Received: 13 May 2022 | Revised: 4 November 2022 | Accepted: 28 December 2022

DOI: 10.1111/geb.13637

Global Ecology and Biogeography | A Journal of Macroecology

WILEY

RESEARCH ARTICLE

Spatial genetic differentiation correlates with species assemblage turnover across tropical reef fish lineages

Maurine Vilcot[1] | Camille Albouy[2,3] | Giulia Francesca Azzurra Donati[3,4] | Thomas Claverie[5,6] | Pagu Julius[7] | Stéphanie Manel[1] | Loïc Pellissier[2,3] | Fabien Leprieur[5,8]

[1]CEFE, Univ Montpellier, CNRS, EPHE-PSL University, IRD, Montpellier, France

[2]Ecosystems and Landscape Evolution, Institute of Terrestrial Ecosystems, Department of Environmental Systems Science, ETH Zürich, Zürich, Switzerland

[3]Swiss Federal Institute for Forest, Snow and Landscape Research WSL, Birmensdorf, Switzerland

[4]Swiss Federal Institute of Aquatic Science & Technology (Eawag), Dübendorf, Switzerland

[5]MARBEC, Univ Montpellier, CNRS, IFREMER, IRD, Montpellier, France

[6]Centre Universitaire de formation et de recherche de Mayotte, Dembeni, France

[7]Mafia Island Marine Park, Mafia, Tanzania

[8]Institut Universitaire de France, Paris, France

Correspondence
Maurine Vilcot, CEFE, Univ Montpellier, CNRS, EPHE-PSL University, IRD, Montpellier, France.
Email: maurine.vilcot@gmail.com

Funding information
SNF-ANR, Grant/Award Number: 310030E-164294

Handling Editor: Andrés Baselga

Abstract

Aim: Evaluating the similarity of diversity patterns across micro- to macroevolutionary scales in natural communities, such as species–genetic diversity correlations (SGDCs), may inform on processes shaping community assembly. However, whether SGDCs not only hold across communities but also across lineages has never been explored so far. Here we investigated SGDCs across co-distributed taxa for different spatial components (α, β, γ), and formally tested the influence of dispersal traits on β-SGDCs.

Location: Western Indian Ocean.

Time period: 2016–2017.

Major taxa studied: Tropical reef fish species with contrasting dispersal traits.

Methods: Using double-digest restriction-site associated DNA sequencing (ddRAD-seq) Single Nucleotide Polymorphism data for 20 tropical reef fishes and distribution data of 2,446 species belonging to 12 families, we analysed the correlations between within-species genetic diversity and within-family species diversity (i.e., lineage diversity) for the three spatial components (α, β, γ-SGDCs). We then related the strength of β-SGDCs per species to proxies of larval dispersal abilities.

Results: We detected positive and significant lineage-based SGDC only for the β component, that is, the families showing the greatest level of species turnover among sites contain the species with the greatest levels of genetic differentiation. We showed that the Monsoon Drift mainly explained the β-diversity patterns at both intraspecific and interspecific levels. Higher β-SGDCs were found for species with short pelagic larval duration and weak larval swimming capacity.

Main conclusions: Our study reveals a strong correlation between genetic and species β-diversity, a result explained by the presence of a 'soft' barrier and mediated by larval dispersal processes. This suggests that vicariance and dispersal limitation are major processes shaping β-diversity patterns from microevolutionary to macroevolutionary scales in tropical reef fishes.

Loïc Pellissier and Fabien Leprieur share senior authorship of this article.

KEYWORDS
ddRADseq, dispersal, genetic diversity, Indian Ocean, macrogenetics, marine barrier, species diversity, tropical reef fishes, β-diversity

1 | INTRODUCTION

Understanding how genetic diversity is related to species diversity has puzzled evolutionary ecologists for over four decades (e.g., Antonovics, 1976; Lino et al., 2021; Vellend, 2002). It has been the focus of numerous studies exploring species–genetic diversity correlations (SGDCs) in both marine and terrestrial ecosystems (Fourtune et al., 2016; Manel et al., 2020; Messmer et al., 2012; Robuchon et al., 2019; Vellend, 2003). According to Antonovics (1976), genetic and species diversity are maintained by parallel processes (i.e., mutation/speciation, genetic drift/ecological drift, gene flow/dispersal, natural selection/environmental and biotic filtering, respectively), which have been conceptually formalized by Vellend (2010) to better understand how natural communities are assembled (Hubbell, 2001; Vellend, 2003, 2005; Vellend & Geber, 2005). Despite this vast bulk of research, whether spatial patterns of diversity are correlated at interspecific and intraspecific levels remains a controversial question (Laroche et al., 2015) and appears to be context dependent (Lamy et al., 2017). In addition, the correlation between species and genetic diversity has been commonly explored across spatially structured communities (SGDC hereafter) but never across different lineages (lineage-based SGDC hereafter). The commonly used SGDC asks whether communities of high species richness are composed of species with high genetic diversity, while the lineage-based SGDC asks whether lineages of high species richness are composed of species with high genetic diversity.

In community ecology and biogeography, species diversity (SD) is partitioned into three main spatial components (α, β and γ), which describe diversity patterns at different spatial scales. (a) α-SD represents the number of species in a given locality, (b) γ-SD the number of species in the total defined region, and (c) β-SD corresponds to differences in species composition among localities (Ellison, 2010; Supporting Information Table S1). Even though often rarely considered in population genetics, the same three components can be used when exploring intraspecific genetic diversity (GD) patterns (Gaggiotti et al., 2018; Tuomisto, 2010): (a) α-GD as the genetic diversity within a single local group of individuals, (b) γ-GD as the genetic diversity across all individuals of the defined region, and (c) β-GD as the genetic differentiation among local groups of individuals that are geographically separated and consequently reflecting the degree of spatial genetic structure (Donati et al., 2021; Supporting Information Table S1). Overall, this diversity partitioning framework allows characterization of complex spatial patterns and thus provides a better understanding of the processes shaping the diversity at both intra- and interspecific levels (Donati et al., 2021; Gaggiotti et al., 2018; Tuomisto, 2010). SGDCs studies generally focus on diversity sensu stricto (α or γ components; e.g., Manel et al., 2020; Vellend, 2005) or differentiation only (β component; e.g., Robuchon et al., 2019), but rarely consider the three spatial components together (but see Fourtune et al., 2016 and Papadopoulou et al., 2011 considering both α and β components).

If neutral processes such as drift and migration exclusively determine diversity patterns across evolutionary scales (Antonovics, 1976), they would similarly influence GD and SD leading to positive SGDCs. The influence of these neutral processes manifest through the well-known correlations between α-diversity and habitat area or isolation at both interspecific and intraspecific levels (Fourtune et al., 2016; Messmer et al., 2012; Papadopoulou et al., 2011; Vellend & Geber, 2005). Similarly, positive β-SGDCs have been demonstrated to occur as a result of differential connectivity among habitats associated with dispersal processes for freshwater fishes (Fourtune et al., 2016; Robuchon et al., 2019). In a pioneer study using haplotype data for multiple species of tenebrionid beetles of the Aegean archipelago, Papadopoulou et al. (2011) also identified a positive β-SGDC, caused by stochastic dispersal processes as expected by neutral theories (Antonovics, 1976; Hubbell, 2001). In the literature, however, the strength and even the sign of SGDCs vary from study to study (Lamy et al., 2017), and such deviation from neutral expectations may be caused by environmental constraints affecting GD and SD in opposite ways. In this context, Baselga et al. (2022) have recently provided a theoretical framework to quantify the role of dispersal and environmental constraints in shaping β-diversity patterns across genealogical scales. Using simulations and empirical data, they showed similar spatial turnover rates at both haplotype and species levels if dispersal limitation processes dominated, while different spatial turnover rates were found if both environmental constraints and dispersal limitation influenced species distribution. However, as pointed out by Robuchon et al. (2019), the observed variations of SGDCs among species has never been formally tested and the role of species traits in shaping the strength of the correlation remains understudied.

Tropical reef fishes belong to one of the most species-rich and morphologically diverse vertebrate groups, with large differences in dispersal ability among species (Claverie & Wainwright, 2014; Luiz et al., 2013; Price et al., 2011), which make it an ideal taxonomic group to explore the role of dispersal processes in shaping SGDCs. Furthermore, SGDCs and the underlying processes have been largely unexplored in the marine environment (but see Manel et al., 2020; Messmer et al., 2012). We focused on tropical reef fishes of the Western Indian Ocean (WIO), selected to maximize the representation of trait variability across species (dispersal potential, reproductive strategy and body size; see Donati et al., 2021). We capitalized on a unique genomic data set derived from an extensive sampling of 20 species and 833 individuals over four sites of the WIO (Mafia Island, Seychelles, Mayotte, Maldives), which are separated by hundreds or even thousands of kilometres (Donati et al., 2021). This sampling design is ideal to evaluate

the role of 'soft' barriers (Cowman & Bellwood, 2013) in shaping biogeographical patterns in the marine realm since the Maldives, in addition to being the most distant studied locality, is also separated from the other sampled localities by a potential 'soft' marine barrier created by the Monsoon Drift. This seasonally varying marine current of the northern Indian Ocean (Schott & McCreary, 2001) is composed of an equatorial counter current north of the Seychelles, that is, an eastward flowing, wind-driven current extending to depths of 100–150 m. 'Soft' barriers such as marine currents or large oceanic distance may impede the movement of adults and/or dispersal of pelagic larvae, but in contrast to 'hard' barriers, 'soft' barriers are considered as permeable (Cowman & Bellwood, 2013).

Based on this large genomic data set of 20 tropical reef fishes, we explore lineage-based SGDCs, namely the correlation between the genetic diversity of a given species (within-species GD) and the species diversity within their respective lineage (within-family SD), at the α, β and γ scales. By doing so, we ask the following questions: (a) Are lineages with high local species richness (α-SD) composed of species with high local genetic diversity (α-GD)? If parallel processes such as speciation/mutation and ecological/genetic drift act at both intra- and interspecific levels of diversity (Antonovics, 1976; Hubbell, 2001), we would expect species with higher α-GD to belong to families with higher α-SD. (b) Are lineages with great levels of species dissimilarity between sites (i.e., species beta diversity, β-SD) composed of species with great levels of genetic differentiation (β-GD)? Under neutral theories (Hubbell, 2001), if limited gene flow resulting from dispersal limitation is pivotal in driving β-diversity at intra- and interspecific levels, we expect positive β-SGDCs and greater levels of β-diversity for the families to which species with the lowest larval dispersal abilities belong. (c) Are lineages with high regional species richness (γ-SD) composed of species with high regional genetic diversity (γ-GD)? We expect neutral processes shaping both α- and β-diversity to cause positive γ-SGDCs. We further explored β-SGDCs for each species independently (see Robuchon et al., 2019), which allows comparison of the strength of β-SGDCs among species with contrasting biological and ecological characteristics, and more particularly those related to dispersal capacity. We predict greater genetic differentiation (β-GD) between the Maldives and the other localities for species with the lowest larval dispersal abilities, the Monsoon Drift being potentially less permeable for those species. Overall, we hypothesize that limited gene flow, resulting from the interplay between the presence of a 'soft' geographical barrier and larval dispersal abilities, is the major process shaping β-diversity patterns from micro- to macroevolutionary scales in tropical reef fishes of the WIO.

2 | METHODS

2.1 | Multi-species data set

We used a data set of 20 reef fish species co-occurring in four populations of the WIO. Species were chosen to maximize the representation of the ecological traits spectrum, so that this multi-species framework is particularly suitable to compare SGDC patterns in species with different life history traits, and thus understand why SGDC studies come up with such different results (Lamy et al., 2017). We used double-digest restriction-site associated DNA sequencing (ddRADseq) data on 20 tropical reef fish species. To expand the data set of Donati et al. (2021), we added ddRADseq data for *Parapercis hexophtalma* and *Naso brevirostris*. We removed *Zanclus cornutus* from the analysis as there are no other species in its family (Zanclidae), which precludes the possibility to estimate within-family SD. The resulting data set contains Single Nucleotide Polymorphism (SNP) data for 833 individuals of 20 tropical reef fish species (Supporting Information Table S2) sampled from four sites in the WIO: Mafia Island (Tanzania), Mayotte Island (Comoros Archipelago), Seychelles, and Maldives (Figure 1a). For more details on sampling strategy and SNP data set preparation, see Donati et al. (2021) or Supporting Information (Supplementary methods). As genomic diversity is directly related to the number of SNPs (Moragues et al., 2010), we randomly down-sampled 999 times the SNPs to the lowest common number of SNPs found across all species (i.e., $n = 4{,}479$ SNPs in *Pseudanthias squamipinnis*; Supporting Information Table S2). To best account for differences in site sampling success across species, we standardized the number of individuals per sampling site to a maximum of 10 (median value of the overall sampling), randomly selected 999 times.

2.2 | Genetic structure at individual level

To investigate the genetic structure at individual level, we applied a discriminant analysis of principal components (DAPC), a multivariate statistical approach for which variance in the sample is partitioned into between-group and within-group components to maximize discrimination between groups (Jombart et al., 2010). We performed the DAPC for each species in the SNP data set by using the *dapc* function implemented in the R package 'adegenet' (Jombart, 2008; Jombart & Ahmed, 2011), with sampling sites of each individual used as prior information (Supporting Information Figure S1). Individuals' membership probability to each cluster was computed and represented as bar plots. In order to perform a multi-species comparison, we retained the same number of principal components analysis (PCA) axes per species in the DAPC. As *Caranx melampygus* is only represented by 15 individuals, we decided to retain five PCA axes to avoid overfitting.

2.3 | Genetic diversity

We estimated the within-species α-GD with the expected heterozygosity H_S (Supporting Information Table S3; Nei, 1987). For the four sampling sites, H_S was calculated using the *basic.stats* function from the 'hierfstat' R package (Goudet & Jombart, 2015). Next, we obtained a measure of the $\bar{\alpha}$-GD as a mean of α-GD across the four sites. Similarly, we estimated the within-species γ-GD over all sampled individuals of

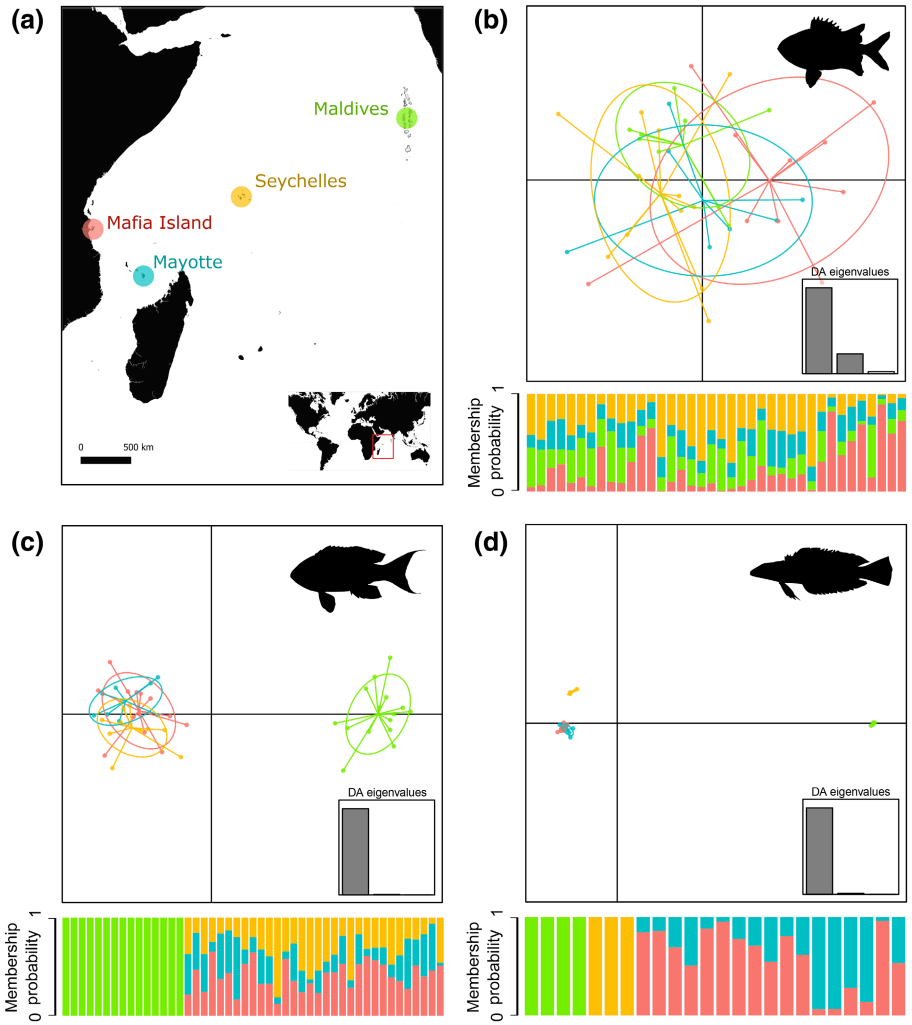

FIGURE 1 Genetic structure in Western Indian Ocean for 20 tropical reef fishes. (a) Map of the four sampling sites; see Donati et al. (2021) for more details. (b–d) Discriminant analyses of principal components (DAPC) on the complete SNP data set (see Supporting Information Table S2 for number of SNPs per species). (Top) Scatter plot of the first and second discriminant functions of DAPC with sampling location priors. Individuals are colour-coded by sampling site. Discriminant analysis (DA) eigenvalues are displayed. (Bottom) Bar plot showing individuals' membership probabilities to the different clusters (i.e., sampling sites). To illustrate different levels of genetic structure among sites, we selected three contrasted example species. From the least to the most differentiated: (b) *Myripristis violacea* (Holocentridae) $G''_{ST} = 0.007$, (c) *Pseudanthias squamipinnis* (Serranidae) $G''_{ST} = 0.027$, and (d) *Parapercis hexophtalma* (Pinguipedidae) $G''_{ST} = 0.27$.

the WIO, using the total heterozygosity H_T estimation of the *basic.stats* function of the 'hierfstat' R package (Supporting Information Table S3; Goudet & Jombart, 2015). For each species, both pairwise (among each pairs of sites) and multiple-site (over the four sites) β-GD were quantified with estimates of Hedrick's G''_{ST}, implemented respectively using *Gst_Hedrick* and *pairwise_Gst_Hedrick* functions of the 'mmod' R package (Supporting Information Table S4; Winter, 2012). G_{ST} is defined as $\frac{H_T - H_S}{H_T}$ (Nei, 1987); it represents the genetic distance between populations, analogous to the common F_{ST} based on biallelic markers but generalized for multiple alleles. G''_{ST} is a corrected version of the standardized Hedrick's $G'_{ST} = \frac{G_{ST}}{G_{ST(max)}}$ (Hedrick, 2005) that includes a correction term to account for when the number of sampled populations is small (Meirmans & Hedrick, 2011). All genetic diversity measures were averaged across the 999 × 4,479 SNP data sets.

2.4 | Species diversity

We obtained the fish species occurrence data from Albouy et al. (2019) that merged data extracted from the Ocean Biogeographic Information System (OBIS, http://www.obis.org) with the GASPAR database (Kulbicki et al., 2013; Parravicini et al., 2013; Pellissier et al., 2014). This database provides information on the occurrences of 14,204 fish species assigned to a spatial location in a grid system covering all oceans at 1° × 1° resolution. From this global database, we extracted the occurrences of species belonging to the 12 families studied, representing 2,446 species worldwide. We restricted the data set to a distance of 1° (111 km at the equator) from the reefs of the four sampling areas by using the function *gbuffer* from the R package 'rgeos' (Bivand & Rundel, 2019) and *st_intersects* from the

package 'sf' (Pebesma, 2018) to produce a species presence/absence matrix by site.

For each family, we assessed α-SD as the number of species present at each site, that is, within-family local species richness (Supporting Information Table S5). The calculation of SD within a family rather than over the entire community allows on the one hand lineage-based SGDCs to be computed, and on the other hand similarity in species past history to be controlled for. ᾱ-SD was obtained by averaging the α-SD obtained for the four sites. Similarly, for each family we computed γ-SD as the number of species from the four sites, that is, within-family regional species richness (Supporting Information Table S5). To quantify the β-SD, we calculated the Jaccard's dissimilarity index (β_{jac}-SD) as well as its turnover component (β_{jtu}-SD; Supporting Information Table S6), which is independent from differences in species richness between assemblages (Baselga, 2010). Both of these two β-diversity indices were computed as pairwise and multiple-site values using the 'betapart' package (Baselga et al., 2018). Both β_{jac}-SD and β_{jtu}-SD are used for the following analysis to compare results obtained with both indices.

2.5 | Drivers of β-diversity patterns

To evaluate whether the Monsoon Drift and geographical distance are factors driving both genetic differentiation and species turnover in tropical reef fishes of the WIO region, we used multiple regression on distance matrices (MRM; Legendre & Legendre, 1998) implemented in the *MRM* function from the R package 'ecodist' (Goslee & Urban, 2007). MRM is similar to linear regression, except that the dependent and independent variables are square distance matrices instead of single vectors (Legendre & Legendre, 1998). For a given explanatory variable, calculation of the standardized partial regression coefficient enabled us to compare their per-unit effects on compositional similarity, while controlling for the effects of the other variables. To overcome the problem of lack of independence between site pairs, we assessed the significance of the standardized partial regression coefficients and the coefficients of multiple determination (R^2) by using a permutation test (n = 999). Overall, we related both pairwise β-GD and pairwise β-SD to (a) geographical distance between sites and a (b) binary distance metric reflecting whether the studied sites are separated or not from the Maldives by the Monsoon Drift (Table 1). (a) Geographical distances between sites were computed as in-water distances, that is, the length of the shortest path within water between two sites with the *lc.dist* function of the 'marmap' R package (Pante & Simon-Bouhet, 2013). (b) The binary distance metric consists of assigning a value of 1 for two sites separated by the assumed 'soft' barrier (i.e., between the Maldives and the three other localities of the WIO), and a value of 0 for two sites that are not separated by this barrier (i.e., within the three south-west sites). This MRM method was performed considering all of the six pairwise distance values of each of the 20 species, that is, 120 non-independent observations.

TABLE 1 Results of multiple regression on distance matrices (MRM) assessing the effects of geographical distance and a marine 'soft' barrier on genetic differentiation (log β-GD) and species turnover (β_{jtu}-SD) from the six pairwise distance values of each of the 20 species (i.e., 120 non-independent observations). Response variable, explanatory variables, and R^2 of each model are reported in the first column.

Variables	Coefficient	*p*-value
Genetic differentiation (log β-GD), R^2 = .2688		
Geographical distance	.0809	.5673
Marine barrier	.8948	**.0021**
Species turnover (β_{jtu}-SD), R^2 = .3617		
Geographical distance	.2182	.1099
Marine barrier	.8122	**.0026**

Note: Significant *p*-values (< .05) are indicated in bold.

2.6 | Species–genetic diversity correlations (SGDCs)

To explore the lineage-based SGDCs for each of the three spatial components of diversity (α, β, γ), we assessed the relationship between within-species GD and their respective within-family SD using a linear regression model (LM). For the lineage-based β-SGDC, estimates of β-GD were log-transformed to fit normality required in linear models. This relationship was tested with either β_{jtu}-SD (Figure 2) or β_{jac}-SD (Supporting Information Figure S2). To further assess the robustness of the lineage-based β-SGDC, we also computed the LM excluding the values for *Parapercis hexophtalma* (Pinguipedidae), as this species stands as an outlier in terms of β-GD compared to the other species of the study (Supporting Information Figure S3).

To quantify β-SGDCs for each species separately, we calculated the correlation coefficient *r* associated with Mantel tests between β-GD distance matrix and β-SD distance matrix of the four sites (Supporting Information Figure S4, Table S7). Because the use of Mantel tests has been debated (Legendre et al., 2015; Peres-Neto & Jackson, 2001; Raufaste & Rousset, 2001), we also conducted a Procrustes analysis that allows the detection of matrix association under a variety of possible scenarios (*protest* R function 'vegan' package; Oksanen et al., 2019). Only Mantel test results are presented in the main text as the Procrustes analysis provided similar results. To test the overall significance of β-SGDCs over the 20 species, we compared the observed mean *p*-value of the 20 β-SGDCs (Mantel test or Procrustes analysis) to a null distribution of mean *p*-values obtained by sampling randomly 99,999 times 20 *p*-values in a uniform distribution in [0, 1] (see Robuchon et al., 2019). 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 (Supporting Information Figure S5, Table S7).

Given the low number of studied localities (n = 4) and because our aim was initially to collect genomic information for the largest number of species covering different ecological strategies and life histories,

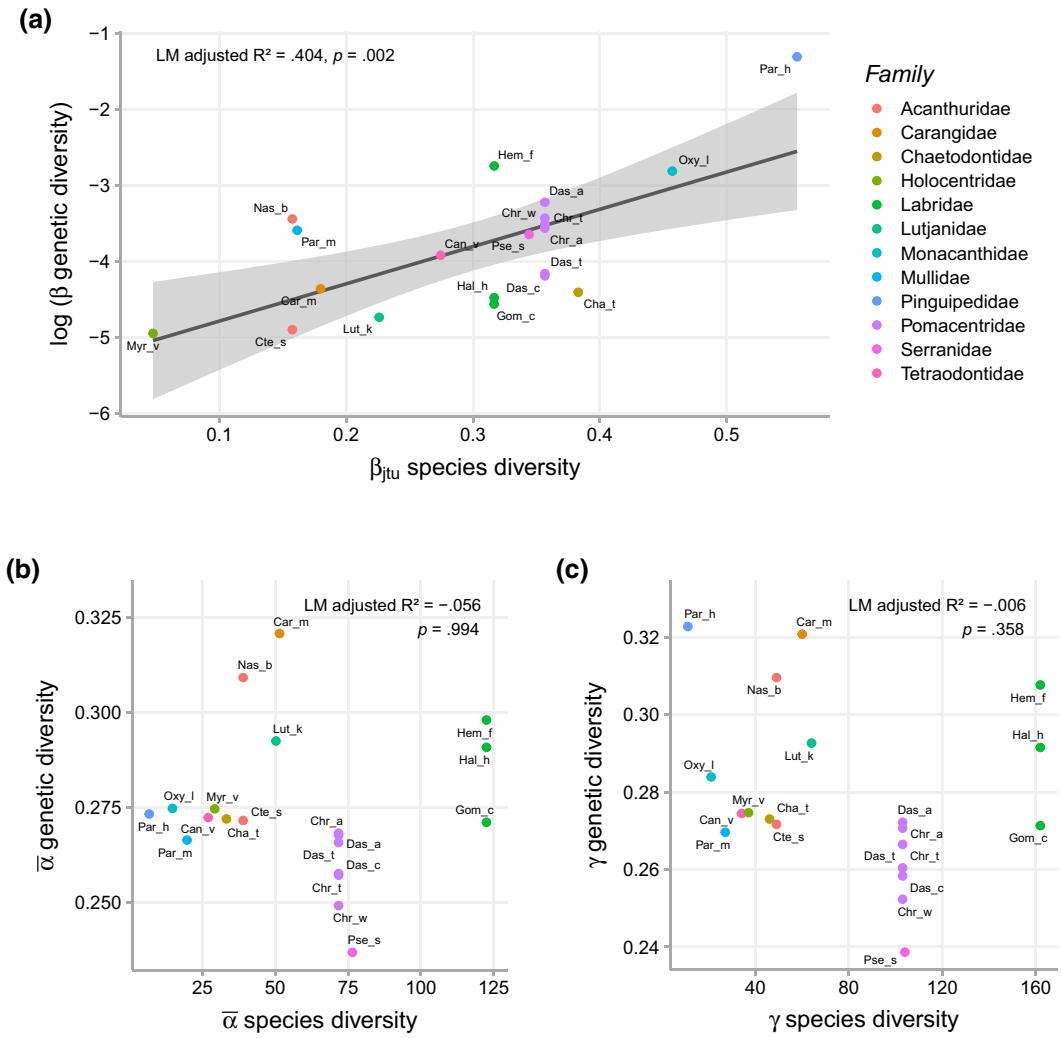

FIGURE 2 Lineage-based species–genetic diversity correlations (SGDCs) in 20 coral reef fish species of the Western Indian Ocean. (a) Lineage-based β-SGDC: linear model (LM) between log-transformed multiple-site β genetic diversity (β-GD) and multiple-site turnover component of Jaccard's dissimilarity index (β_{jtu}-SD). β-GD was log-transformed to fit normality required in linear models. (b) Lineage-based α-SGDC: linear model between mean H_s (ᾱ-GD) and mean local species richness (ᾱ species richness). (c) Lineage-based γ-SGDC: linear model between H_T (γ-GD) and regional species richness (γ-SD). Each dot corresponds to genetic diversity of one species, and the species diversity of its family. Species are indicated by a four-letter code (see Supporting Information Table S2).

we did not assess the statistical significance of SGDCs for the α component (i.e., for each species separately, the relationship between the local within-species α-GD and the local within-family α-SD).

2.7 | Influence of dispersal on β-SGDCs

To explore the link between larval dispersal ability and the strength of β-SGDCs, we extracted pelagic larval duration (PLD) for the 20 reef fish species from Fishbase and the literature (Luiz et al., 2013). The PLD is one of the most widely used indicators of larval dispersal potential (Nanninga & Manica, 2018). It has been shown to contribute to gene flow between geographically isolated populations and thus genetic structure, but often with only weak relationships (Nanninga & Manica, 2018; Riginos et al., 2011). The passive aspect of larval dispersal is now challenged by findings on larval swimming capacities (Fisher et al., 2005). The idea of active dispersal implies that larvae are able to influence their spatial and temporal patterns of dispersal. Critical swimming speed (*U*-crit) measures the maximum larval swimming speed, which reflects such active larval dispersal potential (Hogan et al., 2007). *U*-crit has been shown to have stronger relationships with genetic differentiation and range size than PLD and is a powerful predictor of long-distance dispersal potential (Nanninga & Manica, 2018). Therefore, we gathered larvae critical swimming speed values (*U*-crit) at the scale of the family for 10 of the 12 families from Fisher et al. (2005). Here, we refer to the PLD as a proxy for passive larval dispersal and *U*-crit as a proxy for active larval dispersal (Hogan et al., 2007).

We applied a linear model between the 20 Mantel test's correlation coefficients and estimates of active (*U*-crit) and log-transformed passive (PLD) dispersal abilities (Figure 3). To test whether significant relationships were robust to phylogenetic relatedness, we used

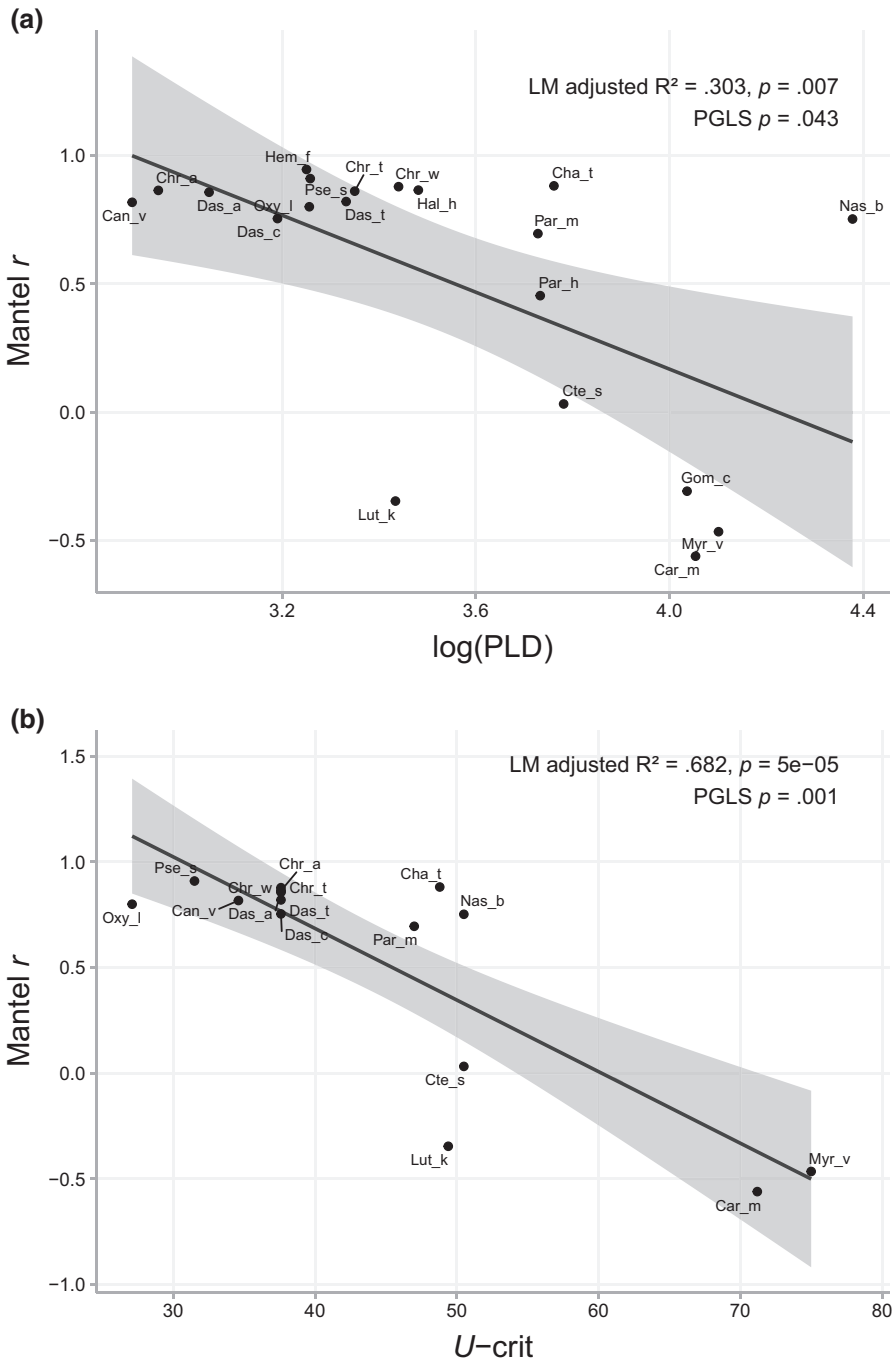

FIGURE 3 Negative correlation between β-species–genetic diversity correlations (β-SGDCs) and larval dispersal ability. Phylogenetic generalized least squares (PGLS) between the correlation coefficient (*r*) of the β-SGDC Mantel test and (a) the log-transformed pelagic larval duration (PLD) used as a proxy of passive larval dispersal abilities, or (b) critical swimming speed (*U*-crit), a proxy of active larval dispersal abilities. The slopes correspond to the regression coefficient estimated by a linear model (LM). The 20 species of the study are represented in (a), while only 16 are present in (b) due to the few values of *U*-crit available in the literature. Species are indicated by a four-letter code (see Supporting Information Table S2).

phylogenetic generalized least squares (PGLS) methods (Freckleton et al., 2002) implemented in the *gls* function of the 'nlme' package (Pinheiro et al., 2021). PGLS is a modification of generalized least squares that accounts for the non-independence of trait values across related taxa (Symonds & Blomberg, 2014). For the PGLS, we pruned the 20 species from the most comprehensive available time-calibrated phylogeny of fishes (Rabosky et al., 2018). The same approach was applied with the Procrustes analysis instead of the Mantel test (Supporting Information Figure S6). All the analyses were performed in R version 4.1.2 (R Core Team, 2021). Figures were made using the 'ggplot2' package (Wickham, 2016).

3 | RESULTS

3.1 | Genetic structure at individual level

By running the DAPC on the SNP data, we identified three different patterns of genetic structure among species. First, for five species such as the lattice soldierfish (*Myripristis violacea*) and the bluefin trevally (*Caranx melampygus*), we did not detect genetic clustering (Figure 1b, Supporting Information Figure S1). Second, for most species (n = 13) such as the sea goldie (*Pseudanthias squamipinnis*) or the Valentin's sharpnose puffer (*Canthigaster valentini*), the Maldives

constitutes one separate cluster, while the other three sampling sites form one homogeneous genetic cluster (Figure 1c). For the 13 species with reduced gene flow, individuals of the Maldivian population are distributed along the first DAPC axis reflecting a gradual and species-specific geographical isolation (Supporting Information Figure S1). We identified three genetic clusters for the speckled sandperch (*Parapercis hexophtalma*) and the harlequin filefish (*Oxymonacanthus longirostris*): (a) Seychelles, (b) Maldives, and (c) Mayotte Island and Mafia Island (Figure 1d).

3.2 | Species and genetic diversities

Both γ-GD (min: *Pseudanthias squamipinnis* 0.239 – max: *Parapercis hexophtalma* 0.323; Supporting Information Table S3) and ᾱ-GD (min: *Pseudanthias squamipinnis* 0.237 – max: *Caranx melampygus* 0.322; Supporting Information Table S3) were not found to vary much among species, while γ-SD (min: Pinguipedidae 11 – max: Labridae 162; Supporting Information Table S5) and ᾱ-SD (min: Pinguipedidae 6.75 – max: Labridae 122.5; Supporting Information Table S5) spanned over one order of magnitude among families. Multiple-site β-GD spanned two orders of magnitude (min: 0.007 – max: 0.270; Supporting Information Table S4). The species with the lowest multiple-site β-GD (0.007 for *Myripristis violacea*; Supporting Information Table S4) was found in the family with the lowest β_{jtu}-SD (0.048 for Holocentridae; Supporting Information Table S6), while *Parapercis hexophtalma* had the highest β-GD (0.268) and belongs to the family with the highest β_{jtu}-SD (0.556 for Pinguipedidae).

3.3 | Drivers of β-diversity patterns

To evaluate the processes acting on both β-GD and β-SD, we implemented a multiple regression on distance matrices (MRM) with two explanatory variables: the Monsoon Drift and geographical distance (Table 1). We observed that β-diversity patterns were associated with this marine 'soft' barrier at both micro- (β-GD, coef = .895, *p*-value = .002) and macroevolutionary scales (β_{jtu}-SD, coef = .812, *p*-value = .003), but not with geographical distance.

3.4 | Species–genetic diversity correlations (SGDCs)

We find no evidence of lineage-based α- or γ-SGDCs, that is, the correlation between within-family SD and within-species GD for both α (LM: adjusted $R^2 = -.056$, $p = .99$; Figure 2b) and γ (LM: adjusted $R^2 = -.006$, $p = .36$; Figure 2c). In contrast, for the lineage-based β-SGDCs, we detected a significant positive linear relationship between the log-transformed multiple-site β-GD and multiple-site β_{jac}-SD (LM: slope = 6.889, $p = .003$, adjusted $R^2 = .366$; Supporting Information Figure S2). This relationship was also found when considering only the turnover component of Jaccard's dissimilarity index, namely β_{jtu}-SD (LM: adjusted $R^2 = .404$, slope = 4.897, $p = .002$; Figure 2a). We also noted that removing the highly structured species from the analysis (*Parapercis hexophtalma*) reduced by half the R^2 of the lineage-based β-SGDC but the relationship was still significant (LM: slope = 3.195, $p = .030$, adjusted $R^2 = .205$; Supporting Information Figure S3).

We also detected an overall positive β-SGDC per species ($p < .001$; Supporting Information Table S7, Figure S5). Mantel test correlations between the β-GD distance matrix and the corresponding β_{jtu}-SD distance matrix were positive for most species (Mantel $r_{mean \pm SD} = .526 \pm .513$, Supporting Information Table S7, Figure S4). This result applies equally to the Procrustes analysis ($r_{mean \pm SD} = .875 \pm .117$, Supporting Information Table S7).

3.5 | Influence of dispersal abilities on β-SGDCs

In order to explain the differences in β-SGDCs observed across species, we examined the influence of dispersal abilities on the strength of β-SGDCs. We found a significant relationship between the correlation coefficient (*r*) of β-SGDC Mantel tests and the log-transformed PLD (LM: adjusted $R^2 = .303$, slope = −0.749, $p = .007$; PGLS: slope = −0.614, $p = .043$; Figure 3a). Species with lower larval dispersal abilities show higher correlation between genetic (β-GD) and species (β-SD) spatial structure. Furthermore, using the *U*-crit values instead of PLD over 16 species almost doubled the value of the correlation (LM: adjusted $R^2 = .682$, slope = −0.034, $p < .001$; PGLS: slope = −0.033, $p = .001$; Figure 3b).

4 | DISCUSSION

This investigation of the lineage-based β-SGDC highlights a strong association between within-species genetic differentiation (β-GD) and within-family species turnover (β-SD) in tropical reef fishes. By contrast, no significant correlation was found between within-species genetic diversity and within-family species diversity at the α and γ scales. Furthermore, our multi-species analysis reveals that the strength of β-SGDCs can be explained by differences in larval dispersal ability among species. Our results suggest that dispersal processes, acting across micro- and macroevolutionary scales, are important drivers of large-scale β-diversity patterns in tropical reef fishes.

4.1 | Isolation as a parallel determinant of genetic differentiation and species beta diversity

Using a lineage-based approach, we found a marked association between β-GD and β-SD in tropical reef fishes of the WIO (Figure 2). This contrasts with what is expected for marine organisms given that they have high dispersal potential and there are few barriers

to dispersal. Indeed, marine organisms tend to have large effective population size (Ne) and very low β-GD among populations due to high levels of gene flow (Nielsen & Kenchington, 2001; Waples et al., 2008), even sometimes across oceans (Crandall et al., 2019). Consequently, one might expect weak β-SGDCs for marine organisms. Our results parallel those observed for North American freshwater fishes that are strongly limited in their dispersal across drainage basins (Fourtune et al., 2016; Robuchon et al., 2019) as well as for tenebrionid beetles in the Aegean Archipelago (Papadopoulou et al., 2011). In a marine context, the Monsoon Drift can be considered as a 'soft' barrier separating the Maldives from the southern part of the WIO region, since we identified the Maldives to be more genetically distant from the other southern locations in most of the studied species. In addition, we showed that both β-GD and β-SD patterns were mainly explained by this marine barrier effect, and not by isolation by distance. This suggests that the Monsoon Drift may restrict the long dispersal of individuals among reefs of the WIO, hence promoting both dissimilarity in allelic composition (β-GD) by limiting gene flow and dissimilarity in species composition (β-SD) by favouring allopatric (or parapatric) speciation (Riginos et al., 2014; Riginos & Nachman, 2001; Selkoe & Toonen, 2011). It is also worth noting that divergence times for most species belonging to the studied families (e.g., Gaboriau et al., 2018; Hodge et al., 2014) pre-date the period of establishment of the Monsoon Drift (circa 13 million years ago) in response to the onset of the Indian Monsoon (see Betzler et al., 2013). Overall, these findings add weight to previous results derived from neutral process-based models, which show that dispersal processes, by mediating vicariance, accurately predict α- and β-diversity patterns (Leprieur et al., 2016) as well as range size variation (Alzate et al., 2019) in tropical reef fishes. More generally, our results highlight the importance of dispersal barriers and neutral processes in the maintenance of both genetic and community structure, as expected by recent theoretical developments (see Baselga et al., 2022).

4.2 | Larval dispersal ability influences β-SGDCs

Our comparative approach allowed us to investigate differences in β-SGDCs across species. We highlight that the strength of β-SGDCs is inversely related to larval dispersal abilities, which means that species with low dispersal abilities showed the strongest correlation between β-GD and β-SD. This result is consistent with the fact that the β-GD is linked to PLD in this data set of WIO tropical fish species (Donati et al., 2021) as well as in the Hawaiian archipelago (Selkoe & Toonen, 2011). With the few *U*-crit values available in the literature (mainly at the family level), we found an even more marked relationship between the strength of β-SGDCs and dispersal abilities, which adds weight to recent findings that challenge the predominant role of passive dispersal in shaping population genetic structure and geographical range size in marine fishes (Nanninga & Manica, 2018). However, better addressing the question would require quantification of active dispersal for more species in future studies. Species with shorter PLD and/or low larval swimming capacity are thus more affected by the 'soft' barrier acting on both genetic and species levels, which leads to stronger β-SGDCs. If we were to consider only species with dispersal abilities that are strong enough to overcome this 'soft' barrier, or if there were no barrier, it would be very unlikely that such effects of dispersal traits over β-SGDCs would be found. Therefore, results of comparative SGDCs studies highly depend on the pool of species chosen.

Here, the detection of positive β-SGDCs relies on the presence of a 'soft' marine barrier, which causes the emergence of strong genetic and species spatial structure. This highlights that capturing the signal of spatial patterns of diversity requires sampling to be done at the appropriate (large enough) spatial scale, especially when studying species with long-distance dispersal (Dalongeville et al., 2018; Manel et al., 2019). The lack of spatial barrier might also explain why other studies do not always find the influence of PLD over genetic structure (Riginos et al., 2011; Riginos & Nachman, 2001; Weersing & Toonen, 2009).

4.3 | Contrasting lineage-based SGDCs among the three spatial components of diversity

Our lineage-based approach to studying SGDCs revealed contrasting results for the three components considered (α, β, γ). We found a positive and marked lineage-based association between β-GD and β-SD, while we found no evidence of this relationship when considering the α and γ components. This means that species with high levels of α- or γ-GD do not necessarily belong to species-rich families. This contrasting finding can be explained by the fact that both α- and γ-GD were found to be negatively associated with regional species abundance (a proxy of population size) for the 20 studied tropical reef fishes (Donati et al., 2021), which is not expected by the neutral theory of molecular evolution (see Romiguier et al., 2014). Such a deviation from neutral expectations implies that processes other than drift, mutation and migration influence the levels of α- and γ-GD differently than those of SD in tropical reef fishes. Adaptation, hybridization and reproductive success could be one of these processes (Donati et al., 2021). Another explanation is that the α- and γ-GD metric used in this study, namely the expected heterozygosity, does not reflect the whole genome diversity and more especially rates of molecular evolution. The use of metrics reflecting chromosomal rearrangement rates, such as the karyotype diversity, based on multiple species for each lineage would likely better uncover the link between genomic diversity, species diversification and ultimately lineage species richness (Martinez et al., 2017; Olmo, 2005).

4.4 | Limitations and perspectives

Our results highlight the importance of 'soft' barriers and neutral dispersal processes in shaping both β-GD and β-SD patterns

in tropical reef fishes of the WIO regions. In the future, a more explicit assessment of the influence of the North Monsoon Drift would require including other sampling sites within and across each side of this marine current. We could then test more specifically the influence of geographical barriers with statistical analyses such as spatial analysis of molecular variance (SAMOVA; Dupanloup et al., 2002), which has been extensively used in population genetics studies (e.g., Ma et al., 2018). Wider biogeographical gradients and other 'soft' barriers (see Cowman & Bellwood, 2013) could also be considered with data from the Red Sea, the Coral Triangle and the Central Pacific. This would permit our results to be generalized to the whole Indo-Pacific region, linking micro- and macroevolutionary processes to help explain the longitudinal gradient of species turnover in the tropical reef fish fauna of this region (see Cowman & Bellwood, 2013; Mouillot et al., 2013). To go beyond the role of neutral dispersal processes in shaping SGDCs, a larger sampling design would allow evaluation of whether ecological or adaptive drivers such as temperature, salinity, habitat heterogeneity and pH could also shape both GD and SD and in turn the strength of the SGDC. This would require the use of both neutral and non-neutral genetic markers (see Matala et al., 2014).

In this study, we explored SGDCs not only across communities as initially proposed by Vellend (2003) but also across lineages. This approach should be extended to other taxonomic groups in both terrestrial and marine environments in order to test the generality of our findings. This research objective is now possible with the development of large-scale databases of genetic diversity, which is the starting point of the emerging field of macrogenetics (Leigh et al., 2021).

5 | CONCLUSIONS

Our results reveal a strong association between genetic and species β-diversity patterns in tropical reef fishes of the Western Indian Ocean region, a result largely explained by the presence of a 'soft' barrier and mediated by larval dispersal processes. These findings suggest that dispersal limitation is a major process shaping β-diversity patterns from micro- to macroevolutionary scales in tropical reef fishes. Overall, this study emphasizes that accounting for the full complexity of spatial components (α, β, γ) and levels (intra-, interspecific) of diversity can substantially improve our understanding of underlying processes spanning from micro- to macroevolutionary scales.

ACKNOWLEDGMENTS

We thank Irthisham Zareer, Shameel Ibrahim (Maldives Whale Shark Research Programme), and Rodney Govinden (Seychelles Fishing Authority) for their precious help with sampling. We also thank Amin Abdallah, Luluesther Samwel, Lucas Simon, Eliad Lukuwi, Oscar Joseph Ngido, Peter Majengo and Ipyana Adamssony, Clara Belmont, Stephanie Marie, Rodney, Maria Cedras, Rodney Melanie, Ahmed Evaan, Séverine Albouy, Patrices Descombes, Loïc Chalmandrier, Nadine Sandau and all the crew of L'Amitié (SFA) and Hope Cruiser, as well as the Mafia Island Diving, in particular David, Danielle and Hamis Mjoc for support and assistance in the field. We are finally grateful to Claudia Michel, Silvia Kobel, Aria Minder, Camille Pitteloud and Livia Gerber for laboratory assistance, as well as to Niklaus Zemp for the bioinformatic analyses, and to the two anonymous referees for providing very constructive comments on the earlier version of this manuscript.

FUNDING INFORMATION

This work was financed by the Swiss National Science Foundation and Agence Nationale Recherche Project 'REEFISH' no° 310030E-164294. [Correction added on 13 February 2023, after first online publication: FUNDING INFORMATION has been updated.]

CONFLICT OF INTEREST

The authors declare no conflict of interest.

DATA AVAILABILITY STATEMENT

Scripts for this study are available on GitHub at https://github.com/mvilcot/reefish_WIO_SGDCs. SNP data and species presence data used in this study are available from the Dryad Digital Repository: https://doi.org/10.5061/dryad.kwh70rz7w.

ETHICS STATEMENT

We thank the local authorities for permits to collect samples and help with field logistics in the Maldives, Mayotte, Seychelles and Mafia Island. Research permits number are: Maldives ((OTHR)30-D/INDIV/2016/538), Mayotte (06/UTM/2016), Seychelles (A0157) and Tanzania (2017-242-NA-2017-87). All the genetic data produced and analysed in this paper have no commercial value and were generated in collaboration with the Genetic Diversity Centre (GDC), ETH Zurich.

ORCID

Maurine Vilcot https://orcid.org/0000-0001-8765-1364
Camille Albouy https://orcid.org/0000-0003-1629-2389
Giulia Francesca Azzurra Donati https://orcid.org/0000-0002-5954-3283
Thomas Claverie https://orcid.org/0000-0002-6258-4991
Stéphanie Manel https://orcid.org/0000-0001-8902-6052
Loïc Pellissier https://orcid.org/0000-0002-2289-8259

REFERENCES

Albouy, C., Archambault, P., Appeltans, W., Araújo, M. B., Beauchesne, D., Cazelles, K., Cirtwill, A. R., Fortin, M. J., Galiana, N., Leroux, S. J., Pellissier, L., Poisot, T., Stouffer, D. B., Wood, S. A., & Gravel, D. (2019). The marine fish food web is globally connected. *Nature Ecology & Evolution*, *3*, 1153–1161.

Alzate, A., Janzen, T., Bonte, D., Rosindell, J., & Etienne, R. S. (2019). A simple spatially explicit neutral model explains the range size distribution of reef fishes. *Global Ecology and Biogeography*, *28*, 875–890.

Antonovics, J. (1976). The input from population genetics: The new ecological genetics. *Systematic Botany*, *1*, 233–245.

Baselga, A. (2010). Partitioning the turnover and nestedness components of beta diversity: Partitioning beta diversity. *Global Ecology and Biogeography*, *19*, 134–143.

Baselga, A., Gómez-Rodríguez, C., Araújo, M. B., Castro-Insua, A., Arenas, M., Posada, D., & Vogler, A. P. (2022). Joint analysis of species and genetic variation to quantify the role of dispersal and environmental constraints in community turnover. *Ecography*, *2022*, e05808.

Baselga, A., Orme, D., Villeger, S., Bortoli, J. D., Leprieur, F., Logez, M., & Henriques-Silva, R. (2018). *betapart: Partitioning beta diversity into turnover and nestedness components*. R package version 1.5.5.

Betzler, C., Lüdmann, T., Hübscher, C., & Fürstenau, J. (2013). Current and sea-level signals in periplatform ooze (Neogene, Maldives, Indian Ocean). *Sedimentary Geology*, *290*, 126–137.

Bivand, R., & Rundel, C. (2019). *rgeos: Interface to geometry engine - open source ('GEOS')*. R package version 0.5-9.

Claverie, T., & Wainwright, P. C. (2014). A morphospace for reef fishes: Elongation is the dominant axis of body shape evolution. *PLoS One*, *9*, e112732.

Cowman, P. F., & Bellwood, D. R. (2013). Vicariance across major marine biogeographic barriers: Temporal concordance and the relative intensity of hard versus soft barriers. *Proceedings of the Royal Society B: Biological Sciences*, *280*, 20131541.

Crandall, E. D., Riginos, C., Bird, C. E., Liggins, L., Treml, E., Beger, M., Barber, P. H., Connolly, S. R., Cowman, P. F., DiBattista, J. D., Eble, J. A., Magnuson, S. F., Horne, J. B., Kochzius, M., Lessios, H. A., Liu, S. Y. V., Ludt, W. B., Madduppa, H., Pandolfi, J. M., … Gaither, M. R. (2019). The molecular biogeography of the Indo-Pacific: Testing hypotheses with multispecies genetic patterns. *Global Ecology and Biogeography*, *28*, 943–960.

Dalongeville, A., Andrello, M., Mouillot, D., Lobreaux, S., Fortin, M.-J., Lasram, F., Belmaker, J., Rocklin, D., & Manel, S. (2018). Geographic isolation and larval dispersal shape seascape genetic patterns differently according to spatial scale. *Evolutionary Applications*, *11*, 1437–1447.

Donati, G. F. A., Zemp, N., Manel, S., Poirier, M., Claverie, T., Ferraton, F., Gaboriau, T., Govinden, R., Hagen, O., Ibrahim, S., Mouillot, D., Leblond, J., Julius, P., Velez, L., Zareer, I., Ziyad, A., Leprieur, F., Albouy, C., & Pellissier, L. (2021). Species ecology explains the spatial components of genetic diversity in tropical reef fishes. *Proceedings of the Royal Society B: Biological Sciences*, *288*, 20211574.

Dupanloup, I., Schneider, S., & Excoffier, L. (2002). A simulated annealing approach to define the genetic structure of populations. *Molecular Ecology*, *11*, 2571–2581.

Ellison, A. M. (2010). Partitioning diversity. *Ecology*, *91*, 1962–1963.

Fisher, R., Leis, J. M., Clark, D. L., & Wilson, S. K. (2005). Critical swimming speeds of late-stage coral reef fish larvae: Variation within species, among species and between locations. *Marine Biology*, *147*, 1201–1212.

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*, 1830–1845.

Freckleton, R. P., Harvey, P. H., & Pagel, M. (2002). Phylogenetic analysis and comparative data: A test and review of evidence. *The American Naturalist*, *160*, 712–726.

Gaboriau, T., Leprieur, F., Mouillot, D., & Hubert, N. (2018). Influence of the geography of speciation on current patterns of coral reef fish biodiversity across the Indo-Pacific. *Ecography*, *41*, 1295–1306.

Gaggiotti, O. E., Chao, A., Peres-Neto, P., Chiu, C.-H., Edwards, C., Fortin, M.-J., Jost, L., Richards, C. M., & Selkoe, K. A. (2018). Diversity from genes to ecosystems: A unifying framework to study variation across biological metrics and scales. *Evolutionary Applications*, *11*, 1176–1193.

Goslee, S. C., & Urban, D. L. (2007). The ecodist package for dissimilarity-based analysis of ecological data. *Journal of Statistical Software*, *22*, 1–19.

Goudet, J., & Jombart, T. (2015). *hierfstat: Estimation and tests of hierarchical F-statistics*. R package version 0.5-10.

Hedrick, P. W. (2005). A standardized genetic differentiation measure. *Evolution*, *59*, 1633–1638.

Hodge, J. R., van Herwerden, L., & Bellwood, D. R. (2014). Temporal evolution of coral reef fishes: Global patterns and disparity in isolated locations. *Journal of Biogeography*, *41*, 2115–2127.

Hogan, J. D., Fisher, R., & Nolan, C. (2007). Critical swimming speed of settlement-stage coral reef fishes from the Caribbean: A methodological and geographical comparison. *Bulletin of Marine Science*, *80*, 219–231.

Hubbell, S. P. (2001). *The unified neutral theory of biodiversity and biogeography (MPB-32)*. Princeton University Press.

Jombart, T. (2008). Adegenet: A R package for the multivariate analysis of genetic markers. *Bioinformatics*, *24*, 1403–1405.

Jombart, T., & Ahmed, I. (2011). adegenet 1.3–1: New tools for the analysis of genome-wide SNP data. *Bioinformatics*, *27*, 3070–3071.

Jombart, T., Devillard, S., & Balloux, F. (2010). Discriminant analysis of principal components: A new method for the analysis of genetically structured populations. *BMC Genetics*, *11*, 94.

Kulbicki, M., Parravicini, V., Bellwood, D. R., Arias-Gonzàlez, E., Chabanet, P., Floeter, S. R., Friedlander, A., McPherson, J., Myers, R. E., Vigliola, L., & Mouillot, D. (2013). Global biogeography of reef fishes: A hierarchical quantitative delineation of regions. *PLoS One*, *8*, e81847.

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*, 759–771.

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., & Borcard, D. (2015). Should the mantel test be used in spatial analysis? *Methods in Ecology and Evolution*, *6*, 1239–1247.

Legendre, P., & Legendre, L. (1998). *Numerical ecology*. Elsevier Science.

Leigh, D. M., van Rees, C. B., Millette, K. L., Breed, M. F., Schmidt, C., Bertola, L. D., Hand, B. K., Hunter, M. E., Jensen, E. L., Kershaw, F., Liggins, L., Luikart, G., Manel, S., Mergeay, J., Miller, J. M., Segelbacher, G., Hoban, S., & Paz-Vinas, I. (2021). Opportunities and challenges of macrogenetic studies. *Nature Reviews Genetics*, *22*, 791–807.

Leprieur, F., Descombes, P., Gaboriau, T., Cowman, P. F., Parravicini, V., Kulbicki, M., Melián, C. J., de Santana, C. N., Heine, C., Mouillot, D., Bellwood, D. R., & Pellissier, L. (2016). Plate tectonics drive tropical reef biodiversity dynamics. *Nature Communications*, *7*, 11461.

Lino, A., Ferreira, E., Fonseca, C., Fischer, E., & Ramos Pereira, M. J. (2021). Species–genetic diversity correlation in phyllostomid bats of the Bodoquena plateau, Brazil. *Biodiversity and Conservation*, *30*, 403–429.

Luiz, O. J., Allen, A. P., Robertson, D. R., Floeter, S. R., Kulbicki, M., Vigliola, L., Becheler, R., & Madin, J. S. (2013). Adult and larval traits as determinants of geographic range size among tropical reef fishes. *Proceedings of the National Academy of Sciences*, *110*, 16498–16502.

Ma, K. Y., van Herwerden, L., Newman, S. J., Berumen, M. L., Choat, J. H., Chu, K. H., & Sadovy de Mitcheson, Y. (2018). Contrasting population genetic structure in three aggregating groupers (Percoidei: Epinephelidae) in the Indo-West Pacific: The importance of reproductive mode. *BMC Evolutionary Biology*, *18*, 180.

Manel, S., Guerin, P.-E., Mouillot, D., Blanchet, S., Velez, L., Albouy, C., & Pellissier, L. (2020). Global determinants of freshwater and marine fish genetic diversity. *Nature Communications*, *11*, 692.

Manel, S., Loiseau, N., Andrello, M., Fietz, K., Goñi, R., Forcada, A., Lenfant, P., Kininmonth, S., Marcos, C., Marques, V., Mallol, S., Pérez-Ruzafa, A., Breusing, C., Puebla, O., & Mouillot, D. (2019).

Long-distance benefits of marine reserves: Myth or reality? *Trends in Ecology & Evolution*, *34*, 342–354.

Martinez, P. A., Jacobina, U. P., Fernandes, R. V., Brito, C., Penone, C., Amado, T. F., Fonseca, C. R., & Bidau, C. J. (2017). A comparative study on karyotypic diversification rate in mammals. *Heredity*, *118*, 366–373.

Matala, A. P., Ackerman, M. W., Campbell, M. R., & Narum, S. R. (2014). Relative contributions of neutral and non-neutral genetic differentiation to inform conservation of steelhead trout across highly variable landscapes. *Evolutionary Applications*, *7*, 682–701.

Meirmans, P. G., & Hedrick, P. W. (2011). Assessing population structure: FST and related measures. *Molecular Ecology Resources*, *11*, 5–18.

Messmer, V., Jones, G. P., Munday, P. L., & Planes, S. (2012). Concordance between genetic and species diversity in coral reef fishes across the Pacific Ocean biodiversity gradient. *Evolution*, *66*, 3902–3917.

Moragues, M., Comadran, J., Waugh, R., Milne, I., Flavell, A. J., & Russell, J. R. (2010). Effects of ascertainment bias and marker number on estimations of barley diversity from high-throughput SNP genotype data. *Theoretical and Applied Genetics*, *120*, 1525–1534.

Mouillot, D., De Bortoli, J., Leprieur, F., Parravicini, V., Kulbicki, M., & Bellwood, D. R. (2013). The challenge of delineating biogeographical regions: Nestedness matters for Indo-Pacific coral reef fishes. *Journal of Biogeography*, *40*, 2228–2237.

Nanninga, G., & Manica, A. (2018). Larval swimming capacities affect genetic differentiation and range size in demersal marine fishes. *Marine Ecology Progress Series*, *589*, 1–12.

Nei, M. (1987). *Molecular evolutionary genetics*. Columbia University Press.

Nielsen, E. E., & Kenchington, E. (2001). A new approach to prioritizing marine fish and shellfish populations for conservation. *Fish and Fisheries*, *2*, 328–343.

Oksanen, J., Blanchet, F. G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., Minchin, P. R., O'Hara, R. B., Simpson, G. L., Solymos, P., Stevens, M. H. H., Szoecs, E., & Wagner, H. (2019). vegan: Community ecology package.

Olmo, E. (2005). Rate of chromosome changes and speciation in reptiles. *Genetica*, *125*, 185–203.

Pante, E., & Simon-Bouhet, B. (2013). Marmap: A package for importing, plotting and analyzing bathymetric and topographic data in R. *PLoS One*, *8*, e73051.

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*, 241–255.

Parravicini, V., Kulbicki, M., Bellwood, D. R., Friedlander, A. M., Arias-Gonzalez, J. E., Chabanet, P., Floeter, S. R., Myers, R., Vigliola, L., D'Agata, S., & Mouillot, D. (2013). Global patterns and predictors of tropical reef fish species richness. *Ecography*, *36*, 1254–1262.

Pebesma, E. (2018). Simple features for R: Standardized support for spatial vector data. *The R Journal*, *10*, 439.

Pellissier, L., Leprieur, F., Parravicini, V., Cowman, P. F., Kulbicki, M., Litsios, G., Olsen, S. M., Wisz, M. S., Bellwood, D. R., & Mouillot, D. (2014). Quaternary coral reef refugia preserved fish diversity. *Science*, *344*, 1016–1019.

Peres-Neto, P. R., & Jackson, D. A. (2001). How well do multivariate data sets match? The advantages of a procrustean superimposition approach over the mantel test. *Oecologia*, *129*, 169–178.

Pinheiro, J., Bates, D., Deb Roy, S., Sarkar, D., & R Core Team. (2021). nlme: Linear and nonlinear mixed effects models.

Price, S. A., Holzman, R., Near, T. J., & Wainwright, P. C. (2011). Coral reefs promote the evolution of morphological diversity and ecological novelty in labrid fishes. *Ecology Letters*, *14*, 462–469.

R Core Team. (2021). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing.

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

Raufaste, N., & Rousset, F. (2001). Are partial man^{TEL} tests adequate? *Evolution*, *55*, 1703–1705.

Riginos, C., Buckley, Y. M., Blomberg, S. P., & Treml, E. A. (2014). Dispersal capacity predicts both population genetic structure and species richness in reef fishes. *The American Naturalist*, *184*, 52–64.

Riginos, C., Douglas, K. E., Jin, Y., Shanahan, D. F., & Treml, E. A. (2011). Effects of geography and life history traits on genetic differentiation in benthic marine fishes. *Ecography*, *34*, 566–575.

Riginos, C., & Nachman, M. W. (2001). Population subdivision in marine environments: The contributions of biogeography, geographical distance and discontinuous habitat to genetic differentiation in a blennioid fish, *Axoclinus nigricaudus*. *Molecular Ecology*, *10*, 1439–1453.

Robuchon, M., Leroy, B., Jézéquel, C., & Hugueny, B. (2019). Correlations between broad-scale taxonomic and genetic differentiations suggest a dominant imprint of historical processes on beta diversities. *Journal of Biogeography*, *46*, 1083–1095.

Romiguier, J., Gayral, P., Ballenghien, M., Bernard, A., Cahais, V., Chenuil, A., Chiari, Y., Dernat, R., Duret, L., Faivre, N., Loire, E., Lourenco, J. M., Nabholz, B., Roux, C., Tsagkogeorga, G., Weber, A. A.-T., Weinert, L. A., Belkhir, K., Bierne, N., … Galtier, N. (2014). Comparative population genomics in animals uncovers the determinants of genetic diversity. *Nature*, *515*, 261–263.

Schott, F. A., & McCreary, J. P. (2001). The monsoon circulation of the Indian Ocean. *Progress in Oceanography*, *51*, 1–123.

Selkoe, K. A., & Toonen, R. J. (2011). Marine connectivity: A new look at pelagic larval duration and genetic metrics of dispersal. *Marine Ecology Progress Series*, *436*, 291–305.

Symonds, M. R. E., & Blomberg, S. P. (2014). A primer on phylogenetic generalised least squares. In L. Z. Garamszegi (Ed.), *Modern phylogenetic comparative methods and their application in evolutionary biology* (pp. 105–130). Springer Berlin Heidelberg.

Tuomisto, H. (2010). A diversity of beta diversities: Straightening up a concept gone awry. Part 1. Defining beta diversity as a function of alpha and gamma diversity. *Ecography*, *33*, 2–22.

Vellend, M. (2002). A pest and an invader: White-tailed deer (*Odocoileus virginianus* Zimm.) as a seed dispersal agent for honeysuckle shrubs (Lonicera L.). *Natural Areas Journal*, *22*, 230–234.

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. (2010). Conceptual synthesis in community ecology. *The Quarterly Review of Biology*, *85*, 183–206.

Vellend, M., & Geber, M. A. (2005). Connections between species diversity and genetic diversity. *Ecology Letters*, *8*, 767–781.

Waples, R. S., Punt, A. E., & Cope, J. M. (2008). Integrating genetic data into management of marine resources: How can we do it better? *Fish and Fisheries*, *9*, 423–449.

Weersing, K., & Toonen, R. J. (2009). Population genetics, larval dispersal, and connectivity in marine systems. *Marine Ecology Progress Series*, *393*, 1–12.

Wickham, H. (2016). *Ggplot2 elegant graphics for data analysis*. Springer-Verlag.

Winter, D. J. (2012). Mmod: An R library for the calculation of population differentiation statistics. *Molecular Ecology Resources*, *12*, 1158–1160.

BIOSKETCH

Maurine Vilcot is a Ph.D. student at the Centre d'Ecologie Fonctionnelle & Evolutive (CEFE) in Montpellier (France). Her research lies between macroecology, population genomics and molecular evolution. She focuses on understanding processes generating patterns of diversity across different evolutionary scales, from genomes to intra- and interspecific scales.

SUPPORTING INFORMATION

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

How to cite this article: Vilcot, M., Albouy, C., Donati, G. F. A., Claverie, T., Julius, P., Manel, S., Pellissier, L., & Leprieur, F. (2023). Spatial genetic differentiation correlates with species assemblage turnover across tropical reef fish lineages. *Global Ecology and Biogeography*, *32*, 535–547. https://doi.org/10.1111/geb.13637