

## Hierarchical partitioning of fish diversity and scale-dependent environmental effects in tropical coastal ecosystems

Ellen Martins Camara, Márcia Cristina Costa de Azevedo, Taynara Pontes Franco, Francisco Gerson Araújo\*

Universidade Federal Rural do Rio de Janeiro, Departamento de Biologia Animal, Laboratório de Ecologia de Peixes, 23897-030, Seropédica, RJ, Brazil



### ARTICLE INFO

#### Keywords:

Hierarchical spatial structure  
Scale dependence  
Additive partitioning  
Multiplicative partitioning  
Multi-scale environmental effects  
Land use  
Isolation  
Fish  
Coastal zone

### ABSTRACT

The spatial structure of the fish diversity and site-scale and landscape-scale environmental effects were investigated across hierarchical levels in tropical coastal ecosystems. Total diversity ( $\gamma$ ) was hierarchically partitioned into  $\alpha$  and  $\beta$  components using both the additive and multiplicative methods. A model selection based on the AICc was applied to generalized linear mixed models relating diversity measures to environmental variables and including random effects for hierarchical levels and season. Short-term seasonal effects were negligible. Spatial effects were more relevant at the site level and negligible at the subregion level, due to the high spatial heterogeneity and the natural pooling of ecosystems, respectively. Site-scale environmental effects were more relevant at the subregion level, with eutrophic conditions (continental influence) favoring the species richness ( $\alpha$  and  $\gamma$ ) and higher absence of species ( $\beta^A$ ) in oligotrophic conditions (marine influence). At the system level, the positive effect of the distance from the ocean on  $\gamma$  and higher  $\beta^A$  in oligotrophic conditions reinforced the positive continental influence on fish diversity. Environmental homogenization processes were most likely associated with the negative effect of the pasture cover on  $\alpha$  at the system level, and  $\gamma$  and  $\beta^A$  at the site level. The negative effect of the forest cover on the later diversity measure evidenced its relevance to maintain richer but more similar assemblages, whereas the positive continental influence was most likely due to the loss of stenohaline marine species. This study evidenced that disentangling spatial, land use, and marine vs. continental effects on diversity components is critical to understand the primary determinants of the fish diversity in tropical coastal ecosystems.

### 1. Introduction

Coastal ecosystems provide critical natural services, including food provision, coastal protection, ocean nourishment, life cycle maintenance for aquatic populations, and recreation and tourism (Liquete et al., 2013). There are substantial evidences that the delivery of ecosystem services is strongly dependent of the biodiversity, which makes coastal ecosystems key areas for global-level biodiversity conservation (Cardinale et al., 2012; Annis et al., 2017). Marine ecosystems worldwide have been long affected by increasing cumulative human impacts, with recent increases more expressive in tropical and subtropical coastal regions, especially in countries with greater growth of coastal population (Halpern et al., 2015). The deleterious impacts of humans on marine organisms, leading to major changes in food webs and ecosystem functioning, are largely recognized (McCauley et al., 2015). In this sense, the increasingly threats to the mega-diverse fish fauna of South America due to land use changes and urbanization (Reis et al.,

2016) make fish species inhabiting coastal areas primary components of that scenario. Therefore, understanding the major environmental determinants of the fish diversity patterns is necessary to set priorities regarding the management of human activities that affect such areas.

Diversity patterns are frequently spatially structured as a result of stochastic processes, historical contingencies, and contemporary factors in broader scales that influence the species ranges and ultimately the regional pool of species (Wang et al., 2015; Reis et al., 2016; Araújo et al., 2018). Local environmental conditions are deemed to select species for assemblages from the regional pool, which is restricted by factors at larger spatial scales (Ricklefs, 1987; Vilar et al., 2013). Therefore, the diversity patterns are a result of the balance between processes acting in multiple and hierarchical spatial and temporal scales (Ricklefs, 2004). That balance is represented by marine and continental influences in coastal ecosystems, since they constitute transitional environments between the ocean, land, and river basins (Pérez-Ruzafa et al., 2011). The relative importance of continental and

\* Corresponding author.

E-mail address: [gerson@ufrj.br](mailto:gerson@ufrj.br) (F.G. Araújo).

<https://doi.org/10.1016/j.marenvres.2019.05.006>

Received 25 February 2019; Received in revised form 23 April 2019; Accepted 3 May 2019

Available online 04 May 2019

0141-1136/ © 2019 Elsevier Ltd. All rights reserved.

marine influences depends on the type of coastal ecosystem, primarily as a result of the degree of isolation from the ocean and the freshwater inflow (Azevedo et al., 2017; Camara et al., 2018). A better understanding of the fish diversity patterns in coastal ecosystems, therefore, depends on disentangling the hierarchical structure of continental and marine environmental effects across different spatial scales.

Species richness, herein representing the number of species per sampling unit, is frequently used as a surrogate for diversity in studies considering large geographical areas (e.g., Tittensor et al., 2010; Huang et al., 2017) and multiple spatial scales (e.g., Francisco-Ramos and Arias-González, 2013; Pasquaud et al., 2015). Despite the limited representation of the complexity inherent to natural assemblages, several studies rely on the species richness because it is a fundamental component of the diversity and a more easily obtained measure for large scales than diversity measures based on the relative abundance of species (Wilsey et al., 2005). The partitioning of the total diversity (gamma;  $\gamma$ ) into alpha ( $\alpha$ ) and beta ( $\beta$ ) components has been increasingly used to quantify patterns across hierarchical spatial scales (e.g., Francisco-Ramos and Arias-González, 2013; Valencia-Méndez et al., 2018). A better understanding of spatially-structured diversity patterns may be achieved by accessing the relative contributions of the  $\alpha$  (e.g., the species richness within sites) and  $\beta$  (e.g., the difference in the species richness between sites) components across hierarchical levels (Zhang et al., 2014; Valencia-Méndez et al., 2018). The diversity partitioning may be additive, with  $\beta$  expressing the number of species absent from a hierarchical level, or multiplicative, with  $\beta$  expressing the number of distinct assemblages at a hierarchical level (Crist et al., 2003; Jost, 2007). Both methods of diversity partitioning are relevant to investigate diversity patterns across different spatial scales because they emphasize different properties of species data (Anderson et al., 2011).

Other relevant point regarding investigations based on multiple spatial scales is that they must rely on statistical methods that explicitly account for the spatial structure of environmental data, disentangling spatial and environmental effects (Fortin et al., 2012). Generalized linear mixed models (GLMMs) can control the spatial and/or temporal dependence of data using random effects, producing more accurate estimates of the species-environment relationships (Gelman and Hill, 2007). In addition, GLMMs are flexible methods to analyze non-normal data, overcoming common problems, such as pseudo-replication and heteroscedastic variance, by allowing for random effects and response variables from different distributions (Bolker et al., 2009). Concurrently, model selection approaches based on the information theory are useful alternatives for the traditional hypothesis-testing approaches, which often perform poorly for observational studies (Burnham and Anderson, 2002). The information-theoretic approach is focused on the selection of a best model or several parsimonious models given the data and the set of a priori models representing the scientific hypotheses of interest (Burnham and Anderson, 2002). That is a different paradigm than the search of a true model based on the acceptance or rejection of null hypotheses by statistical tests and significance levels (Burnham and Anderson, 2001). Information-theoretic approaches are, therefore, more effective and informative methods for statistical inferences in ecological studies, in which multiple relative effects are usually observed (Burnham and Anderson, 2001). Therefore, the investigation of multi-scale environmental effects on fish diversity patterns in coastal ecosystems may strongly benefit from the use of GLMMs and information-theoretic approaches.

This study evaluated multi-scale environmental effects on diversity patterns of fish assemblages in different types of coastal ecosystems (i.e. coastal lagoons, bays, and oceanic beaches) of the Southeastern Brazil. These three types of coastal ecosystems have different environmental conditions and degrees of connectivity with marine waters: (1) oceanic beaches are entirely connected with marine waters and have more habitat homogeneity, stable salinities and high physical dynamism due to wave exposure; (2) bays are partially connected with marine waters and have greater habitat diversity, slight salinity gradients and great

tidal influences; and (3) coastal lagoons have limited connections with marine waters, suitable conditions for trapping matter and energy, thus increasing productivity, stable salinity gradients, and well-protected conditions. Despite the expressive local environmental differences, these coastal ecosystems share several biogeographical and physiographical features, and large-scale environmental conditions (Petty et al., 2016; Araújo et al., 2018). Therefore, environmental and diversity patterns are most likely spatially structured and scale-dependent.

We accessed the relative effects of environmental variables measured at site and landscape scales on the diversity patterns at four hierarchical spatial scales, herein named site, system, subregion, and region levels. Gamma diversity was partitioned into  $\alpha$  and  $\beta$  components across the four hierarchical levels to better investigate the mechanisms determining the diversity patterns. The relationships between the diversity measures and environmental variables were modeled by GLMMs, including the hierarchical level as a random effect. We hypothesized that due to mechanisms associated with environmental filters and the degree of isolation: (i) the site-scale environmental conditions are most likely more relevant to determine the diversity patterns at the site level and the land use at the higher hierarchical levels; and (ii) increasing isolation most likely determine decreases in the  $\alpha$  component and increases in the  $\beta$  component at all hierarchical levels. Following these scientific hypotheses and using a model selection approach based on the information theory, this study look forward to select parsimonious models as a basis for statistical inferences regarding the primary determinants of the fish diversity patterns in tropical coastal ecosystems.

## 2. Materials and methods

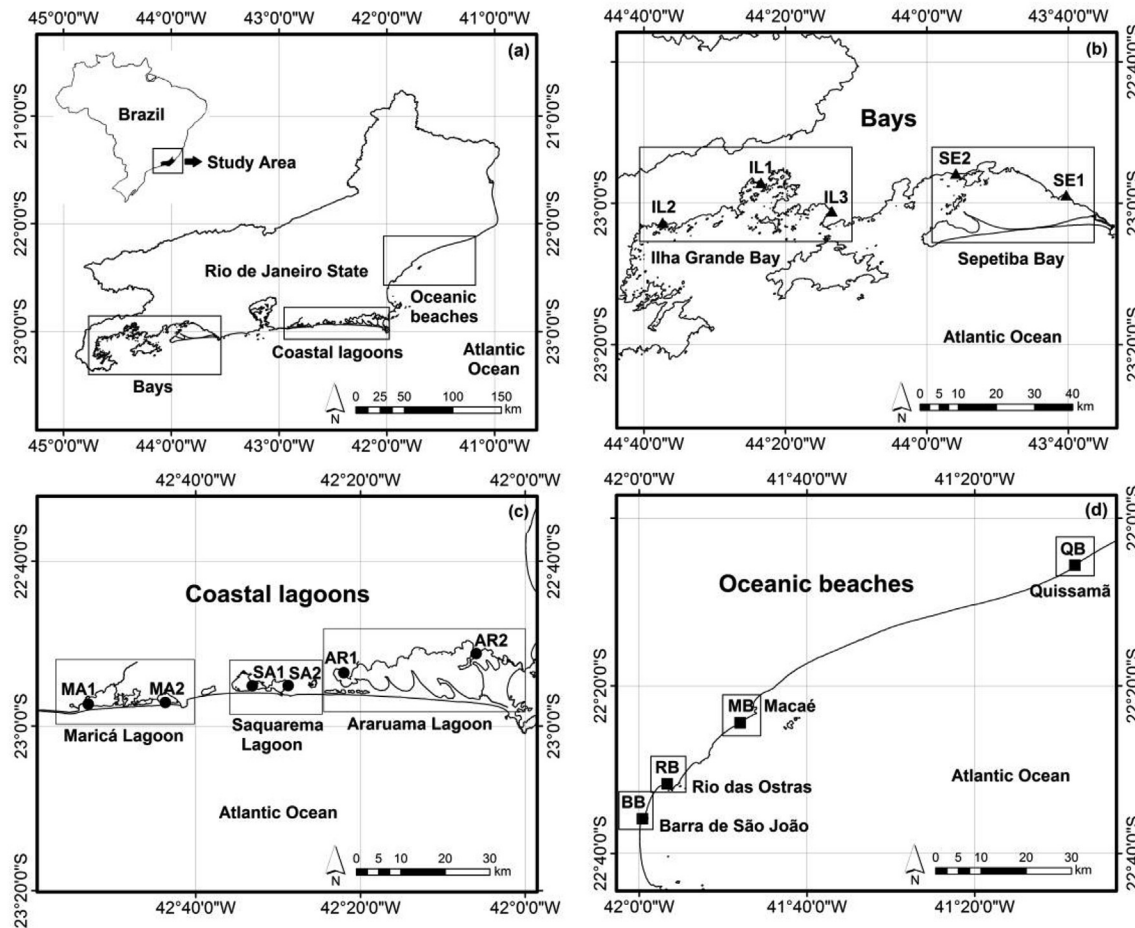
### 2.1. Study area

The study area comprises different coastal ecosystems located along the coast of the Rio de Janeiro State, Southeastern Brazil (Fig. 1a). Climate is tropical, with annual mean temperature of 22 °C and annual total rainfall ranging from 1000 to 1600 mm (Alvares et al., 2013). Rainfall is primarily concentrated from October to March (wet season) compared with April to September (dry season) (Alvares et al., 2013). The coastal ecosystems present several environmental differences, mainly as a result of the different degrees of connection with the ocean, continental influence, and human activities (Azevedo et al., 2017; Franco et al., 2019).

Each type of ecosystem is located at a different subregion, creating natural hierarchical levels (Fig. 1a), as follows: one subregion comprises large bays, represented by two bay ecosystems, the Ilha Grande bay, with three sites (IL1, IL2, and IL3), and the Sepetiba bay, with two sites (SE1 and SE2) (Fig. 1b); other subregion comprises coastal lagoons with permanent connections with the ocean and minimum freshwater input (Franco et al., 2019), represented by three coastal lagoon ecosystems with two sites each, the Maricá lagoon (MA1 and MA2), the Araruama lagoon (AR1 and AR2), and the Saquarema lagoon (SA1 and AR2) (Fig. 1c); and a third subregion comprises oceanic beaches, represented by four beaches with one site each, the Barra de São João beach (BB), the Rio das Ostras beach (RB), the Macaé beach (MB), and the Quissamã beach (QB) (Fig. 1d).

### 2.2. Fish sampling

Fish sampling was performed during the 2011 dry season and 2012 wet season across the geographical range encompassed by the study ( $\pm 400$  km). A total of 72 samples were obtained at the sampling locations (Fig. 2), as follows: 36 samples in coastal lagoons (3 sampling locations x 2 sites x 2 seasons x 3 coastal lagoons); 16 samples in oceanic beaches (2 sampling locations x 1 site x 2 seasons x 4 oceanic beaches); and 20 samples in bays, 12 in the Ilha Grande bay (2 sampling



**Fig. 1.** Location of the (a) coastal ecosystems (bays, coastal lagoons, and oceanic beaches) within the study area in Rio de Janeiro State. Location of the sites within the (b) bay ecosystems (Ilha Grande Bay and Sepetiba Bay), (c) coastal lagoon ecosystems (Maricá Lagoon, Saquarema Lagoon, and Araruama Lagoon), and oceanic beach ecosystems (Barra de São João, Rio das Ostras, Macaé, and Quissamã).

locations  $\times$  3 sites  $\times$  2 seasons) and 8 in the Sepetiba Bay (2 sampling locations  $\times$  2 sites  $\times$  2 seasons). The unbalanced sampling design aimed to properly consider the environmental heterogeneity intrinsic to each type of ecosystem.

Our study was focused on nearshore fish assemblages inhabiting shallow and semi-enclosed areas in unconsolidated substrate. For that reason, fishes were collected with a beach seine (12  $\times$  2.5 m; 5-mm mesh size) set parallel to the shore at approximately 1.5-m depth, dragging perpendicular to the shore (30-m long) for about 15 min and covering a swept area of approximately 300 m<sup>2</sup>. A total of 8 replicates were performed at each site per season, 2–3 replicates per sampling location in coastal lagoons and 4 replicates per sampling location in bays and oceanic beaches. The collected fishes were fixed in 10% formalin, and after 48 h, preserved in 70% ethanol. All fishes were identified at the species level, and vouchers specimens were deposited in the Ichthyological Collection of the Laboratório de Ecologia de Peixes of the Universidade Federal Rural do Rio de Janeiro. Given the unbalanced sampling design, with different numbers of sampling locations per site and sites per system, and the unequal number of replicates per sampling location in coastal lagoons, estimates of the diversity patterns were corrected by the sampling effort, as described in the “2.3. Diversity measures” subsection.

### 2.3. Diversity measures

The diversity patterns were investigated at four hierarchical levels, with the respective samples pooled as follows: (1) site level, with 15

sites per season (Fig. 1b–d); (2) system level, including the five original coastal lagoon and bay ecosystems and the four oceanic beaches (pooled within two different systems), totaling seven systems per season (Fig. 1b–c); (3) subregion level, the three types of ecosystem (i.e. coastal lagoon, bay, and oceanic beach) in each season (Fig. 1a); and (4) region level, the entire study area in each season (Fig. 1a). As each oceanic beach was represented by a single site, the four oceanic beaches were pooled within the system level as follows: one system comprised BB and RB, which are separated by a distance of about 8 km and located in highly urbanized areas; the other system comprised MB and QB, which are closer from a conservation area with several coastal lagoons presenting expressive freshwater inputs, and temporary and sporadic connections with the sea (Di Dario et al., 2013), despite subjected to different urbanization pressures and 70 km from each other (Fig. 1d).

The species richness was used as a surrogate for diversity because it is less sensible to possible differences in the sampling efficiency between different types of ecosystem than diversity measures including the species abundance, and to favor comparisons with studies developed in different spatial scales that frequently rely on that measure. For each season (wet; dry), rarefaction curves based on the number of species per number of individuals were used to evaluate if the observed species richness was representative of species pool in the study area. This method avoids possible underestimates of the species richness as a result of the spatial or temporal autocorrelation causing the non-random occurrence of species among samples (Gotelli and Colwell, 2001). According to individual-based rarefaction curves, the regional pool of species was effectively estimated for both the dry and wet

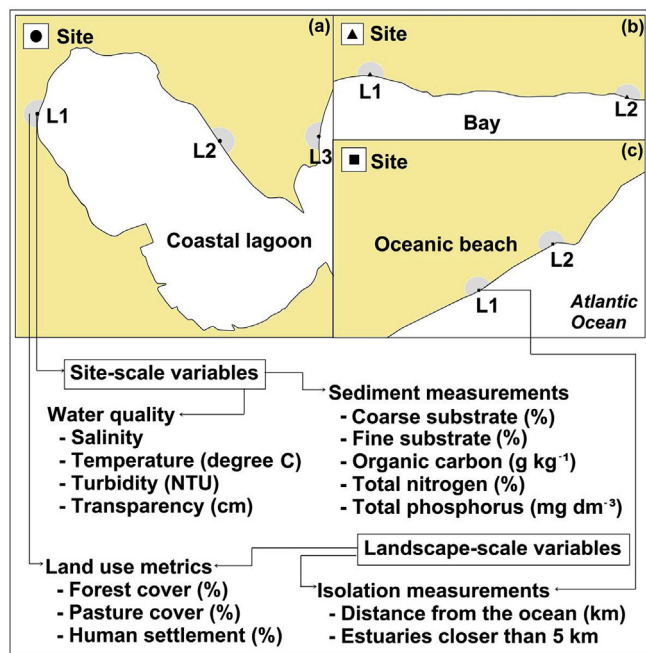


Fig. 2. Conceptual illustration representing the categories of environmental variables measured at the site and landscape scales in the different types of coastal ecosystems. Site-scale environmental variables and isolation measurements were obtained at three sampling locations (L1, L2, and L3) in the (a) coastal lagoons, and at two sampling locations (L1 and L2) in (b) bays and (c) oceanic beaches. For each site, land use metrics were obtained within a 200-m radius buffer (grey semi-circle) at each sampling location, totaling three measures per coastal lagoon and two measures per bay or coastal lagoon.

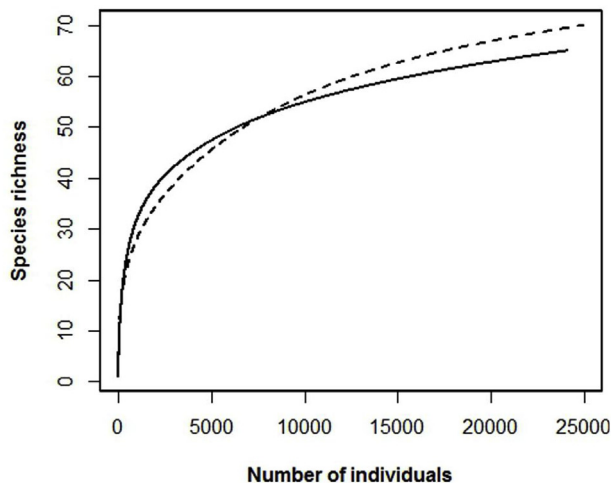


Fig. 3. Individual-based rarefaction curves representing the expected species richness in random re-samplings of all individuals collected in the study area per season. Dashed and solid lines represent the samples obtained during the dry and wet seasons, respectively.

seasons, with no relevant seasonal differences (Fig. 3). For both seasons, despite the species richness is still increasing, the curve is relatively near from reach an asymptote and the slight increase most likely reflects the environmental instability in oceanic beaches, where the fish assemblages are more susceptible to the effects of tides, waves, and oceanic currents. The rarefaction curves representing the expected species richness in random re-samplings of all individuals collected in the study area were based on Hurlbert's (1971) formulation and performed in the R environment (version 3.5.2; R Core Team, 2018), using the *vegan* package (version 2.5–3; Oksanen et al., 2018).

For each hierarchical level and season, the gamma diversity ( $\gamma$ ) was partitioned into ( $\alpha$ ) alpha and beta components (Zhang et al., 2014). The  $\alpha$  component for a sample in a given hierarchical level was calculated as the sum of the number of species times the number of replicates in each sampling unit (from the immediately inferior hierarchical level) pooled within the sample divided by the sum of the replicates. Therefore, the  $\alpha$  component represented the weighted mean number of species per sampling location within a site ( $\alpha_{site}$ ), per site within a system ( $\alpha_{sys}$ ), per system within a subregion ( $\alpha_{sub}$ ) or per subregion within the region ( $\alpha_{reg}$ ). For each hierarchical level, the  $\beta$  component was calculated based on the additive and multiplicative methods of diversity partitioning (Crist et al., 2003; Jost, 2007). In both cases, the  $\gamma$  diversity at the region level ( $\gamma_{reg}$ ) was considered as the species richness of all samples pooled in the study area per season.

In the additive partitioning, the  $\gamma$  diversity at a given hierarchical level was considered as the  $\alpha$  component at the immediately superior level. Therefore, the  $\alpha_{sys}$  corresponds to the  $\gamma$  diversity at the site level ( $\gamma_{site}$ ), the  $\alpha_{sub}$  corresponds to the  $\gamma$  diversity at the system level ( $\gamma_{sys}$ ), and the  $\alpha_{reg}$  corresponds to the  $\gamma$  diversity at the subregion level ( $\gamma_{sub}$ ). The  $\beta$  component was calculated as the mean number of species absent from a random sampling location within a site ( $\beta^A_{site}$ ), a random site within a system ( $\beta^A_{sys}$ ), a random system within a subregion ( $\beta^A_{sub}$ ) or a random subregion within the region ( $\beta^A_{reg}$ ), as follows:

$$\beta^A_{reg} = \gamma_{reg} - \alpha_{reg}$$

and

$$\beta^A_{level} = \alpha_{level+1} - \alpha_{level}$$

In the multiplicative partitioning, the  $\gamma$  diversity was considered as the total species richness at a site ( $\gamma_{site}$ ), system ( $\gamma_{sys}$ ), and subregion ( $\gamma_{sub}$ ). The  $\beta$  component was calculated as the number of assemblages that do not share species within a site ( $\beta^M_{site}$ ), system ( $\beta^M_{sys}$ ), subregion ( $\beta^M_{sub}$ ), and region ( $\beta^M_{reg}$ ), as follows:

$$B^M_{level} = \gamma_{level} / \alpha_{level}$$

#### 2.4. Environmental measures

The environmental measures, herein considered as the variables representative of the physical and chemical conditions at the site scale and the types of land use and vegetal cover, the marine influence, and the availability of estuarine habitats at the landscape scale, were chosen based on their relevance for assemblages of coastal fishes according to studies encompassing the same or similar ecosystems and comparable spatial and temporal scales at the study area (e.g., Franco et al., 2019), in broader regions (e.g., Araújo et al., 2018), and worldwide (e.g., Teichert et al., 2018). The validity of the spot measures at the site scale is supported by the spatial and temporal replication of the sampling design and the effectiveness of the statistical approach to avoid the occurrence of spurious relationships, as described in the “2.5. Data analysis” subsection.

##### 2.4.1. Site-scale variables

Environmental measurements of water quality and sediment were recorded concurrently with the fish sampling (Fig. 2). Variables representative of water quality included temperature (degree C), salinity, turbidity (Nephelometric Turbidity Units – NTU), and transparency (cm). Temperature, salinity, and turbidity were taken with a Horiba U-50 multiprobe (Horiba Trading Co. Ltd., Shanghai) immersed approximately 0.5 m under the water surface. The transparency was measured with a Secchi disk.

Sediment measurements included granulometric parameters and nutrient concentrations. For the granulometric and nutrient analysis, sediment samples were collected using a PVC corer (10 cm in diameter and 50 cm in length) in a collecting area of 0.00785 m<sup>2</sup> at a depth of

15 cm using 4 replicates. The collected sediment was weighed (precision of 0.01 g) and dried at 80 °C in a stove. A portion of the sediment (300 g) was used in the granulometric analyses, and another part (150 g) was used in the nutrient analyses. The granulometric parameters were calculated according to Folk and Ward (1957) and classified according to Shepard (1954). The mean granule size was determined from each granulometric fraction weight retained in each sieve using the software SysGran 3.0 (Camargo, 2006). The silt and clay fractions were grouped together. We considered as coarse substrate (%) the sum of gravel, very coarse sand, and coarse sand, and as fine substrate (%) the sum of medium sand, fine sand, very fine sand and clay + silt.

The concentrations of the following nutrients in the sediment were analyzed: organic carbon ( $\text{g. kg}^{-1}$ ), total nitrogen (%) and total phosphorus ( $\text{mg. dm}^{-3}$ ). The concentration of organic carbon was determined using the method of Walkley and Black (1934). The organic carbon was measured by the oxidation of wet organic matter with potassium dichromate in sulfuric acid medium, employing the heat given off from the sulfuric acid and/or applied heat as the energy source. The excess dichromate after oxidation was titrated with a standard solution of ammoniacal ferrous sulfate (Mohr salt). The concentration of total nitrogen in the sediment was determined using the Kjeldahl nitrogen method with a diffusion camera. The total nitrogen was converted to ammonium sulfate by oxidation with a mixture of  $\text{CuSO}_4$ ,  $\text{H}_2\text{SO}_4$  and  $\text{Na}_2\text{SO}_4$  or  $\text{K}_2\text{SO}_4$  (mineralization). Later, in alkaline medium, the ammonium sulfate converted from the organic matter released ammonia, which was complexed in a boric acid solution containing a mixed indicator in a diffusion chamber and was finally determined using acidimetry ( $\text{H}_2\text{SO}_4$  or  $\text{HCl}$ ). The concentration of total phosphorus was determined using a spectrophotometer after digestion with  $\text{HNO}_3\text{-HCl}$  (3:1, V/V) at 200 °C. The solubilization of the mineral and organic phosphate forms was conducted using 1:1  $\text{H}_2\text{SO}_4$  (Bowman, 1988). The phosphorus contained in the sulfuric extract represented the total concentration of this element.

The mean values of the environmental variables obtained at each sampling location were calculated for each site.

#### 2.4.2. Landscape-scale variables

Landscape-scale variables representative of land use and isolation were obtained for each site (Fig. 2) from vectorial layers of hydrography (1:50,000 scale; Instituto Brasileiro de Geografia e Estatística) and land use (1:100,000 scale; 2010 Landsat sensor multispectral images; Instituto Estadual do Ambiente do Estado do Rio de Janeiro) using a geographic information system (ArcGIS v. 10.2; ESRI, 2013).

To better estimate the land use metrics, as 200 m corresponds to the shorter distance between two sampling locations within a same site, a 200-m radius buffer was defined for each sampling location within a site per season. For each site, the land use metrics (forest cover, FC; pasture cover, PC; and human settlements, HS) were obtained as the total area ( $\text{km}^2$ ) within all 200-m radius buffers (two buffers per site in bays and oceanic beaches, and three buffers per site in coastal lagoons; Fig. 2). Land use metrics were then calculated as percentages of the total buffer area at each site. The land use metrics were calculated as the mean values for the system, subregion, and region levels (Table 1).

The isolation measurements were also obtained for each sampling location within each site per season (Fig. 2). The distance from the ocean (DO; km) corresponds to the mean distance of each site from the ocean. Estuaries closer than 5 km (ES) corresponds to the total number of estuaries located within 5 km of distance from each site. The distance of 5 km was considered because it corresponds to approximately half of the shorter distance between two sites. The mean values of DO were calculated for the system, subregion, and region levels, whereas the sum of the values of ES was calculated for each hierarchical level (Table 1).

#### 2.5. Data analysis

A principal component analysis (PCA) was performed with the water quality variables and substrate measurements, in order to account for the multicollinearity and reduce the number of variables by summarizing the predominant patterns of original data in new latent variables. The PCA was based on a correlation matrix of the centered and standardized data (Legendre and Legendre, 2012). The most relevant axes to explain the environmental variance were selected based on the broken-stick criterion (Peres-Neto et al., 2003), and used as latent environmental variables. Considering that the squares of all loadings for the original variables in an individual principal component sum to one, the variables that most contributed for the explained variance were considered as those with loadings larger than the hypothetical equal contribution of all variables (i.e. the square root of 1 divided by 9 variables;  $r = 0.33$ ). For the system, subregion, and region levels, the latent environmental variables were calculated as the mean values of the scores. The same procedure was not applied to the landscape-scale variables because the land use metrics and isolation measurements include few and informative metrics. However, all environmental variables were tested for multicollinearity before included in the models, as described later.

Generalized linear mixed models (GLMMs) were used to investigate the effects of site-scale variables (PCA axes representing latent environmental variables) and landscape-scale variables (land use metrics, FC, PC, and HS; and isolation measurements, DO and ES) on the  $\gamma$  diversity and  $\alpha$  and  $\beta$  components at three hierarchical levels (site, system, and subregion). The environmental effects were not investigated for the region level due to the reduced number of samples. The hierarchical level and the season (dry; wet) were included as random effects to control for the possible effects of the spatial dependence and temporal variability on the residual variance, respectively (Gelman and Hill, 2007; Bolker et al., 2009). Site was included as random effect in the site-based models, whereas system was included as random effect in the system-based models, and subregion was included as random effect in the subregion-based models. Therefore, the effects of the sampling units were removed from the species-environment relationships, and the samples from different seasons obtained at same sampling units were considered as independent samples. In all models, the diversity values were included as Poisson variables, since the Poisson distribution is appropriate for count data, with a log link function that makes the expected response linear and the expected variance homogeneous (Gelman and Hill, 2007; Bolker et al., 2009).

In order to avoid multicollinearity between the predictor variables, for each hierarchical level and diversity measure, the variance inflation factor (VIF) was calculated for each predictor variable in the full model. Generalized linear models (GLMs) were used for the variable selection, with the full models including all predictor variables as fixed effects (Gelman and Hill, 2007). For the subregion level, the full models did not include all predictor variables because of the reduced degrees of freedom. In this case, the variable selection was first based on full models including only one type of variable (i.e. site-scale environmental conditions, land use metrics or distance measurements). In a second step, the variable selection was based on full models including all the remaining variables. Following the approach suggested by Zuur et al. (2010), the variable with highest VIF was removed from the full model and the procedure was repeated until no variable present VIF > 4. After removing variables with highest multicollinearity, almost all variables presented VIF < 2. These procedures resulted in different types and/or numbers of predictor variables in the final full models for the different diversity measures and/or hierarchical levels.

For each hierarchical level and diversity measure, GLMMs with all possible combinations of the remaining predictor variables in the full models were run. Therefore, for each hierarchical level and diversity measure the candidate models were subsets of the full models. All models, including null models with no fixed effect, included the

**Table 1**

Site-scale and landscape-scale environmental variables included in the generalized linear mixed models. Code, median, lower and upper quartiles of the samples obtained during the dry and wet seasons for each hierarchical level (site, system, subregion, and region).

Variable	Code	Site		System		Subregion		Region					
		Median	Quartiles		Median	Quartiles		Median	Quartiles				
			Lower	Upper		Lower	Upper		Lower	Upper			
Site scale													
Environmental conditions													
Principal component 1	E1	0.06	-1.26	0.95	0.14	-1.12	0.47	-0.16	-0.60	0.38	-0.15	-0.32	0.01
Principal component 2	E2	-0.23	-1.19	1.01	-0.11	-0.78	0.82	-0.46	-0.65	0.24	-0.10	-0.25	0.04
Landscape scale													
Land use metrics													
Forest cover (%)	FC	7	0	50	9	2	26	13	12	39	23	22	24
Pasture cover (%)	PC	11	0	22	11	7	14	11	10	18	14	13	14
Human settlement (%)	HS	33	0	47	25	18	60	36	28	43	35	35	36
Distance measurements													
Distance from the ocean (km)	DO	15.6	0.5	33.7	13.4	1.4	30.1	16.3	4.1	30.7	17.3	17.3	17.3
Estuaries closer than 5 km	ES	1	1	5	2	0	11	6	5	28	39	39	39

hierarchical level as a random effect. Other models with the same fixed effects and random effects for the hierarchical levels also included the season as a random effect, in order to evaluate its influence on the models fitting. The pseudo- $R^2$  for GLMMs was calculated to express the variance explained by the fixed effects (marginal  $R^2$ ,  $\text{marg. } R^2$ ) and the entire model, including both fixed and random effects (conditional  $R^2$ ,  $\text{cond. } R^2$ ) (Nakagawa and Schielzeth, 2013; Johnson, 2014). A model selection based on the information theory was applied to compare the fitted models and identify the best-supported models (Burnham and Anderson, 2002). We did not exclude the candidate models with negligible values of pseudo- $R^2$  because the model selection was focused on evaluating the performance of each model compared with the data set included in the full model.

The model selection was based on the corrected Akaike information criterion (AICc), which correct for the bias resulting from small sample sizes (Burnham and Anderson, 2002). The best model is the one with lower AICc (i.e. with less information loss and simpler structure) because adding more variables to improve the goodness of fit of a model (i.e. its likelihood) is penalized by the increasing number of parameters, which discourages overfitting (Burnham and Anderson, 2002). Models were ranked according the AICc weight ( $w_i$ ) that represents the probability that the model is the best between the set of candidate models, i.e. the relative likelihood of the model (Wagenmakers and Farrel, 2004). The  $\Delta\text{AICc}$  is the difference between the lowest AICc and the AICc of the model, and represents the probability that the model minimizes the information loss (Burnham and Anderson, 2002; Wagenmakers and Farrel, 2004). All models with  $\Delta\text{AICc} < 2$  were considered with substantial support for interpretation (Burnham and Anderson, 2002).

A model averaging approach was applied when more than one model was selected to explain a given diversity measure at given hierarchical level. This procedure combined inferences across the selected models by calculating model-averaged parameter estimates and the associated confidence intervals (Burnham and Anderson, 2002). The model averaging approach estimated the strength of the environmental effects based on their contributions for average model. We used 85% confidence intervals because the model selection using the AICc supports additional variables over a null model at this level (Arnold, 2010). Therefore, a parameter was considered to be informative if the 85% confidence interval did not overlap 0. The relative variable importance (RVI) for the parameter estimates in the average model was also calculated by summing the  $w_i$  of the selected models (i.e. calculated again without the other candidate models) including the predictor variable (Burnham and Anderson, 2002).

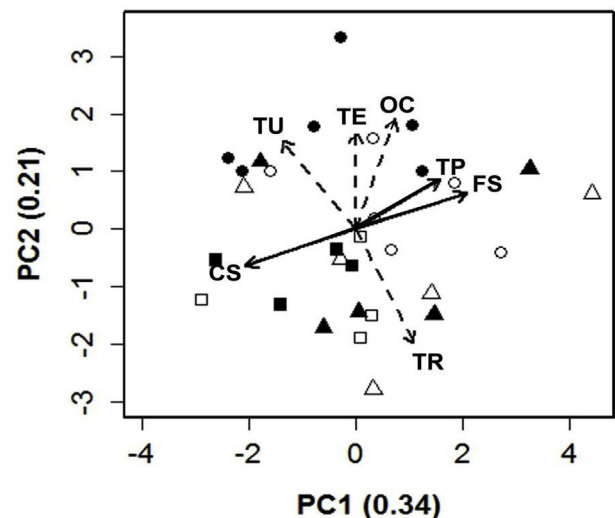
For all models, the predictor variables were centered and standardized (Legendre and Legendre, 2012). All analyses were performed in

the R environment (version 3.5.2; R Core Team, 2018), with the packages *vegan* (version 2.5–3; Oksanen et al., 2018), *car* (version 3.0–2; Fox and Weisberg, 2011), *lme4* (version 1.1–19; Bates et al., 2015), *AICcmodavg* (version 2.1–1; Mazerolle, 2017), and *MuMIn* (version 1.42.1, Barton, 2018).

### 3. Results

#### 3.1. Site- and landscape-scale environmental effects

In the PCA performed with the water quality variables and substrate measurements, the first and second axes explained 55% of the environmental variance (Fig. 4), and were selected to be used as latent environmental variables (E1 and E2). The latent environmental variable E1 expressed primarily the gradient of sediment measurements, with negative scores associated with coarser substrates and positive scores



**Fig. 4.** Principal component analysis ordination based on site-scale environmental measures of salinity (SA), transparency (TR), total nitrogen (TN), total phosphorous (TP), temperature (TE), turbidity (TU), organic carbon (OC), coarse substrate (CS), and fine substrate (FS) obtained in coastal lagoons (circles), bays (triangles), and oceanic beaches (squares) during the dry season (empty symbols) and wet season (filled symbols). Only variables that most explained the variance in each axis ( $r > 0.33$ ) are shown and the arrow lengths correspond to the strength of the correlation of the variables with the axes. Solid and dashed lines indicate the variables most relevant in the first and second axes, respectively.

with finer substrates and higher concentrations of total phosphorous (Fig. 4). The latent environmental variable E2 expressed primarily the gradient of water quality, with positive scores associated with higher turbidity, temperature, and concentrations of organic carbon, and negative scores associated with higher transparency (Fig. 4).

The latent environmental variables E1 and E2 were calculated as the mean values of the scores for the system, subregion, and region levels (Table 1). The coastal lagoons were primarily associated with the positive scores of E1 and E2, whereas oceanic beaches were primarily associated with the negative scores of E1 and E2 (Fig. 4). The bays, in turn, were primarily associated with intermediate to positive values of E1, and intermediate to negative scores of E2 (Fig. 4). For all cases, the seasonal effects were not expressive.

The variable selection applied to the site- and landscape-scale variables resulted in final full models with six predictor variables at the site level, four to five at the system level, and three at the subregion level. At the site level, E1, E2, PC, HS, DO, and ES were included in the full model for the  $\alpha$  and  $\beta^M$  components, and the measures of  $\gamma$  diversity considered in both methods of partitioning (i.e. additive and multiplicative), whereas E1, E2, FC, PC, DO, and ES were included in the full model for the  $\beta^A$  component. At the system level, E1, E2, FC, PC, and DO were included in the full model for the  $\alpha$  component, E1, E2, PC, and DO were included in the full model for the  $\beta^A$  component, and E1, E2, HS, and DO were included in the full model for the  $\beta^M$  component and both measures of  $\gamma$  diversity. At the subregion level, E1, E2, and FC were included in the full models for all diversity measures. For each diversity measure and hierarchical level, the smaller models were subsets of the aforementioned full models.

### 3.2. Hierarchical patterns of fish diversity

For all diversity measures, considering the samples within the same hierarchical level, in general, a greater variation in the values was observed during the wet season, but no differences were observed between samples of the dry and wet seasons (Fig. 5). Total diversity ( $\gamma$ ) increased across the hierarchical levels, especially from the subregion to the region level (Fig. 5a–c). The  $\alpha$  and  $\beta^A$  components also increased across the hierarchical levels, with more expressive increases from the subregion to the region level (Fig. 5d–e). According to the values of the  $\beta^A$  component, more species were absent from assemblages at the higher hierarchical levels compared with the lower hierarchical levels.

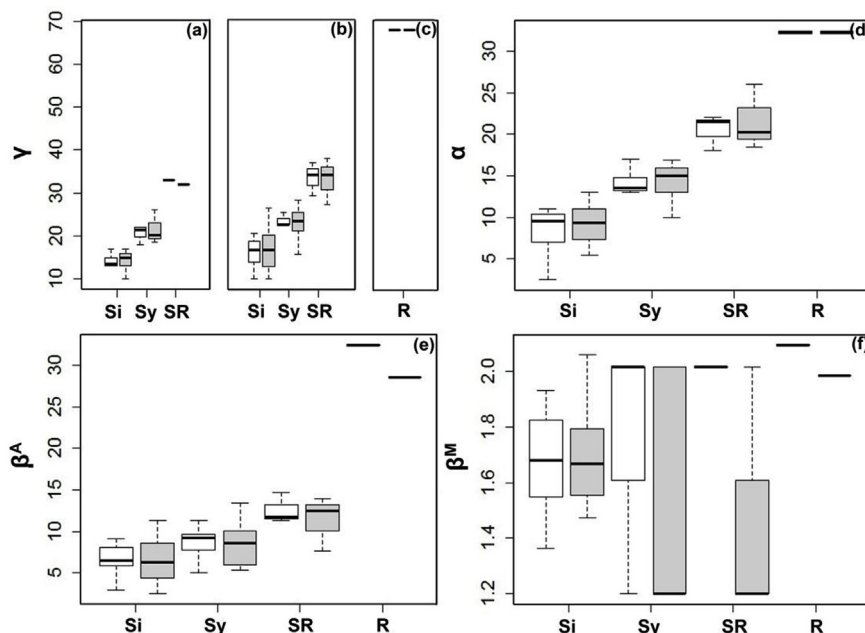


Fig. 5. Box-plots (median, lower and upper quartiles, and minimum and maximum values) of the variation in the values of the total diversity ( $\gamma$ ) calculated as (a) the mean species richness and (b, c) the species richness, and the components (c) alpha ( $\alpha$ ), (d) additive beta ( $\beta^A$ ), and (e) multiplicative beta ( $\beta^M$ ) at the site (Si), system (Sy), subregion (SR), and region (R) levels. Different colors represent samples from the dry (white) and wet (grey) seasons.

The  $\beta^M$  component did not present expressive changes from the site to the subregion level (Fig. 5f). A slight increase was observed from the subregion to the region level, indicating that more assemblages do not share species at the subregions within the region level compared with the other hierarchical levels (Fig. 5f).

The  $\gamma$  diversity representing the mean richness varied more within the system level, whereas the  $\gamma$  diversity values representing the total species richness varied more within the site and subregion levels (Fig. 5a–c). A slightly more expressive variation was also observed for the values of  $\alpha_{\text{site}}$  and  $\alpha_{\text{sub}}$  compared with  $\alpha_{\text{sys}}$  and especially  $\alpha_{\text{reg}}$ , indicating higher differences between the mean richness per sampling locations within sites and per system within subregions (Fig. 5d). The same trend was observed for values of the  $\beta^A$  component, with the mean number of absent species in assemblages varying more between sampling locations within sites and systems within subregions compared with the other levels (Fig. 5e). The higher variation in the values of  $\beta^M_{\text{sys}}$  in turn, indicated higher differences in the number of assemblages that do not share species between sites (Fig. 5f). The  $\beta^M_{\text{site}}$  and  $\beta^M_{\text{sub}}$  (during the wet season), also presented a more expressive variation compared with  $\beta^M_{\text{reg}}$ , indicating lower differences in the number of assemblages that do not share species between subregions within the study area (Fig. 5f).

### 3.3. Spatial structure and seasonal variation of environmental effects

According to the model selection, in general, more than one model was plausible to explain the diversity measures at the site and system levels ( $\Delta\text{AICc} < 2$ ; Table 2). The season was included in selected models for the subregion level, but its random effect on the residual variance, as well as the random effect of the hierarchical level, was negligible (marg.  $R^2 = \text{cond. } R^2$ ; Table 2). The random effect of the hierarchical level, in turn, was more relevant in the selected models for the site level, primarily on the residual variance associated with the values of  $\gamma_{\text{site}}$  represented by the total species richness within sites (Table 2). The random effect of the system level was more expressive on the residual variance of  $\beta^A_{\text{sys}}$  (Table 2).

Only the site effect was included in the best supported model to explain the values of  $\alpha_{\text{site}}$  ( $\text{AICc } wi = 0.18$ ;  $\text{cond. } R^2 = 0.16$ ; Table 2), and in the other selected models the site effect was proportional to the environmental effects (Table 2). The positive effect of E1, with higher values of  $\alpha_{\text{site}}$  in finer substrates and higher concentrations of total

**Table 2**

Poisson regression models (generalized linear mixed models) for the relationships between the total diversity (gamma –  $\gamma$ ) considered in the additive and multiplicative methods of partitioning and its components (alpha –  $\alpha$ ; additive beta –  $\beta^A$ ; and multiplicative beta –  $\beta^M$ ) and the site and landscape environmental variables in each hierarchical level (site – Si; system – Sy; and subregion – SR). The different hierarchical levels and the season were random effects. Results are shown only for models with the  $\Delta AICc < 2$ . The best-supported model given the data set and the candidate models is in bold. Codes for environmental variables included as fixed effects in Table 1.

Dependent variable	Model	Effects		marg. R <sup>2</sup>	cond. R <sup>2</sup>	K	LL	AICc	$\Delta AICc$	AICc wi	
		Fixed	Random								
Site $\alpha$	<b>m1</b>	-	Si	-	<b>0.16</b>	<b>2</b>	<b>-68.46</b>	<b>141.39</b>	<b>0.00</b>	<b>0.18</b>	
	m2	E1	Si	0.07	0.15	3	-67.56	142.13	0.74	0.12	
	m3	PC	Si	0.07	0.15	3	-67.72	142.44	1.05	0.11	
Additive $\beta^A$	<b>m1</b>	<b>FC, PC</b>	Si	<b>0.29</b>	<b>0.37</b>	<b>4</b>	<b>-59.83</b>	<b>129.40</b>	<b>0.00</b>	<b>0.11</b>	
	m2	E2	Si	0.16	0.32	3	-61.37	129.73	0.33	0.09	
	m3	E2, FC, PC	Si	0.40	0.40		-58.66	130.05	0.65	0.08	
	m4	PC	Si	0.14	0.38	3	-61.66	130.31	0.91	0.07	
	m5	-	Si	-	0.36	2	-63.09	130.67	1.27	0.06	
$\Gamma$	<b>m1</b>	<b>PC</b>	Si	<b>0.19</b>	<b>0.19</b>	<b>3</b>	<b>-66.39</b>	<b>139.77</b>	<b>0.00</b>	<b>0.25</b>	
	m2	E2, PC	Si	0.24	0.24	4	-65.54	140.81	1.04	0.14	
Multiplicative $\beta^M$	<b>m1</b>	-	Si	-	<b>0.01</b>	<b>2</b>	<b>-28.68</b>	<b>61.83</b>	<b>0.00</b>	<b>0.25</b>	
	$\Gamma$	-	Si	-	<b>0.44</b>	<b>2</b>	<b>-80.43</b>	<b>165.34</b>	<b>0.00</b>	<b>0.20</b>	
	m2	E1	Si	0.08	0.40	3	-79.58	166.16	0.82	0.13	
System $\alpha$	<b>m1</b>	<b>PC</b>	Sy	<b>0.28</b>	<b>0.28</b>	<b>3</b>	<b>-32.53</b>	<b>73.46</b>	<b>0.00</b>	<b>0.30</b>	
	Additive $\beta^A$	<b>m1</b>	-	Sy	-	<b>0.37</b>	<b>-33.55</b>	<b>72.19</b>	<b>0.00</b>	<b>0.28</b>	
		m2	E2	Sy	0.24	0.38	3	-32.13	72.67	0.48	0.22
$\Gamma$	<b>m1</b>	-	Sy	-	<b>0.01</b>	<b>2</b>	<b>-36.28</b>	<b>77.65</b>	<b>0.00</b>	<b>0.36</b>	
	m2	DO	Sy	0.15	0.15	3	-35.13	78.66	1.01	0.22	
Multiplicative $\beta^M$	<b>m1</b>	-	Sy	-	0.01	<b>2</b>	<b>-17.60</b>	<b>40.29</b>	<b>0.00</b>	<b>0.42</b>	
	$\Gamma$	-	Sy	-	0.06	<b>2</b>	<b>-39.00</b>	<b>83.10</b>	<b>0.00</b>	<b>0.28</b>	
	m2	DO	Sy	0.20	0.20	3	-37.48	83.36	0.26	0.25	
Subregion $\alpha$	<b>m1</b>	<b>E1, E2, FC</b>	<b>SR, S</b>	<b>0.24</b>	<b>0.24</b>	<b>6</b>	<b>-14.90</b>	<b>-42.19</b>	<b>0.00</b>	<b>1.00</b>	
	Additive $\beta^A$	<b>m1</b>	<b>E1, E2, FC</b>	<b>SR, S</b>	<b>0.39</b>	<b>0.39</b>	<b>6</b>	<b>-13.05</b>	<b>-45.89</b>	<b>0.00</b>	<b>1.00</b>
		$\gamma$	<b>E1, E2, FC</b>	<b>SR, S</b>	<b>0.01</b>	<b>0.01</b>	<b>6</b>	<b>-16.99</b>	<b>-40.04</b>	<b>0.00</b>	<b>1.00</b>
Multiplicative $\beta^M$	<b>m1</b>	<b>E1, E2, FC</b>	<b>SR, S</b>	<b>0.07</b>	<b>0.07</b>	<b>6</b>	<b>-7.49</b>	<b>-57.02</b>	<b>0.00</b>	<b>1.00</b>	
	$\gamma$	<b>E1, E2, FC</b>	<b>SR, S</b>	<b>0.37</b>	<b>0.37</b>	<b>6</b>	<b>-16.01</b>	<b>-39.99</b>	<b>0.00</b>	<b>1.00</b>	

Marg. R<sup>2</sup>, marginal coefficient of determination representing the variance explained by fixed effects; cond. R<sup>2</sup>, conditional coefficient of determination representing the variance explained by the entire model, including both fixed and random effects; K, number of factors (including intercept); LL, log likelihood; AICc, corrected Akaike Information Criterion;  $\Delta AICc$ , the difference between the value of each model and the model with lower AICc; AICc wi, AICc weight.

phosphorus, was slight and the parameter estimates overlapped zero (Tables 2 and 3). The negative effect of PC was more expressive, but the parameter estimates also overlapped zero (Tables 2 and 3). At the system level, only one model was selected to explain the values of the  $\alpha$  component, with the negative effect of PC supported by the parameter estimates (Table 3), and a negligible effect of the system (AICc wi = 0.30; marg. R<sup>2</sup> = 0.28; cond. R<sup>2</sup> = 0.28; Table 2). In the best supported model for the subregion level, the  $\alpha$  component was positively associated with E1, E2, and FC, with a negligible effect of the hierarchical level (AICc wi = 1; marg. R<sup>2</sup> = 0.24; cond. R<sup>2</sup> = 0.24; Table 2). Higher values of  $\alpha_{sub}$  occurred in higher values of FC, finer substrates, and higher concentrations of total phosphorus (Table 3). The effect of E2, favoring  $\alpha_{sub}$  in higher concentrations of organic carbon, turbidity, and temperature, and lower transparency, was slight and the parameter estimates overlapped zero (Tables 2 and 3).

In the set of plausible models selected to explain  $\beta^A_{site}$ , the site effect was more expressive in models with less parameters (marg. R<sup>2</sup> < cond. R<sup>2</sup>; Table 2), indicating that the random effects were most likely related to the variables not included in the models. The  $\beta^A_{site}$  was negatively associated with PC and FC, and positively E2 (Tables 2 and 3), all environmental effects were relevant in the average model according to the parameter estimates (Table 3). Therefore,  $\beta^A_{site}$  decreased in higher values of PC and FC, whereas more species were

absent within sites with higher turbidity, temperature, and concentration of organic carbon, and lower transparency (Tables 2 and 3). The PC followed by FC were more relevant to explain the observed patterns in the average model (RVI = 63 and 46, respectively; Table 3).

The best supported model to explain the  $\beta^A$  component at the system level included only the random effect of the hierarchical level (AICc wi = 0.28; cond. R<sup>2</sup> = 0.37; Table 2), which was also expressive in the other selected model (marg. R<sup>2</sup> < cond. R<sup>2</sup>; Table 2). The negative effect of E2 was supported by the parameter estimates, indicating that less species were absent from assemblages in higher turbidity, temperature, and concentration of organic carbon, and lower transparency (marg. R<sup>2</sup> = 0.24; Tables 2 and 3). At the subregion level, the only model selected to explain the values of  $\beta^A_{sub}$  included negative effects of E1, E2, and FC, all supported by the parameter estimates (AICc wi = 1; marg. R<sup>2</sup> = 0.39; cond. R<sup>2</sup> = 0.39; Tables 2 and 3). According to that model, more species were absent from assemblages in lower values of FC, coarser substrates, lower concentrations of total phosphorus and organic carbon, lower turbidity and temperature, and higher transparency (Table 3).

At the site and system levels, the values of the  $\beta^M$  component were not explained by the fixed effects and the random effects of the hierarchical levels were negligible (cond. R<sup>2</sup> = 0.01; Table 2). According to the best supported model for the subregion level,  $\beta^M_{sub}$  was positively



**Table 3**  
Parameter estimates (with confidence intervals using a nominal level of 85%) and relative variable importance (RVI) based on the average of the generalized linear mixed models (Poisson regressions) selected (ΔAICc < 2) for the relationships between the total diversity (gamma - γ) considered in the additive and multiplicative methods of partitioning and its components (alpha - α; additive beta - β<sup>A</sup>; and multiplicative beta - β<sup>M</sup>) and the site and landscape environmental variables in each hierarchical level (site; system; and subregion). Codes for environmental variables included as fixed effects in Table 1. The number of models (n) including each parameter was indicated when more than one model was selected. Estimates with confidence intervals that do not overlap zero in bold.

Parameter	α			β <sup>A</sup>			γ (additive)			β <sup>M</sup>			γ (multiplicative)		
	Estimate	RVI		Estimate	RVI		Estimate	RVI		Estimate	RVI		Estimate	RVI	
Site															
Intercept	2.158 (2.050, 2.267)	-		1.455 (1.289, 1.622)	-		2.599 (2.522, 2.676)	-		0.035 (0.000, 0.284)	-		2.641 (2.526, 2.756)	-	
E1	0.061 (-0.004, 0.127)	0.30 (n = 1)		-	-		-	-		-	-		0.059 (-0.005, 0.122)	0.40 (n = 1)	
E2	-	-		<b>0.141 (0.026, 0.256)</b>	0.42 (n = 2)		0.052 (-0.007, 0.111)	0.37 (n = 1)		-	-		-	-	
FC	-	-		<b>-0.418 (-0.758, -0.077)</b>	0.46 (n = 2)		-	-		-	-		-	-	
PC	<b>-0.185 (-0.407, 0.036)</b>	0.26 (n = 1)		<b>-0.449 (-0.788, -0.111)</b>	0.63 (n = 3)		<b>-0.252 (-0.422, -0.082)</b>	1.00 (n = 2)		-	-		-	-	
System															
Intercept	2.601 (2.487, 2.714)	-		1.895 (1.683, 2.107)	-		3.036 (2.946, 3.127)	-		0.452 (0.128, 0.744)	-		3.045 (2.952, 3.138)	-	
E2	-	-		<b>-0.200 (-0.378, -0.023)</b>	0.44 (n = 1)		-	-		-	-		-	-	
PC	<b>-0.331 (-0.591, -0.070)</b>	0.61 (n = 1)		-	-		-	-		-	-		-	-	
DO	-	-		-	-		0.184 (-0.004, 0.371)	0.38 (n = 1)		-	-		<b>0.212 (0.025, 0.399)</b>	0.47 (n = 1)	
Subregion															
Intercept	3.045 (3.004, 3.085)	-		2.325 (2.277, 2.373)	-		3.483 (3.375, 3.588)	-		0.539 (0.025, 0.967)	-		3.481 (3.439, 3.522)	-	
E1	<b>0.077 (0.001, 0.152)</b>	-		<b>-0.084 (-0.158, 0.001)</b>	-		0.023 (-0.175, 0.216)	-		0.253 (-0.552, 1.044)	-		<b>0.079 (0.002, 0.155)</b>	-	
E2	0.025 (-0.020, 0.069)	-		<b>-0.049 (-0.097, -0.002)</b>	-		<b>-0.008 (-0.124, 0.103)</b>	-		<b>0.057 (-0.477, 0.528)</b>	-		<b>0.048 (0.003, 0.093)</b>	-	
FC	<b>0.160 (0.054, 0.266)</b>	-		<b>-0.393 (-0.784, -0.003)</b>	-		<b>-0.024 (-0.302, 0.257)</b>	-		<b>-0.372 (-1.624, 0.861)</b>	-		<b>0.171 (0.065, 0.279)</b>	-	

associated with E1 and E2, and negatively associated with FC, indicating that more assemblages most likely do not share species in finer substrates, higher concentrations of total phosphorous and organic carbon, higher turbidity and temperature, and lower transparency and values of FC (AICc  $w_i = 1$ ; marg.  $R^2 = 0.07$ ; cond.  $R^2 = 0.07$ ; Table 2). However, these effects were not supported by the parameter estimates, which overlapped zero (Tables 2 and 3).

The random effect of the site was negligible in the models selected to explain the  $\gamma_{site}$  representing the mean species richness within systems (marg.  $R^2 =$  cond.  $R^2$ ; Table 2). The negative effect of PC was of primary relevance to explain  $\gamma_{site}$  in the average model (RVI = 1.00; Table 3), whereas the positive effect of E2 was not supported by the parameter estimates (Table 3). The effect of the site was expressive in the models selected to explain the  $\gamma_{site}$  representing the total richness (marg.  $R^2 <$  cond.  $R^2$ ; Table 2), with the best supported model including only that random effect (AICc  $w_i = 0.20$ ; cond.  $R^2 = 0.44$ ; Table 2). The positive effect of E1 was slight and not supported by precise parameter estimates, which overlapped zero (Table 3).

At the system level, the values of  $\gamma_{sys}$  representing both the mean species richness within subregions and total species richness were explained by two models (Table 2). In both cases, the best supported models included only the system effect, which influenced only slightly the  $\gamma_{sys}$  representing the total species richness (AICc  $w_i = 0.28$ ; cond.  $R^2 = 0.06$ ) and was negligible for the  $\gamma_{sys}$  representing the mean species richness (AICc  $w_i = 0.36$ ; cond.  $R^2 = 0.01$ ; Table 2). The positive effect of DO was relevant for both measures of  $\gamma_{sys}$  (RVI > 0.30; Table 3), but that effect was supported by the parameter estimates only for the  $\gamma_{sys}$  representing the total species richness (Table 3).

The measures of  $\gamma_{sub}$  were also associated with E1, E2, and FC, with a negligible effect of the subregion (marg.  $R^2 =$  cond.  $R^2$ ; Table 2). The  $\gamma_{sub}$  representing the mean species richness within the region was positively associated with E1, and negatively associated with E2 and FC, but the fixed effects were negligible and not supported by the parameter estimates (marg.  $R^2 = 0.01$ ; Tables 2 and 3). The  $\gamma_{sub}$  representing the total species richness was positively associated with E1, E2, and FC, with all effects supported by precise parameter estimates (marg.  $R^2 = 0.37$ ; Tables 2 and 3). Therefore, higher values of  $\gamma_{sub}$  representing the total species richness occurred in higher values of FC, finer substrates, and higher concentrations of total phosphorus and organic carbon, higher turbidity and temperature, and lower transparency (Table 3).

#### 4. Discussion

The general increase in the diversity measures across the hierarchical levels evidenced the spatial structure of diversity patterns in the coastal ecosystems. These patterns expressed differences in the species richness, primarily associated with the  $\alpha$  component and the number of species absent from assemblages, represented by the  $\beta^A$  component. The  $\alpha$  component most likely reflected the increasing diversity of habitats available from the lower to the higher hierarchical level, with more niche opportunities supporting richer assemblages (Messmer et al., 2011; Heino et al., 2015). Based on the variation in the values of the  $\alpha$  component within the hierarchical levels, differences regarding the habitat availability are most likely more expressive within sites and between systems. As the habitat heterogeneity is expected to increase with the area, the higher variation of  $\alpha_{site}$  was most likely associated with the limited representation of the environmental heterogeneity covered by each sampling location within each site. Indeed, Franco et al. (2019) observed that the structure of the fish assemblages and environmental conditions at the coastal lagoons considered in the present study were significantly influenced by the sampling sites, herein pooled as the sampling locations within sites. Evidences for the former possibility, in turn, were provided by Azevedo et al. (2017) that observed higher taxonomic and functional distinctiveness in coastal lagoons compared with bays and oceanic beaches in

the study area.

The  $\beta$  component may express both the nestedness of assemblages, when assemblages of depauperate sites are subsets of the assemblages of sites with more species, and the species turnover, with the replacement of species between sites (Baselga, 2012). As the biogeographical regionalization ensures that the regional pool of species is almost the same for the entire study area (Henriques et al., 2017; Araújo et al., 2018), the increasing values of  $\beta^A$  across the hierarchical levels suggests that the species loss causing nestedness is a relevant process, especially at the region level. This pattern reinforces the relevance of the aforementioned ecological differences between the types of ecosystem. Concurrently, as each type of ecosystem is located at a different sub-region, the observed pattern most likely is also influenced by the geographical proximity and the dispersal limitation of species (Cowen et al., 2006; Bender et al., 2017). It is likely that both dispersal and environmental filters have a higher relevance at the region level (Vilar et al., 2013), supporting a greater species loss between subregions compared with the lower hierarchical levels.

Less expressive increases in the number of effective assemblages that do not share species ( $\beta^M$ ) across the hierarchical levels suggest that the species turnover also influences the diversity patterns. This possibility is reinforced by the coincidence between the higher values of  $\beta^M_{reg}$  and the expressive taxonomic and functional distinctiveness between the three types of systems (Azevedo et al., 2017). Concurrently, this provides evidences regarding the relevant roles of dispersal and environmental filters to determine also the species turnover between subregions. Different assemblages may be dominated by the species replacement or species loss, and may present the same species richness, but different species compositions (Baselga, 2010). Therefore, more accurate estimates of the relative contribution of each process must rely on dissimilarity measures based on the species composition (Baselga, 2012; Legendre, 2014). Despite that, assuming the same species pool for the entire study area, and that the prevalence of the species turnover is more presumable between assemblages in geographically distant areas or high-endemism areas (Leprieur et al., 2011; Bender et al., 2017), differences in the species richness may provide effective tips regarding the relative contribution of each process. Considering the higher differences between the values of  $\beta^A$  across the hierarchical levels compared with the values of  $\beta^M$ , the relevance of species loss most likely increased with the hierarchical level, whereas the species turnover presented a more proportional contribution. Likewise, the higher values of  $\beta^A_{reg}$  and  $\beta^M_{reg}$  suggest a greater contribution of both processes at the region level, with a more expressive effect of the species loss.

Site-scale environmental effects were more relevant to determine the diversity patterns at the subregion level, evidencing their prevalent roles as environmental filters to select species for the ecosystems from the regional pool. The negligible effect of the subregion in all selected models was most likely influenced by the natural pooling of the different types of ecosystem, creating a higher environmental similarity and possibility of species sharing due to lower dispersal constraints (Bender et al., 2017). The positive effects of the gradient of sediment measurements (E1) on  $\alpha_{sub}$  and the  $\gamma_{sub}$  representing the total richness, and the gradient of water quality (E2) on the later diversity measure, evidenced the relevance of more eutrophic conditions associated with the continental influence to support a higher species richness, most likely as a result of a greater variety of feeding resources (Vorwerk and Froneman, 2009; Abrantes et al., 2015). The concurrent positive effect of the forest cover (FC) evidenced its relevance to provide a more complex habitat structure and the feeding resources necessary to support richer assemblages (Azevedo et al., 2017; Whitfield, 2017).

The negative effects of E1 and E2 on the values of  $\beta^A_{sub}$  suggest a higher species loss within subregions with more oligotrophic conditions resulting from the prevalent marine influence, most likely due to the loss of freshwater and stenohaline species between systems (Franco et al., 2019). On the other hand, the lower values of  $\beta^A_{sub}$  under a higher degree of continental influence were most likely a result of the

sharing of species with different environmental affinities and tolerances within subregions (Azevedo et al., 2017; Camara et al., 2018). The negative effect of FC on  $\beta^A_{sub}$ , in turn, reinforces its relevance for the maintenance of the higher habitat heterogeneity necessary to support richer assemblages within subregions (Casatti et al., 2009; Whitfield, 2017), and with more species in common due to the shared regional pool (Araújo et al., 2018). These relationships reflected the differences between the types of ecosystems, and evidenced that the degree of continental and marine influence is a primary determinant of the diversity patterns at the subregion level (Azevedo et al., 2017; Franco et al., 2019).

A more proportional contribution of environmental effects at site and landscape scales was observed at the system level, but also primarily as a result of different degrees of marine and continental influences. The positive effect of the distance from the ocean (DO) on the  $\gamma_{sys}$  representing the total species richness per system was most likely a result of the occurrence of species with different environmental affinities and tolerances in ecosystems with more continental influence, such as the coastal lagoons (Azevedo et al., 2017; Camara et al., 2018). Concurrently, the negligible effect of the system evidenced the high environmental similarity between the same types of ecosystem. The strong effect of the system observed for  $\beta^A_{sys}$ , in turn, suggests relevant differences in the species loss within different systems. These differences were primarily associated with the negative effect of E2 because the variance explained by both fixed and random effects did not increase expressively when that fixed effect was included in the model (Table 2). Therefore, more species were absent within systems with more oligotrophic conditions, and considering the negligible effects of the system and the environmental conditions on  $\beta^M_{sys}$ , the species loss is most likely a more determinative process for the diversity patterns within ecosystems. In this sense, the water quality and DO are most likely the prevalent filters to select species for ecosystems according to their affinities to higher continental or marine influences, with richer assemblages supported by a higher continental influence. In addition, the negative effect of PC on  $\alpha_{sys}$  was most likely associated with the relevance of that variable to determine processes of environmental homogenization within ecosystems due to a lower variety of habitats and feeding resources (Casatti et al., 2009).

The relevance of the environmental homogenization for the diversity patterns was reinforced by the negative effect of PC on  $\gamma_{site}$  representing the mean species richness. Likewise, the negative effect of PC on  $\beta^A_{site}$  suggested a lower species loss within sites in increasing values of PC, most likely also as result of homogenization processes. In this sense, the negligible site effect and the irrelevance of environmental effects on  $\beta^M_{site}$  suggest that the species loss was indeed the prevalent process determining differences in the species richness within sites. The negative effect of FC on  $\beta^A_{site}$ , in turn, evidenced that the land use effects support similar environmental conditions within sites. Therefore, considering the role of the forest cover on the maintenance of the habitat heterogeneity (Casatti et al., 2009; Whitfield, 2017), its conservation is of primary relevance to support richer assemblages at the site level. The less relevant positive effect of E2 on  $\beta^A_{site}$ , compared with the landscape-scale effects, evidenced a more slight influence of more eutrophic conditions on the species loss, most likely due to the loss of stenohaline marine species within sites (Franco et al., 2019). These relationships summarized the primary relevance of the land use and, to a lesser extent, the water conditions representative of the continental influence as environmental filters to select species for different sites. The expressive effect of the site in the best supported models was most likely associated with the existence of relevant missing environmental variables as a result of the spatial heterogeneity (Teichert et al., 2018), as well as factors at larger scales that may increase the environmental variability between more distant sites (Vilar et al., 2013). The strong effect of the site on the variance of  $\gamma_{site}$  representing the total species richness provides a reliable evidence for these possibilities.

Contrary to the expected, the land use was more relevant to

determine the diversity patterns at the site level, whereas the site-scale environmental effects were more relevant at the subregion level. Regardless of that, the site-scale environmental conditions, especially the water quality, influenced the diversity patterns at all hierarchical levels, most likely based on the species affinities and tolerances to the continental and marine influences. The land use also influenced the diversity patterns at all hierarchical levels, evidencing the relevance of the opposite effects of FC, supporting a higher habitat heterogeneity, and PC, leading to processes of environmental homogenization, culminating in richer and depauperate assemblages, respectively. The irrelevance of the percentage of human settlements (HS) to determine the diversity patterns, in turn, evidenced that the conservation degree of the vegetal cover is the primary component of the land use effects. Also contrary to the expected, the degree of isolation did not determined decreases in the  $\alpha$  component and increases in the  $\beta$  component at neither hierarchical level. The relevance of DO to determine the  $\gamma_{\text{sys}}$  representative of the total species richness suggested that its effect was most likely associated with the prevalence of continental or marine influences on the selection of species for the different systems. That possibility was reinforced by the aforementioned higher proximity between the same types of system, naturally pooled within different subregions. The irrelevance of the number of estuaries closer than 5 km (ES) for the diversity patterns, in turn, was most likely influenced by the similar number of estuaries close to the sites and systems pooled within the same subregions. The similar proportion of potentially available habitats and higher probability of species dispersal within the subregions most likely damped the effects of ES on the diversity patterns.

Despite the limited number of samples pooled within the hierarchical levels, due to the spatially-replicated sampling design, the samples are based on ecologically representative measures of the species richness and environmental variables. Likewise, the validity of the observed relationships is supported by the analytical framework, which produced reliable estimates of the environmental effects by considering the intrinsic structure of data. The GLMMs disentangled the random effects of the hierarchical levels and seasons and the fixed environmental effects (Bolker et al., 2009), whereas the information-theoretic approach minimized the possible bias resulting from the small sampling sizes and reduced the probability of model overfitting (Burnham and Anderson, 2002). Therefore, the observed relationships were valid to support reliable inferences regarding the environmental effects on the diversity patterns.

The slight environmental effects were most likely associated with the high environmental variability in oceanic beaches (i.e. higher susceptibility to the effects of tides, waves, and oceanic currents), which limits the observation of patterns in a short-term perspective. In this sense, it is possible that a greater temporal replication may reinforce the observed relationships. Concurrently, the relevance of the site and system effects for the diversity patterns may suggest the existence of relevant environmental predictors not considered in the study. Based on the environmental effects observed, it is likely that missing variables relevant within the site and system levels are also representative of the degrees of marine and continental influences and conservation of the vegetal cover. We do not fully dismiss any possible relevant influence of limitations in the sampling design on the observed relationships, but if relevant, these effects were most likely also included in the random effects for the hierarchical levels and season, not affecting the validity of the environmental effects on the diversity patterns. Considering that most of the variance in the diversity patterns was associated with neither the random nor the fixed effects included in the models, it is likely that other factors operating at larger spatial and/or temporal scales, such as historical contingencies and changes in oceanic currents due unevenly rising of oceanic temperature (Leprieur et al., 2011; Magurran et al., 2015), also influence the observed patterns.

Several recent studies support that changes in diversity patterns are primarily associated with changes in the species composition (e.g., Magurran et al., 2015; Teichert et al., 2018). According to Magurran

et al. (2015), the lack of trends in local species richness may mask expressive changes in assemblages, such as spatial homogenization. Vargas-Fonseca et al. (2016) observed that populations of species with different feeding strategies and habitat requirements presented different responses to urbanization and connectivity with estuaries. Therefore, it is possible that stronger and more clear relationships may emerge by considering the species identity, possibly reinforcing the relevance of the environmental effects observed and evidencing the relevance of other factors, such as ES and HS, for the diversity measures based on the species richness.

Regardless of the possible limitations of measures based on the species richness, we are not aware of other studies using a hierarchical partitioning of fish diversity to investigate multi-scale environmental effects on such different types of coastal ecosystems. Recent studies focused on assemblages of coastal fishes are usually based on a single type of ecosystem, primarily on coral reefs (e.g., Messmer et al., 2011; Francisco-Ramos and Arias-González, 2013; Bender et al., 2017) and estuaries (e.g., Pasquaud et al., 2015; Henriques et al., 2017; Teichert et al., 2018). Therefore, the relationships observed in this study are of primary relevance to improve the understanding of the relative effects of environmental changes on the diversity patterns of coastal fishes. A better understanding of the assembly processes associated with coastal fishes in tropical ecosystems is of major relevance to support imperative conservation and management actions, and to guide experiments able to replicate the complexity of high-diversity assemblages and investigate its relationships with ecosystem functioning, thus reducing the lack of such information for the tropics (Clarke et al., 2017).

This study evidenced the hierarchical structure of the diversity patterns and the scale dependence of the environmental effects measured at the site and landscape scales. The relevance of each set of environmental variables was dependent on the hierarchical level, as well as the diversity measure. Therefore, the observed patterns reinforced the relevance of decomposing the  $\gamma$  diversity into  $\alpha$  and  $\beta$  components and analyzing each component separately to better access the possible mechanisms associated with the diversity patterns. Likewise, the hierarchical diversity partitioning using different methods (i.e. additive and multiplicative) was essential to better understand the possible processes (i.e. species replacement or species loss) generating the diversity patterns across the coastal ecosystems. Future investigations must benefit from including measures of  $\beta$  diversity based on the dissimilarity between the species composition, as well as functional traits and phylogenetic lineages, in order to better estimate the contribution of the species replacement and gradients of species richness for the diversity patterns (Baselga, 2012; Cardoso et al., 2014; Legendre, 2014). These approaches may also reveal relevant temporal effects not observed in the present study. Regardless of that, the hierarchical approach adopted in this study provided relevant insights regarding the multi-scale influences of the conservation of the forest cover and the maintenance of freshwater inputs on the fish diversity in tropical coastal ecosystems.

## Funding

This work was partially supported by CNPq–Conselho Nacional de Desenvolvimento Científico e Tecnológico (Grants for the last author, Process: 304954/2011-0) and by FAPERJ – Fundação Carlos Chagas de Amparo à Pesquisa do Estado do Rio de Janeiro, CNE Process: E – 26/102.997/2011. FUNBIO – Fundo Brasileiro para a Biodiversidade also give financial support for the conclusion of the project (Proc. 05/2016).

## Conflicts of interest

The authors declare that they have no conflict of interest.

## Ethical approval

All applicable international, national, and/or institutional guidelines for the care and use of animals were followed, including Universidade Federal Rural do Rio de Janeiro, Brazil, Animal Care Protocol (# 11874).

## Acknowledgements

We thank several students from the Laboratory of Fish Ecology for helping in fieldwork, especially to Gustavo Henrique Soares Guedes, Rafaela de Sousa Gomes-Gonçalves, Tailan Moretti Mattos and Wagner Uehara. We thank the Instituto Estadual do Ambiente do Estado do Rio de Janeiro (INEA) for providing the vectorial layers of hydrography and land use for GIS analyses. Grants for E.M. Camara from the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) and the Fundação Carlos Chagas Filho de Amparo à Pesquisa do Estado do Rio de Janeiro (FAPERJ). This research was conducted under SISBIO Collection of Species Permit number 10707 issued by ICMBio, Brazilian Environmental Agency.

## Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.marenvres.2019.05.006>.

## References

- Vargas-Fonseca, E., Olds, A.D., Gilby, B.L., Connolly, R.M., Schoeman, D.S., Huijbers, C.M., Hyndes, G.A., Schlacher, T.A., 2016. Combined effects of urbanization and connectivity on iconic coastal fishes. *Divers. Distrib.* 22 (12), 1328–1341. <https://doi.org/10.1111/ddi.12495>.
- Abrantes, K.G., Johnston, R., Connolly, R.M., Sheaves, M., 2015. Importance of mangrove carbon for aquatic food webs in wet-dry tropical estuaries. *Estuar. Coasts* 38 (1), 383–399. <https://doi.org/10.1007/s12237-014-9817-2>.
- Alvares, C.A., Stape, J.L., Sentelhas, P.C., de Moraes, G., Leonardo, J., Sparovek, G., 2013. Köppen's climate classification map for Brazil. *Meteorol. Z.* 22 (6), 711–728. <https://doi.org/10.1127/0941-2948/2013/0507>.
- Anderson, M.J., Crist, T.O., Chase, J.M., Vellend, M., Inouye, B.D., Freestone, A.L., Sanders, N.J., Cornell, H.V., Comita, L.S., Davies, K.F., Harrison, S.P., Kraft, N.J.B., Stegen, J.C., Swenson, N.G., 2011. Navigating the multiple meanings of  $\beta$  diversity: a roadmap for the practicing ecologist. *Ecol. Lett.* 14 (1), 19–28. <https://doi.org/10.1111/j.1461-0248.2010.01552.x>.
- Annis, G.M., Pearsall, D.R., Kahl, K.J., Washburn, E.L., May, C.A., Taylor, R.F., Cole, J.B., Ewert, D.N., Game, E.T., Doran, P.J., 2017. Designing coastal conservation to deliver ecosystem and human well-being benefits. *PLoS One* 12 (2), 1–21. <https://doi.org/10.1371/journal.pone.0172458>.
- Araújo, F.G., Rodrigues, F.L., Teixeira-Neves, T.P., Vieira, J.P., Azevedo, M.C., Guedes, A.P.P., Garcia, A.M., Pessanha, A.L.M., 2018. Regional patterns in species richness and taxonomic diversity of the nearshore fish community in the Brazilian coast. *Estuar. Coast Shelf Sci.* 208, 9–22. <https://doi.org/10.1016/j.ecss.2018.04.027>.
- Arnold, T.W., 2010. Uninformative parameters and model selection using Akaike's Information Criterion. *J. Wildl. Manag.* 74 (6), 1175–1178. <https://doi.org/10.1111/j.1937-817.2010.tb01236.x>.
- Azevedo, M.C.C., de Sousa Gomes-Gonçalves, R., Mattos, T.M., Uehara, W., Guedes, G.H.S., Araújo, F.G., 2017. Taxonomic and functional distinctness of the fish assemblages in three coastal environments (bays, coastal lagoons and oceanic beaches) in Southeastern Brazil. *Mar. Environ. Res.* 129, 180–188. <https://doi.org/10.1016/j.marenvres.2017.05.007>.
- Barton, K., 2018. MuMIn: multi-model inference. R package. version 1.42.1. <https://CRAN.R-project.org/package=MuMIn>.
- Baselga, A., 2010. Partitioning the turnover and nestedness components of beta diversity. *Glob. Ecol. Biogeogr.* 19 (1), 134–143. <https://doi.org/10.1111/j.1466-8238.2009.00490.x>.
- Baselga, A., 2012. The relationship between species replacement, dissimilarity derived from nestedness, and nestedness. *Glob. Ecol. Biogeogr.* 21 (12), 1223–1232. <https://doi.org/10.1111/j.1466-8238.2011.00756.x>.
- Bates, D., Maechler, M., Bolker, B., Walker, S., 2015. Fitting linear mixed-effects models using lme4. *J. Stat. Softw.* 67 (1), 1–48. <https://doi.org/10.18637/jss.v067.i01>.
- Bender, M.G., Leprieux, F., Mouillot, D., Kulbicki, M., Parravicini, V., Pie, M.R., Barneche, D.R., Oliveira-Santos, L.G.R., Floeter, S.R., 2017. Isolation drives taxonomic and functional nestedness in tropical reef fish faunas. *Ecography* 40 (3), 425–435. <https://doi.org/10.1111/ecog.02293>.
- Bolker, B.M., Brooks, M.E., Clark, C.J., Geange, S.W., Poulsen, J.R., Stevens, M.H.H., White, J.S.S., 2009. Generalized linear mixed models: a practical guide for ecology and evolution. *Trends Ecol. Evol.* 24 (3), 127–135. <https://doi.org/10.1016/j.tree.2008.10.008>.
- Bowman, R.A., 1988. A rapid method to determine total phosphorus in soils. *Soil Sci. Soc. Am. J.* 52, 1301–1304. <https://doi.org/10.2136/sssaj1988.03615995005200050016x>.
- Burnham, K.P., Anderson, D.R., 2001. Kullback-Leibler information as a basis for strong inference in ecological studies. *Wildl. Res.* 28 (2), 111–119. <https://doi.org/10.1071/WR99107>.
- Burnham, K.P., Anderson, D.R., 2002. *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach*, vol. 496p Springer Science and Business Media, New York-387-95364-7.
- Camara, E.M., Caramaschi, E.P., Di Dario, F., Petry, A.C., 2018. Short-term changes in two tropical coastal lagoons: effects of sandbar openings on fish assemblages. *J. Coast. Res.* 34 (1), 90–105. <https://doi.org/10.2112/JCOASTRES-D-16-00026.1>.
- Camargo, M.G., 2006. SYSGRAN: a system of open code for granulometric analyses. *Rev. Bras. Geociencias* 36, 371–378.
- Cardinale, B.J., Duffy, J.E., Gonzalez, A., Hooper, D.U., Perrings, C., Venail, P., Narwani, A., Mace, G.M., Tilman, D., Wardle, D.A., Kinzing, A.P., Daily, G.C., Loreau, M., Grace, J.B., Larigauderie, A., Srivastava, D.S., Naeem, S., 2012. Biodiversity loss and its impact on humanity. *Nature* 486 (7401), 59–67. <https://doi.org/10.1038/nature11148>.
- Cardoso, P., Rigal, F., Carvalho, J.C., Fortelius, M., Borges, P.A., Podani, J., Schmera, D., 2014. Partitioning taxon, phylogenetic and functional beta diversity into replacement and richness difference components. *J. Biogeogr.* 41 (4), 749–761. <https://doi.org/10.1111/jbi.12239>.
- Casatti, L., de Paula Ferreira, C., Carvalho, F.R., 2009. Grass-dominated stream sites exhibit low fish species diversity and dominance by guppies: an assessment of two tropical pasture river basins. *Hydrobiologia* 632 (1), 273–283. <https://doi.org/10.1007/s10750-009-9849-y>.
- Clarke, D.A., York, P.H., Rasheed, M.A., Northfield, T.D., 2017. Does biodiversity–ecosystem function literature neglect tropical ecosystems? *Trends Ecol. Evol.* 32 (5), 320–323. <https://doi.org/10.1016/j.tree.2017.02.012>.
- Cowen, R.K., Paris, C.B., Srinivasan, A., 2006. Scaling of connectivity in marine populations. *Science* 311 (5760), 522–527. <https://doi.org/10.1126/science.1122039>.
- Crist, T.O., Veech, J.A., Gering, J.C., Summerville, K.S., 2003. Partitioning species diversity across landscapes and regions: a hierarchical analysis of  $\alpha$ ,  $\beta$ , and  $\gamma$  diversity. *Am. Nat.* 162 (6), 734–743. <https://doi.org/10.1086/378901>.
- Di Dario, F., Petry, A.C., Pereira, M.M.D.S., Mincarone, M.M., Agostinho, L.S., Camara, E.M., Caramaschi, E.P., Britto, M.R.D., 2013. An update on the fish composition (teleostei) of the coastal lagoons of the restinga de Jurubatiba national park and the Imboassica lagoon, northern Rio de Janeiro state. *Acta Limnol. Bras.* 25 (3), 257–278. <https://doi.org/10.1590/S2179-975X2013000300006>.
- ESRI (Environmental Systems Research Institute), 2013. ArcGIS for Desktop (ArcInfo). (Redlands, CA) Version 10.2.
- Folk, R.L., Ward, W.C., 1957. Brazos river bar: a study of significance of grain size parameters. *J. Sediment. Petrol.* 27, 3–26.
- Fortin, M.J., James, P.M.A., Mackenzie, A., Melles, S.J., Rayfield, B., 2012. Spatial statistics, spatial regression and graph theory in ecology. *Spat. Stat.* 1, 100–109. <https://doi.org/10.1016/j.spa.2012.02.004>.
- Fox, J., Weisberg, S., 2011. *An R Companion to Applied Regression*. Sage, Thousand Oaks, California. <http://socserv.socsci.mcmaster.ca/jfox/Books/Companion>.
- Francisco-Ramos, V., Arias-González, J.E., 2013. Additive partitioning of coral reef fish diversity across hierarchical spatial scales throughout the Caribbean. *PLoS One* 8 (10), 1–14. <https://doi.org/10.1371/journal.pone.0078761>.
- Franco, T.P., Neves, L.M., Araújo, F.G., 2019. Better with more or less salt? The association of fish assemblages in coastal lagoons with different salinity ranges. *Hydrobiologia* 828 (1), 83–100. <https://doi.org/10.1007/s10750-018-3804-8>.
- Gelman, A., Hill, J., 2007. *Data Analysis Using Regression and Multilevel/Hierarchical Models*. Cambridge University Press, pp. 625p eBook ISBN-13 978-0-511-26878-6.
- Gotelli, N.J., Colwell, R.K., 2001. Quantifying biodiversity: procedures and pitfalls in the measurement and comparison of species richness. *Ecol. Lett.* 4 (4), 379–391. <https://doi.org/10.1046/j.1461-0248.2001.00230.x>.
- Halpern, B.S., Frazier, M., Potapenko, J., Casey, K.S., Koenig, K., Longo, C., Lowndes, J.S., Rockwood, R.C., Selig, E.R., Selkoe, K.A., Walbridge, S., 2015. Spatial and temporal changes in cumulative human impacts on the world's ocean. *Nat. Commun.* 6, 1–7. <https://doi.org/10.1038/ncomms8615>.
- Heino, J., Melo, A.S., Siqueira, T., Soinen, J., Valanko, S., Bini, L.M., 2015. Metacommunity organisation, spatial extent and dispersal in aquatic systems: patterns, processes and prospects. *Freshw. Biol.* 60 (5), 845–869. <https://doi.org/10.1111/fwb.12533>.
- Henriques, S., Cardoso, P., Cardoso, I., Laborde, M., Cabral, H.N., Vasconcelos, R.P., 2017. Processes underpinning fish species composition patterns in estuarine ecosystems worldwide. *J. Biogeogr.* 44 (3), 627–639. <https://doi.org/10.1111/jbi.12824>.
- Huang, D., Goldberg, E.E., Chou, L.M., Roy, K., 2018. The origin and evolution of coral species richness in a marine biodiversity hotspot. *Evolution* 72 (2), 288–302. <https://doi.org/10.1111/evo.13402>.
- Hurlbert, S.H., 1971. The nonconcept of species diversity: a critique and alternative parameters. *Ecology* 52, 577–586. <https://doi.org/10.2307/1934145>.
- Johnson, P.C., 2014. Extension of Nakagawa & Schielzeth's  $R^2_{GLMM}$  to random slopes models. *Methods Ecol. Evol.* 5 (9), 944–946. <https://doi.org/10.1111/2041-210X.12225>.
- Jost, L., 2007. Partitioning diversity into independent alpha and beta components. *Ecology* 88 (10), 2427–2439. <https://doi.org/10.1890/06-1736.1>.
- Legendre, P., 2014. Interpreting the replacement and richness difference components of beta diversity. *Glob. Ecol. Biogeogr.* 23 (11), 1324–1334. <https://doi.org/10.1111/geb.12207>.
- Legendre, P., Legendre, L., 2012. *Numerical Ecology*. Elsevier, Amsterdam9780444538697 1006p. eBook.

- Leprieux, F., Tedesco, P.A., Hugué, B., Beauchard, O., Dürr, H.H., Brosse, S., Oberdorff, T., 2011. Partitioning global patterns of freshwater fish beta diversity reveals contrasting signatures of past climate changes. *Ecol. Lett.* 14 (4), 325–334. <https://doi.org/10.1111/j.1461-0248.2011.01589.x>.
- Liquete, C., Piroddi, C., Drakou, E.G., Gurney, L., Katsanevakis, S., Charef, A., Ego, B., 2013. Current status and future prospects for the assessment of marine and coastal ecosystem services: a systematic review. *PLoS One* 8 (7), 1–15. <https://doi.org/10.1371/journal.pone.0067737>.
- Magurran, A.E., Dornelas, M., Moyes, F., Gotelli, N.J., McGill, B., 2015. Rapid biotic homogenization of marine fish assemblages. *Nat. Commun.* 6, 8405. <https://doi.org/10.1038/ncomms9405>.
- Mazerolle, M.J., 2017. AICcmodavg: model selection and multimodel inference based on (Q)AIC(c). R package version 2.1-1. <https://cran.r-project.org/package=AICcmodavg>.
- McCauley, D.J., Pinsky, M.L., Palumbi, S.R., Estes, J.A., Joyce, F.H., Warner, R.R., 2015. Marine defaunation: Animal loss in the global ocean. *Science* 347 (6219), 1255641. <https://doi.org/10.1126/science.1255641>.
- Messmer, V., Jones, G.P., Munday, P.L., Holbrook, S.J., Schmitt, R.J., Brooks, A.J., 2011. Habitat biodiversity as a determinant of fish community structure on coral reefs. *Ecology* 92 (12), 2285–2298. <https://doi.org/10.1890/11-0037.1>.
- Nakagawa, S., Schielzeth, H., 2013. A general and simple method for obtaining  $R^2$  from generalized linear mixed-effects models. *Methods Ecol. Evol.* 4 (2), 133–142. <https://doi.org/10.1111/j.2041-210x.2012.00261.x>.
- Oksanen, J., Blanchet, F.G., Friendly, M., Kindt, R., Legendre, P., McGlenn, D., Minchin, P.R., O'Hara, R.B., Simpson, G.L., Solymos, P., Stevens, M.H.H., Szoecs, E., Wagner, H., 2018. Vegan: community ecology package. R package version 2.5-1. <https://CRAN.R-project.org/package=vegan>.
- Pasqualetto, S., Vasconcelos, R.P., França, S., Henriques, S., Costa, M.J., Cabral, H., 2015. Worldwide patterns of fish biodiversity in estuaries: effect of global vs. local factors. *Estuar. Coast Shelf Sci.* 154, 122–128. <https://doi.org/10.1016/j.ecss.2014.12.050>.
- Peres-Neto, P.R., Jackson, D.A., Somers, K.M., 2003. Giving meaningful interpretation to ordination axes: assessing loading significance in principal component analysis. *Ecology* 84 (9), 2347–2363. <https://doi.org/10.1890/00-0634>.
- Pérez-Ruzafa, Á., Marcos, C., Pérez-Ruzafa, I.M., 2011. Recent advances in coastal lagoons ecology: evolving old ideas and assumptions. *Transitional Waters Bull.* 5 (1), 50–74. <https://doi.org/10.1285/i1825229Xv5n1p50>.
- Petry, A.C., Guimarães, T.F.R., Vasconcellos, F.M., Hartz, S.M., Becker, F.G., Rosa, R.S., Goyenola, G., Caramaschi, E.P., Díaz de Astarloa, J.M., Sarmiento-Soares, L.M., Vieira, J.P., Garcia, A.M., Teixeira de Melo, F., De Melo, F.A.G., Meerhoff, M., Attayde, J.L., Menezes, R.F., Mazzeo, N., Di Dario, F., 2016. Fish composition and species richness in eastern South American coastal lagoons: additional support for the freshwater ecoregions of the world. *J. Fish Biol.* 89 (1), 280–314. <https://doi.org/10.1111/jfb.13011>.
- Reis, R.E., Albert, J.S., Di Dario, F., Mincarone, M.M., Petry, P., Rocha, L.A., 2016. Fish biodiversity and conservation in South America. *J. Fish Biol.* 89 (1), 12–47. <https://doi.org/10.1111/jfb.13016>.
- Ricklefs, R.E., 1987. Community diversity: relative roles of local and regional processes. *Science* 235 (4785), 167–171. <https://doi.org/10.1126/science.235.4785.167>.
- Ricklefs, R.E., 2004. A comprehensive framework for global patterns in biodiversity. *Ecol. Lett.* 7 (1), 1–15. <https://doi.org/10.1046/j.1461-0248.2003.00554.x>.
- Shepard, F.P., 1954. Nomenclature based on sand-silt-clay ratios. *J. Sediment. Petrol.* 24, 151–158. <https://doi.org/10.1306/D4269774-2B26-11D7-8648000102C1865D>.
- Teichert, N., Lepage, M., Chevillat, X., Lobry, J., 2018. Environmental drivers of taxonomic, functional and phylogenetic diversity (alpha, beta and gamma components) in estuarine fish communities. *J. Biogeogr.* 45 (2), 406–417. <https://doi.org/10.1111/jbi.13133>.
- Tittensor, D.P., Mora, C., Jetz, W., Lotze, H.K., Ricard, D., Berghe, E.V., Worm, B., 2010. Global patterns and predictors of marine biodiversity across taxa. *Nature* 466 (7310), 1098–1101. <https://doi.org/10.1038/nature09329>.
- Valencia-Méndez, O., Rodríguez-Zaragoza, F.A., Calderon-Aguilera, L.E., Domínguez-Domínguez, O., López-Pérez, A., 2018. Gamma-diversity partitioning of gobiid fishes (Teleostei: Gobiidae) ensemble along of Eastern Tropical Pacific: biological inventory, latitudinal variation and species turnover. *PLoS One* 13 (8), 1–20. <https://doi.org/10.1371/journal.pone.0202863>.
- Vilar, C.C., Joyeux, J.C., Giarrizzo, T., Spach, H.L., Vieira, J.P., Vaske-Junior, T., 2013. Local and regional ecological drivers of fish assemblages in Brazilian estuaries. *Mar. Ecol. Prog. Ser.* 485, 181–197. <https://doi.org/10.3354/meps10343>.
- Vorwerk, P.D., Froneman, P.W., 2009. The importance of estuarine-derived carbon for the nearshore marine environment: studies on two contrasting South African estuaries. *Afr. J. Aquat. Sci.* 34 (2), 137–146. <https://doi.org/10.2989/AJAS.2009.34.2.4.891>.
- Wagenmakers, E.J., Farrell, S., 2004. AIC model selection using Akaike weights. *Psychon. B. Rev.* 11 (1), 192–196. <https://doi.org/10.3758/BF03206482>.
- Walkley, A., Black, I.A., 1934. An examination of Degtjareff method for determining soil organic matter, and proposed modification of the chromic acid titration method. *Soil Sci.* 37, 29–38. <https://doi.org/10.1097/00010694-193401000-00003>.
- Wang, D., Gouhier, T.C., Menge, B.A., Ganguly, A.R., 2015. Intensification and spatial homogenization of coastal upwelling under climate change. *Nature* 518 (7539), 390–394. <https://doi.org/10.1038/nature14235>.
- Whitfield, A.K., 2017. The role of seagrass meadows, mangrove forests, salt marshes and reed beds as nursery areas and food sources for fishes in estuaries. *Rev. Fish Biol. Fish.* 27 (1), 75–110. <https://link.springer.com/article/10.1007/s11160-016-9454-x>.
- Wilsey, B.J., Chalcraft, D.R., Bowles, C.M., Willig, M.R., 2005. Relationships among indices suggest that richness is an incomplete surrogate for grassland biodiversity. *Ecology* 86 (5), 1178–1184. <https://doi.org/10.1890/04-0394>.
- Zhang, J., Crist, T.O., Hou, P., 2014. Partitioning of alpha and beta diversity using hierarchical Bayesian modeling of species distribution and abundance. *Environ. Ecol. Stat.* 21 (4), 611–625. <https://doi.org/10.1007/s10651-013-0271-2>.
- Zuur, A.F., Ieno, E.N., Elphick, C.S., 2010. A protocol for data exploration to avoid common statistical problems. *Methods Ecol. Evol.* 1 (1), 3–14. <https://doi.org/10.1111/j.2041-210X.2009.00001.x>.