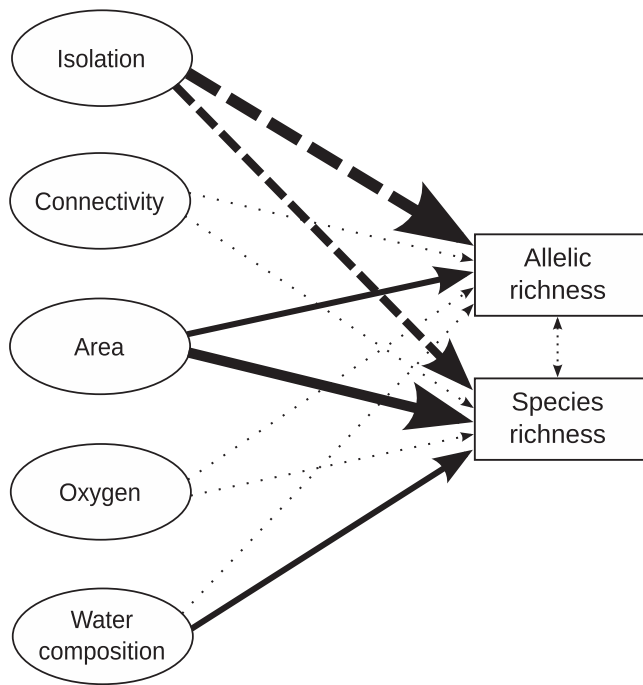


**Fig. 4** Jost's  $D$  (genetic  $\beta$ -diversity) of *Barbatula barbatula* (a), *Gobio occitaniae* (b), *Phoxinus phoxinus* (c) and *Squalius cephalus* (d) plotted against species richness (species  $\beta$ -diversity) with Mantel's  $r$  and associated  $P$ -values.

target species we used to estimate genetic  $\alpha$ -diversity may not be different enough to generate  $\alpha$ -SGDCs of different strength or sign. This is rather surprising given that these species occupy very different ecological niches and strongly vary in abundance and for many biological traits (Keith *et al.*, 2011), with some of these traits having previously been shown to affect the strength and sign of SGDCs (Vellend, 2005; Laroche *et al.*, 2015). Nonetheless, the use of path analysis in our study, combined with a meta-analytic approach, demonstrated that such similar  $\alpha$ -SGDCs arose from common mechanisms involving geographical isolation and the area of the sampled habitat.

We demonstrated that hypotheses linking species and genetic richness through a causal relationship were discarded for all species. On the contrary, we found that similar effects of environmental variables on both levels

of  $\alpha$ -diversity (hypothesis 1 of Vellend & Geber, 2005) were the main drivers of positive  $\alpha$ -SGDCs. First, the geographical isolation of sites negatively impacted both levels of  $\alpha$ -diversity, such that species and allelic richness tended to be lower in high altitude sites, located far away from the outlet. These negative effects of isolation on  $\alpha$ -diversity are consistent with previous findings (Buisson *et al.*, 2008; Paz-Vinas *et al.*, 2015), and may be explained with two non-mutually exclusive mechanisms. First, sites at low altitude and geographically close to river outlets are expected to experience greater immigration (i.e. sink populations; Gotelli & Taylor, 1999; Cadotte, 2006; Paz-Vinas *et al.*, 2015), which might provide both new species and novel alleles to these sites, thus counteracting the effect of drift and increasing  $\alpha$ -diversity (Vellend & Geber, 2005). Second, high altitude and increased distance from river outlets can reflect



**Fig. 5** Graphical representation of the meta-analysis results obtained from partial least square path modelling on  $\alpha$ -diversity indices over all species. Hatched arrows correspond to significant negative mean effect sizes and plain arrows correspond to significant positive mean effect sizes; their size is proportional to the absolute value of their mean effect sizes. Thin dotted arrows represent non-significant mean effect sizes. Variable names enclosed in squares depict observed variables while variable names enclosed in circles depict latent variables. For the sake of clarity, the paths linking environmental variables are not shown.

colonisation gradients, resulting from historical glacial refugia (Costedoat & Gilles, 2009; Blanchet *et al.*, 2014; Paz-Vinas *et al.*, 2015), which likely contributes to species and genetic  $\alpha$ -diversity gradients. For instance, Paz-Vinas *et al.* (2015) demonstrated that the downstream increase in genetic  $\alpha$ -diversity generally observed in rivers was mainly due to past colonisation pathways from downstream to upstream in many freshwater organisms. Additionally, we found that site area (i.e. the river width) positively impacted both levels of  $\alpha$ -diversity. Sites of higher river width are expected to sustain high fish density, thus increasing (i) the effective population size of a species, which is known to correlate positively with genetic  $\alpha$ -diversity (Frankham, 1996; Raeymaekers *et al.*, 2008) and (ii) the number of species that can cohabit a site (Jackson, Peres-Neto & Olden, 2001). Furthermore, large stretches of rivers are expected to display greater environmental heterogeneity, which might increase species and genetic  $\alpha$ -diversity through diversifying selection (Vellend & Geber, 2005).

#### $\beta$ -SGDCs in riverine fish species

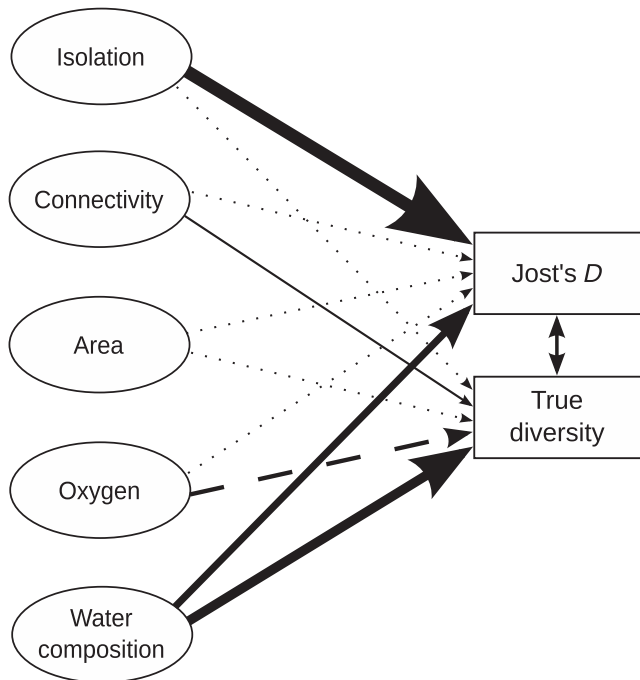
$\beta$ -SGDCs were weaker in strength than  $\alpha$ -SGDCs, although positive and within the same order of magnitude among species. This result was unexpected since most SGDCs studies investigating both  $\alpha$ - and  $\beta$ -diversity found, in average, stronger  $\beta$ -SGDCs than  $\alpha$ -SGDCs (see, for example, Odat *et al.*, 2004; Sei *et al.*, 2009).

We provided strong evidence that mechanisms driving  $\beta$ -SGDCs differ to some extent from those driving  $\alpha$ -SGDCs. We found evidence that hypotheses linking genetic differentiation to species differentiation could not be discarded. Although we were not able to tease apart the direction (genetic to taxonomy or taxonomy to genetic) of the relationship between species and genetic  $\beta$ -diversity, two non-exclusive mechanisms can be considered. Species  $\beta$ -diversity may influence genetic  $\beta$ -diversity if taxonomically distinct communities generate differential selective pressures acting on local populations, hence favouring different genotypes (Vellend & Geber, 2005). Given that we focused on neutral genetic markers, therein observing evidence of neutral population divergence, this first explanation is unlikely to explain our results. It is now well acknowledged that genetically differentiated populations, even at small spatial scales, can influence differential structure among communities and ecosystems (e.g. Harmon *et al.*, 2009; Bassar *et al.*, 2010). Therefore, a second explanation is that genetic drift led to morphological, physiological or behavioural divergences among populations, thereby influencing the structure of the local ecosystems and communities. Either way, our work suggests that population and community differentiation are interrelated. An important next step will be to further unravel the exact mechanisms underlying the relationships between different levels of diversity.

We further showed that hypothesis 1 from Vellend & Geber (2005) was also likely to explain observed patterns, since differences in water composition had a positive influence on both genetic and species  $\beta$ -diversity. This result might reflect a mechanism of isolation-by-environment (Sexton *et al.*, 2014; Wang & Bradburd, 2014), by which gene flow between environmentally different sites is limited by the maladaptation of immigrants. Although this mechanism strongly relies on the effects of local selective pressures, it can also affect neutral genetic diversity when gene flow is reduced between sites differing in their local environments (Sexton *et al.*, 2014). Similarly, the structure and composition of a community is expected to vary with environmental features, here water composition, when species show strong habitat preferences. At the community level, environmental filtering may strongly

	$\bar{E}$	95% CI	$Q_t$	$P$
(a) $\alpha$ -diversity				
Isolation → Allelic richness	<b>-0.464</b>	<b>[-0.684; -0.244]</b>	19.863	<0.001
Area → Allelic richness	<b>0.238</b>	<b>[0.017; 0.458]</b>	6.647	0.084
Connectivity → Allelic richness	0.079	[-0.141; 0.299]	1.767	0.622
Water composition → Allelic richness	0.121	[-0.099; 0.341]	3.245	0.355
Oxygen → Allelic richness	0.175	[-0.045; 0.395]	3.628	0.304
Isolation → Species richness	<b>-0.322</b>	<b>[-0.542; -0.102]</b>	0.352	0.950
Area → Species richness	<b>0.405</b>	<b>[0.185; 0.625]</b>	3.354	0.340
Connectivity → Species richness	-0.089	[-0.309; 0.131]	0.850	0.837
Water composition → Species richness	<b>0.238</b>	<b>[0.018; 0.458]</b>	3.822	0.281
Oxygen → Species richness	0.175	[-0.395; 0.045]	0.382	0.944
Allelic richness → Species richness	-0.012	[-0.232; 0.208]	0.755	0.862
(b) $\beta$ -diversity				
Isolation → Jost's $D$	<b>0.316</b>	<b>[0.276; 0.356]</b>	344.950	<0.001
Area → Jost's $D$	-0.025	[-0.065; 0.015]	15.381	0.001
Connectivity → Jost's $D$	0.019	[-0.021; 0.059]	63.807	<0.001
Water composition → Jost's $D$	<b>0.175</b>	<b>[0.136; 0.215]</b>	240.067	<0.001
Oxygen → Jost's $D$	-0.018	[-0.058; 0.022]	38.508	<0.001
Isolation → True diversity	-0.011	[-0.051; 0.029]	34.910	<0.001
Area → True diversity	-0.035	[-0.075; 0.004]	6.416	0.093
Connectivity → True diversity	<b>0.053</b>	<b>[0.013; 0.092]</b>	6.319	0.097
Water composition → True diversity	<b>0.235</b>	<b>[0.195; 0.275]</b>	47.929	<0.001
Oxygen → True diversity	<b>-0.107</b>	<b>[-0.147; -0.067]</b>	5.621	0.132
Jost's $D$ → True diversity	<b>0.086</b>	<b>[0.046; 0.126]</b>	12.639	0.005

**Table 3** Mean effect size ( $\bar{E}$ ), 95% confidence interval (95% CI), total heterogeneity of a sample ( $Q_t$ ) and corresponding  $P$ -value ( $P$ ) computed from the path coefficients obtained from partial least-square path modelling applied to (a)  $\alpha$ -diversity indices and (b)  $\beta$ -diversity indices. Values in bold are significant mean effect sizes and the corresponding 95% confidence intervals.



**Fig. 6** Graphical representation of the meta-analysis results obtained from partial least square path modelling on  $\beta$ -diversity indices over all species. See legend in Fig. 5 for details.

drive community composition, which has been demonstrated in diverse freshwater organisms, including fish (Blanchet *et al.*, 2014; Heino, Melo & Bini, 2015).

Environmental filtering might, thus, occur both at the population and community levels to drive co-variation between community and population compositions.

In addition to observing a common influence of environmental filtering on species and genetic diversity, we also found that genetic and species  $\beta$ -diversity was driven by independent processes. For instance, our results suggest that genetic  $\beta$ -diversity was also driven by isolation-by-distance for all four species, which was not the case for species  $\beta$ -diversity. This shows that the equilibrium between genetic drift and gene flow in these four species is affected by geographical isolation, although not similarly as the effect size of isolation on genetic differentiation was heterogeneous among species. Because species  $\beta$ -diversity was not impacted by isolation, community and population differentiations may be partly driven by independent evolutionary and ecological processes, which may explain why  $\beta$ -SGDCs tended to be weaker than  $\alpha$ -SGDCs. Similarly, the heterogeneity in the effects of geographical isolation on genetic differentiation among species may explain why  $\beta$ -SGDCs vary among species (in terms of strength and significance).

#### *The role of spatial extent in quantifying SGDCs*

It is noteworthy that we observed heterogeneous effect sizes among the four species for several pathways



explaining species  $\alpha$ - and  $\beta$ -diversity. This is a rather unexpected outcome as the species lists were extracted from a unique database, and we expected species  $\alpha$ - and  $\beta$ -diversity indices to be related to the same variables, irrespectively of the target species. This heterogeneity could result from spatial-scale effects. Indeed, although the target populations showed a relatively high degree of spatial congruence, there were noticeable differences in spatial extents and environmental variabilities for the datasets used for each target species. These differences in spatial extents may explain the heterogeneity in effect sizes of several environmental variables on species  $\alpha$ - and  $\beta$ -diversity between target species (Lira-Noriega *et al.*, 2007; Cushman & Landguth, 2010). However, these spatial extent effects seemed to be of minor consequences and did not prevent the identification of the main mechanisms driving SGDCs as significant mean effect sizes were identified.

We demonstrated that combined effects of isolation and habitat area at the population and community levels are likely to be responsible for the positive  $\alpha$ -SGDCs observed in the four freshwater fish species. We further showed that positive  $\beta$ -SGDCs are likely to be generated by two co-occurring mechanisms, although additional mechanisms uniquely affecting one of the two levels of  $\beta$ -diversity probably contribute to make  $\beta$ -SGDCs weaker than  $\alpha$ -SGDCs. To sum up, while species and genetic  $\alpha$ -diversity were likely to be driven by similar landscape features, it appeared that species and genetic  $\beta$ -diversity were underlined by a combination of common and unique eco-evolutionary processes.

Alpha and  $\beta$ -SGDCs often vary strongly among studied organisms, even in a single landscape (e.g. Struebig *et al.*, 2011; Taberlet *et al.*, 2012), which was unexpectedly not verified in our study. We believe that the next important challenge of SGDC studies will be to better quantify the level of variation in life-history traits needed to generate different (or similar)  $\alpha$ - and  $\beta$ -SGDCs among species. To solve this issue empirically, it seems relevant to focus on several target species, although researchers will need to carefully design the study to encompass large ranges of life-history strategies (as in Taberlet *et al.*, 2012). The meta-analytic approach we used allowed highlighting both common and unique mechanisms driving  $\alpha$ - and  $\beta$ -diversity in these four species. All four fish species we considered here are from a similar trophic level (secondary consumers), but they varied strikingly in many traits. From our results, we can conclude that dissimilarities in species traits have no or little impact on the two mechanisms driving  $\alpha$ -SGDCs, namely the effects of isolation and habitat area,

whereas they may generate slight dissimilarities in the mechanisms driving genetic  $\beta$ -diversity.

## Acknowledgments

We are very grateful to all the ONEMA staff who contributed to the field sampling. Leslie Faggiano, Charlotte Evangelista, Christine Lauzeral, Loïc Tudesque, Roselyne Etienne, Sébastien Villéger, Julien Cucherousset, Vincent Dubut, Charlotte Veyssière and Alain Blanchet. Charlotte Veyssière, Roselyne Etienne and Vincent Dubut are also thanked for their help with molecular analyses. We warmly thank Mathew Seymour and Florian Altermatt for constructive discussions and for commenting/correcting the manuscript. Two reviewers provided insightful comments that greatly improve the quality of the manuscript. LF was financially supported by a MESR ('Ministère de l'Enseignement Supérieur et de la Recherche') PhD scholarship during this study. This project has been carried out with financial support from the Commission of the European Communities, specific RTD programme 'TWRMNET'. This work has been done in two research units (EDB & SETE) that are part of the 'Laboratoire d'Excellence' (LABEX) entitled TULIP (ANR-10-LABX-41).

## References

- Aljanabi S.M. & Martinez I. (1997) Universal and rapid salt-extraction of high quality genomic DNA for PCR-based techniques. *Nucleic Acids Research*, **25**, 4692–4693.
- Altermatt F. (2013) Diversity in riverine metacommunities: a network perspective. *Aquatic Ecology*, **47**, 365–377.
- Antonovics J. (1976) The input from population genetics: "the new ecological genetics". *Systematic Botany*, **1**, 233–245.
- Bassar R.D., Marshall M.C., López-Sepulcre A., Zandonà E., Auer S.K., Travis J. *et al.* (2010) Local adaptation in Trinidadian guppies alters ecosystem processes. *Proceedings of the National Academy of Sciences of the United States of America*, **107**, 3616–3621.
- Belkhir K., Borsa P., Chikhi L., Raufaste N. & Bonhomme F. (1996) GENETIX 4.05, Logiciel sous Windows TM pour la Génétique des Populations. Laboratoire Génome, Populations, Interactions, CNRS UMR 5000, Université de Montpellier II, Montpellier.
- Blanchet S., Helmus M.R., Brosse S. & Grenouillet G. (2014) Regional vs local drivers of phylogenetic and species diversity in stream fish communities. *Freshwater Biology*, **59**, 450–462.
- Blum M.J., Bagley M.J., Walters D.M., Jackson S.A., Daniel F.B., Chaloud D.J. *et al.* (2012) Genetic diversity and species diversity of stream fishes covary across a land-use gradient. *Oecologia*, **168**, 83–95.

- Buisson L., Blanc L. & Grenouillet G. (2008) Modelling stream fish species distribution in a river network: the relative effects of temperature versus physical factors. *Ecology of Freshwater Fish*, **17**, 244–257.
- Cadotte M.W. (2006) Dispersal and species diversity: a meta-analysis. *The American Naturalist*, **167**, 913–924.
- Campbell Grant E.H., Lowe W.H. & Fagan W.F. (2007) Living in the branches: population dynamics and ecological processes in dendritic networks. *Ecology Letters*, **10**, 165–175.
- Costedoat C. & Gilles A. (2009) Quaternary pattern of freshwater fishes in Europe: comparative phylogeography and conservation perspective. *The Open Conservation Biology Journal*, **3**, 36–48.
- Cushman S.A. & Landguth E.L. (2010) Spurious correlations and inference in landscape genetics. *Molecular Ecology*, **19**, 3592–3602.
- Derry A.M., Arnott S.E., Shead J.A., Hebert P.D.N. & Boag P.T. (2009) Ecological linkages between community and genetic diversity in zooplankton among boreal shield lakes. *Ecology*, **90**, 2275–2286.
- Duminil J., Fineschi S., Hampe A., Jordano P., Salvini D., Vendramin G.G. *et al.* (2007) Can population genetic structure be predicted from life-history traits? *The American Naturalist*, **169**, 662–672.
- Estrada E. & Bodin Ö. (2008) Using network centrality measures to manage landscape connectivity. *Ecological Applications*, **18**, 1810–1825.
- Evanno G., Castella E., Antoine C., Paillat G. & Goudet J. (2009) Parallel changes in genetic diversity and species diversity following a natural disturbance. *Molecular Ecology*, **18**, 1137–1144.
- Frankham R. (1996) Relationship of genetic variation to population size in wildlife. *Conservation Biology*, **10**, 1500–1508.
- Frankham R. (2015) Genetic rescue of small inbred populations: meta-analysis reveals large and consistent benefits of gene flow. *Molecular Ecology*, **24**, 2610–2618.
- Freeman L. (1977) A set of measures of centrality based on betweenness. *Sociometry*, **40**, 35–41.
- Gotelli N.J. & Taylor C.M. (1999) Testing metapopulation models with stream-fish assemblages. *Evolutionary Ecology Research*, **1**, 835–845.
- Hagberg A.A., Schult D.A. & Swart P.J. (2008) Exploring network structure, dynamics, and function using NetworkX. In: *Proceedings of the 7th Python in Science Conference (SciPy2008)* (Eds G. Varoquaux, T. Vaught & J. Millman), pp. 11–15, Pasadena. Available at [https://hal.archives-ouvertes.fr/hal-00502586/file/SciPy2008\\_proceedings.pdf](https://hal.archives-ouvertes.fr/hal-00502586/file/SciPy2008_proceedings.pdf).
- Harmon L.J., Matthews B., Des Roches S., Chase J.M., Shurin J.B. & Schluter D. (2009) Evolutionary diversification in stickleback affects ecosystem functioning. *Nature*, **458**, 1167–1170.
- He T. & Lamont B.B. (2010) Species versus genotypic diversity of a nitrogen-fixing plant functional group in a meta-community. *Population Ecology*, **52**, 337–345.
- He T., Lamont B.B., Krauss S.L., Enright N.J. & Miller B.P. (2008) Covariation between intraspecific genetic diversity and species diversity within a plant functional group: covariation between species and genetic diversity. *Journal of Ecology*, **96**, 956–961.
- Heino J., Melo A.S. & Bini L.M. (2015) Reconceptualising the beta diversity-environmental heterogeneity relationship in running water systems. *Freshwater Biology*, **60**, 223–235.
- Jackson D.A., Peres-Neto P.R. & Olden J.D. (2001) What controls who is where in freshwater fish communities – the roles of biotic, abiotic, and spatial factors. *Canadian Journal of Fisheries and Aquatic Sciences*, **58**, 157–170.
- Jombart T. (2008) adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics*, **24**, 1403–1405.
- Jost L. (2006) Entropy and diversity. *Oikos*, **113**, 363–375.
- Jost L. (2008) GST and its relatives do not measure differentiation. *Molecular Ecology*, **17**, 4015–4026.
- Jurasinski G. & Retzer V. (2012) *simba: A Collection of Functions for Similarity Analysis of Vegetation Data*. R package version 0.3-5. <http://CRAN.R-project.org/package=simba>
- Keith P., Persat H., Feunteun E., Adam B. & Geniez M. (2011) *Les Poissons d'eau douce de France*. Muséum National d'Histoire Naturelle and Publications Biotope, Paris and Mèze.
- Kelly R.P. & Palumbi S.R. (2010) Genetic structure among 50 species of the northeastern Pacific rocky intertidal community. *PLoS ONE*, **5**, e8594.
- Lamy T., Jarne P., Laroche F., Pointier J.-P., Huth G., Segard A. *et al.* (2013) Variation in habitat connectivity generates positive correlations between species and genetic diversity in a metacommunity. *Molecular Ecology*, **22**, 4445–4456.
- 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**, 59–69.
- Legendre P. & Fortin M.-J. (2010) Comparison of the Mantel test and alternative approaches for detecting complex multivariate relationships in the spatial analysis of genetic data. *Molecular Ecology Resources*, **10**, 831–844.
- Lira-Noriega A., Soberón J., Navarro-Sigüenza A.G., Nakazawa Y. & Peterson A.T. (2007) Scale dependency of diversity components estimated from primary biodiversity data and distribution maps. *Diversity and Distributions*, **13**, 185–195.
- Loreau M. (2000) Biodiversity and ecosystem functioning: recent theoretical advances. *Oikos*, **91**, 3–17.
- Mazé-Guilmo E., Blanchet S., McCoy K.D. & Loot G. (2016) Host dispersal as the driver of parasite genetic structure: a paradigm lost? *Ecology Letters*, **19**, 336–347.
- Múrrica C., Rugenski A.T., Whiles M.R. & Vogler A.P. (2015) Long-term isolation and endemism of Neotropical aquatic insects limit the community responses to recent amphibian decline. *Diversity and Distributions*, **21**, 938–949.

- Nakagawa S. & Cuthill I.C. (2007) Effect size, confidence interval and statistical significance: a practical guide for biologists. *Biological Reviews*, **82**, 591–605.
- Nychka D., Furrer R. & Sain S. (2015) *fields: Tools for Spatial Data*. R package version 8.4-1. [www.image.ucar.edu/fields](http://www.image.ucar.edu/fields).
- Odat N., Jetschke G. & Hellwig F.H. (2004) Genetic diversity of *Ranunculus acris* L. (Ranunculaceae) populations in relation to species diversity and habitat type in grassland communities. *Molecular Ecology*, **13**, 1251–1257.
- Paz-Vinas I. & Blanchet S. (2015) Dendritic connectivity shapes spatial patterns of genetic diversity: a simulation-based study. *Journal of Evolutionary Biology*, **28**, 986–994.
- Paz-Vinas I., Loot G., Stevens V.M. & Blanchet S. (2015) Evolutionary processes driving spatial patterns of intraspecific genetic diversity in river ecosystems. *Molecular Ecology*, **24**, 4586–4604.
- Pella H., Lejot J., Lamouroux N. & Snelder T. (2012) Le réseau hydrographique théorique (RHT) français et ses attributs environnementaux. *Géomorphologie: Relief, Processus, Environnement*, **3**, 317–336.
- Poulet N., Beaulaton L. & Dembski S. (2011) Time trends in fish populations in metropolitan France: insights from national monitoring data. *Journal of Fish Biology*, **79**, 1436–1452.
- Puşçaş M., Taberlet P. & Choler P. (2008) No positive correlation between species and genetic diversity in European alpine grasslands dominated by *Carex curvula*. *Diversity and Distributions*, **14**, 852–861.
- Quantum GIS Development Team (2015) *Quantum GIS Geographic Information System*. Open Source Geospatial Foundation Project. <http://www.qgis.org/>
- Radinger J. & Wolter C. (2014) Patterns and predictors of fish dispersal in rivers. *Fish and Fisheries*, **15**, 456–473.
- Raeymaekers J.A.M., Maes G.E., Geldof S., Hontis I., Nackaerts K. & Volckaert F.A.M. (2008) Modeling genetic connectivity in sticklebacks as a guideline for river restoration. *Evolutionary Applications*, **1**, 475–488.
- Rosenberg M.S., Adams D.C. & Gurevitch J. (2000) *MetaWin: Statistical Software for Meta-Analysis, Version 2.0*. Sinauer Associates, Sunderland.
- Rousset F. (2008) genepop'007: a complete re-implementation of the genepop software for Windows and Linux. *Molecular Ecology Resources*, **8**, 103–106.
- Saccheri I., Kuussaari M., Kankare M., Vikman P., Fortelius W. & Hanski I. (1998) Inbreeding and extinction in a butterfly metapopulation. *Nature*, **392**, 491–494.
- Sanchez G., Trinchera L. & Russolillo G. (2015) *plspm: Tools for Partial Least Squares Path Modeling (PLS-PM)*. R package version 0.4.7. <http://CRAN.R-project.org/package=plspm>
- Sei M., Lang B.K. & Berg D.J. (2009) Genetic and community similarities are correlated in endemic-rich springs of the northern Chihuahuan Desert. *Global Ecology and Biogeography*, **18**, 192–201.
- Sexton J.P., Hangartner S.B. & Hoffmann A.A. (2014) Genetic isolation by environment or distance: which pattern of gene flow is most common? *Evolution*, **68**, 1–15.
- Seymour M., Seppälä K., Mächler E. & Altermatt F. (2016) Lessons from the macroinvertebrates: species-genetic diversity correlations highlight important dissimilar relationships. *Freshwater Biology*, doi:10.1111/fwb.12816, in press.
- Struebig M.J., Kingston T., Petit E.J., Le Comber S.C., Zubaid A., Mohd-Adnan A. *et al.* (2011) Parallel declines in species and genetic diversity in tropical forest fragments: parallel diversity declines in forest fragments. *Ecology Letters*, **14**, 582–590.
- Szpiech Z.A., Jakobsson M. & Rosenberg N.A. (2008) ADZE: a rarefaction approach for counting alleles private to combinations of populations. *Bioinformatics*, **24**, 2498–2504.
- Taberlet P., Zimmermann N.E., Englisch T., Tribsch A., Holderegger R., Alvarez N. *et al.* (2012) Genetic diversity in widespread species is not congruent with species richness in alpine plant communities. *Ecology Letters*, **15**, 1439–1448.
- Tenenhaus M., Vinzi V.E., Chatelin Y.-M. & Lauro C. (2005) PLS path modeling. *Computational Statistics & Data Analysis*, **48**, 159–205.
- Van Oosterhout C., Hutchinson W.F., Wills D.P.M. & Shipley P. (2004) Micro-checker: software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes*, **4**, 535–538.
- Van Valen L. (1965) Morphological variation and width of ecological niche. *The American Naturalist*, **99**, 377–390.
- Vannote R.L., Minshall G.W., Cummins K.W., Sedell J.R. & Cushing C.E. (1980) The river continuum concept. *Canadian Journal of Fisheries and Aquatic Sciences*, **37**, 130–137.
- Vellend M. (2003) Island biogeography of genes and species. *The American Naturalist*, **162**, 358–365.
- Vellend M. (2005) Species diversity and genetic diversity: parallel processes and correlated patterns. *The American Naturalist*, **166**, 199–215.
- Vellend M. & Geber M.A. (2005) Connections between species diversity and genetic diversity. *Ecology Letters*, **8**, 767–781.
- 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**, 2890–2901.
- Wang I.J. & Bradburd G.S. (2014) Isolation by environment. *Molecular Ecology*, **23**, 5649–5662.
- Winter D.J. (2012) MMOD: an R library for the calculation of population differentiation statistics. *Molecular ecology resources*, **12**, 1158–1160.
- Wright S. (1921) Correlation and causation. *Journal of Agricultural Research*, **20**, 557–585.
- Xu W., Liu L., He T., Cao M., Sha L., Hu Y. *et al.* (2016) Soil properties drive a negative correlation between species

diversity and genetic diversity in a tropical seasonal rainforest. *Scientific Reports* 6, 20652.

### Supporting Information

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

**Table S1.** Minimum, mean, maximum and standard deviation of environmental variables.

**Table S2.** List of the 51 species of the regional pool, their family and percentage of occurrence over the 81 sites. Species in bold are the four target species of the study.

**Table S3.** Mean effect size ( $\bar{E}$ ), 95% confidence interval (95% CI), total heterogeneity of a sample ( $Q_t$ ) and corre-

sponding  $P$ -value ( $P$ ) computed from the path coefficients obtained from partial least-square path modelling applied to (a)  $\alpha$ -diversity indices and (b)  $\beta$ -diversity indices computed after the removal of non-native species. Values in bold are significant mean effect sizes and the corresponding 95% confidence intervals.

**Table S4.** Sampling information on microsatellite loci and multiplexed PCR used in this study for the four species.

**Table S5.** Correlation values between the latent variables of the eight models.

**Table S6.** Tables for Hardy–Weinberg and linkage disequilibrium tests for all species.

(Manuscript accepted 1 August 2016)