Rarity and beta diversity assessment as tools for guiding conservation strategies in marine tropical subtidal communities

Lélis A. Carlos-Júnior1,2 | Matthew Spencer2 | Danilo Mesquita Neves3,4 | Timothy Peter Moulton1 | Débora de Oliveira Pires5 | Clovis Barreira e Castro5 | Carlos Renato Rezende Ventura5 | Carlos Eduardo Leite Ferreira6 | Cristiana Silveira Serejo5 | Simone Oigman-Pszczol7 | Fernanda Araújo Casares1,7 | Marcelo Checoli Mantelatto1 | Joel Christopher Creed1

Departamento de Ecologia, Universidade do Estado do Rio de Janeiro, Rio de Janeiro, Brazil
School of Environmental Sciences, University of Liverpool, Liverpool, UK
Department of Ecology and Evolutionary Biology, University of Arizona, Tucson, Arizona
Department of Botany, Federal University of Minas Gerais, Belo Horizonte, Brazil
Museu Nacional, Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brazil
Departamento de Biologia Marinha, Universidade Federal Fluminense, Niterói, Brazil
Instituto Brasileiro de Biodiversidade, Rio de Janeiro, Brazil

Abstract

Aim: Our aim was to uncover patterns of distribution of marine subtidal rocky reef communities across six taxonomic groups and decompose the relative roles of species loss and turnover in total community variation. Additionally, we propose an easily calculated index that can be used to highlight areas with unique species composition for conservation planning. We estimated the strengths of associations between environmental factors and species richness and rarity.

Location: Ilha Grande Bay, Brazil, covering about 150,000 ha harbouring different marine habitats.

Methods: We used the Marine Rapid Assessment Protocol at 42 sites to gather information on environmental variables and species in six subtidal marine groups. We determined “singular” sites as the regions harbouring higher numbers of rare species. Then, we estimated the roles of species loss and turnover on the observed total variation among sites. We used Generalized Linear Model to partition the relative importance of the selected environmental factors in driving variation in species richness and singularity.

Results: The singularity index and richness showed that the bay could be divided into three subregions for subtidal communities. Richness and rarity were structured at different spatial scales and associated with environmental variables related to water productivity and nutrients but varied among taxonomic groups. Community variation over space was largely associated with turnover of species.

Main conclusions: Higher singularity and richness on the western side of the bay and around the main island suggested that these regions should be conservation priorities, but high species turnover across the whole bay indicated that portions of the central channel should be included in conservation strategies. This draws attention to the importance of community variation rather than just species numbers in conservation and management planning. The high species turnover indicated that these
1 | INTRODUCTION

Our current knowledge of global biodiversity points to an ongoing major species-loss crisis (Pimm et al., 2014). Although this trend seems pervasive among different organisms and habitats (IUCN, 2014), the estimations are based on assessments using information on a fraction of the total number of species, many of which remain undescribed or lack distributional information (Carpenter et al., 2008; Peters, O’Leary, Hawkins, Carpenter, & Roberts, 2013). With many species yet to be discovered (Pimm et al., 2014) and the increasing rate of extinctions caused and/or exacerbated by anthropogenic activities (McCauley et al., 2015; Pandolfi, 2003), it is paramount to understand and explain diversity patterns across ecological systems (Von Der Heyden, 2011).

Lack of comprehensive distributional data leads marine species to be severely underrepresented. For example, according to the IUCN Red List, they comprise less than 12% of all studied taxa, although nearly a third of all eukaryotes are thought to be marine (IUCN, 2014; Mora, Tittensor, Adl, Simpson, & Worm, 2011; Peters et al., 2013). Few studies have tried to assess community organization in marine systems, which precludes strong inferences and robust syntheses (Heino et al., 2015; and see Moritz et al., 2013; Okuda, Noda, Yamamoto, Hori, & Nakaoka, 2010; Yamada, Tanaka, Era, & Nakaoka, 2014 for exceptions). This gap is detrimental not only to management/conservation efforts but also impairs the determination of what drives variation in diversity patterns in marine systems. Except for some general approaches and recent advances in inventorying databases (Briggs, 1974, 1995; Costello & Chaudhary, 2017; Costello et al., 2017; Spalding et al., 2007), most of our current biogeographical knowledge for marine ecosystems is still restricted to single taxonomic groups (e.g., bryozoans, Clarke & Lidgard, 2000; corals, Cornell, Arlson, & Hughes, 2007; fish Kulbicki et al., 2013), restricted to temperate, less diverse regions (Clarke & Lidgard, 2000) and/or does not account for differential responses amongst taxonomic groups (Soininen, 2014). Addressing these gaps is no easy task, but recent development in ecological analyses has provided the means to better explore the variety of biodiversity dimensions across multiple spatial scales.

One important trait of communities is the relationship between local (α) and regional (γ) diversity. Beta diversity was originally defined as “the extent of change in community composition” estimated from the ratio of gamma to alpha diversity (sensu Whittaker, 1960), although a variety of definitions were subsequently proposed (Anderson et al., 2011; Baselga, 2012; Tuomisto, 2010). We explored beta diversity (sensu Baselga, 2010, 2012, also defined as community turnover; see Tuomisto, 2010) patterns across the region by decomposing beta diversity into its nestedness and turnover components, the two distinct processes that cause variation in community composition, as explained elsewhere (Baselga, 2010; Baselga & Orme, 2012; Harrison, Ross, & Lawton, 1992). In nestedness, variation in composition between two or more sites occurs due to species loss or gain, such that species-poor sites are subsets of richer sites. Turnover is variation caused by the replacement of some species by others, usually associated with stochasticity and/or spatial/environment constraints (Baselga, 2010; Qian, Ricklefs, & White, 2005), including stressors and impact. Therefore, analysing beta diversity components also helps to recognize potential drivers of diversity differentiation among sites within a metacommunity, defined here as a set of local communities significantly linked by the dispersal of multiple species (Leibold et al., 2004).

Parallel to our considerations of beta diversity patterns in the marine benthos and reef fishes, we also wanted to identify areas characterized by faunas or florae composed of less frequent species. Our challenge was to propose a simple mechanism for assessing areas with high “rarity” in species composition when compared to other sites within the same metacommunity. The description of such locations is relevant for focussing management and conservation efforts, since human activities alter habitat availability and change species composition (Halpern et al., 2008; Pauly, Watson, & Alder, 2005). The concept of rarity is intuitive but often difficult to define, since there is a continuum from commonness to rareness (Usher, 1986). With that in mind, we wanted an index that was simple to interpret, especially by the non-scientific public, and was biologically meaningful. Also, we wanted to keep unavoidable subjectivity to a minimum in the mathematical designation of what “rarity” meant, making it clear, reproducible in other situations, and not strongly correlated with species richness in order to show patterns not necessarily caused by differences in the number of species.

Although described as the richest marine habitats (Costello & Chaudhary, 2017), tropical coastal areas are still under-studied (Cox, Spalding, & Foster, 2017; Kaehler & Williams, 1996) when compared to temperate shores or coral reefs (e.g., Mieszkowska et al., 2006). We performed descriptive analyses of diversity in marine subtidal rocky reef communities in a tropical region (sensu Spalding et al., 2007) of Rio de Janeiro state, Brazil. Our goal was to test: (a) whether it was possible to highlight distinctive areas, in terms of species spatial composition, especially those areas with higher prevalence of rare species, which

---

**Keywords**

alpha beta gamma diversity, benthos, community composition, marine community, marine ecology, metacommunities, rare species, tropical rocky reefs
we deemed singular areas; (b) the relative importance of environmentally versus spatially structured factors in driving variation in species richness and singularity; and (c) the relative roles of species turnover and nestedness in total beta diversity. For marine systems in general, those questions have hardly ever been investigated in an intergroup approach, let alone in the tropics. This is the first time the datasets available in Supporting Information (Tables S1 and S2) are published for an international readership (inventories were published in Portuguese, Creed et al., 2007) and their exploration will improve our understanding of tropical marine systems.

2 | METHODS

2.1 | Study site

Ilha Grande Bay (Baía da Ilha Grande—BIG, Figure 1) is located in the south of the state of Rio de Janeiro, southeast Brazil. The bay covers around 150,000 ha and is situated between the two most industrialized regions of the country—Rio de Janeiro and São Paulo. The diversity of different faunas/floras results from the distinctive geomorphology of the region, which harbours different types of terrestrial, freshwater and marine habitats, such as sand beaches, estuaries, mangroves and rocky reefs (Bastos & Callado, 2009). The bay’s location is associated with multiple potential anthropogenic pressures that threaten its diversity, such as intensive fishing, extensive occupation of shore areas, domestic and industrial waste, unregulated tourism, extensive circulation of ships and oil/gas platforms with several marinas and shipyards, and even the operation of an oil terminal and a nuclear power plant in Angra dos Reis, on the northern coast of the bay (near site 17 in Figure 1). The large centrally positioned island, Ilha Grande, has an important influence on the bay.

Using only nautical charts, 42 sites were pre-chosen in order that sampling sites would be distributed more or less evenly throughout the coastline and islands (n = 360) of the region and to represent a suite of differential environmental and subtidal marine benthic communities. Most sites had never been studied before. At all 42 sites (Figure 1), samples were taken to measure physiochemical properties of the water as well as to obtain information about sediment and geomorphology (Creed et al., 2007, chapters 4 and 5). In total, 31 environmental variables were measured, and they are available as supporting information, including a brief description of data collection (Supporting Information Table S2). In summary, the region is characterized by shallower waters on the west side of the bay, with deeper sites located in its central channel and on the outer side of the main island. The bottom temperatures sampled at these regions also differ considerably. The western side showed higher quantities of some types of sediment suspended in the water. This side is less exposed to wave action, whereas the southern sites of Ilha Grande and some exposed sites in the central channel were more exposed to wave action.

2.2 | Biological data collection

Species composition (presence/absence) data were collected in 2003–2004 by specialists using protocols developed for a Rapid Assessment Program (RAP) for three hard substrate (or hard/soft substrate interface) benthic groups (Macroalgae—hereafter called algae, Cnidaria—called corals henceforth although including some sea-anemones, and Echinodermata), two soft substrate benthic groups (Mollusca and Crustacea) and reef fish. All sampling was carried out using SCUBA. The RAP approach consists of short expeditions led by specialists into regions of biological importance in order to examine the status of the region’s biodiversity by selecting...
some groups which best represent the biota. The health of local ecosystems is also assessed, and management strategies proposed. Although not specifically designed for aquatic habitats, it has been used to assess marine systems around the world by Conservation International, who refer to it as the Marine Rapid Assessment Program (e.g., Dutra, Allen, Werner, & McKenna, 2005; McKenna & Allen, 2002, 2003).

For the benthos on hard substrate and on hard/soft substrate interface, the assessment was made through visual censuses on transects of approximately 100 m parallel to the coastline, and fish presence/absence was recorded using three 20 × 2 m transects per site. Each census was carried out from the littoral fringe to the depth at which the substrate changed from rock to soft bottom. As the selected sampling sites varied substantially in depth (min = 1 m, max = 27 m, mean = 10 m), which affects diving time, each dive was restricted to a minimum of 45 min and a maximum of 90 min to avoid significant differences in sampling effort. For detailed methods of data collection, see chapters 6, 7 and 11 in Creed et al. (2007). Corals could not be assessed at one site (17), so 41 sample sites are available for corals.

The benthos of soft substrata (Mollusca and Crustacea) were sampled using a sediment corer. At each site, five core samples (100 mm diameter × 150 mm height) were collected at each of two stations, one close to the rocky shore and the other 100 m away. The sediment was sieved, and fauna identified. For detailed methods, see chapters 8 and 10 in Creed et al. (2007).

We used species accumulation curves (Colwell, Chang, & Chang, 2004; Kindt, Van Damme, & Simons, 2006; Ugland, Gray, & Ellingsen, 2003) for all six taxa to ensure adequacy of our sampling effort.

### 2.3 An index for site singularity and richness

In order to identify distinctive sites, that is sites with more uncommon species composition, we devised "Singularity," a measure based on the number of rare species present at a local site within a metacommunity. We defined a rare species as one present at fewer than k out of n sites, where k is some number between 2 and the integer part of n/2. We defined the singularity of a site j (Sj) for a given rarity threshold as the proportion of species at that site that were rare. We used the proportion of rare species in order to avoid species richness of the site or individual taxonomic groups strongly influencing the results. In our study, we calculated the mean singularity value over all possible k thresholds, in order to avoid making an arbitrary choice of threshold. For thresholds above 4–5 sites (10%), the correlation between the mean singularity and the proportion of rare species at any given threshold was between 0.7 and 0.9 for all taxonomic groups. Thus, mean singularity was a good proxy for singularity over thresholds of rarity from 10% to 40% (4 and 17 sites, respectively) and therefore provided a good representation of rarity for our system. The R script for computing rarity for multiple thresholds, as well as checks on the performance of the mean singularity against any particular thresholds, is available as supporting information (S3).

Similarly, general (considering all taxa) richness was also determined for each site j taking into account the large intergroup variability in regional species richness. Let n_j be the number of species from group i at site j, n_i be the total number of species from group i in the region, and n_j^i be the total number of species at site j. Then, the proportion of species in group i that occur at site j is p_i^j = n_i^j / n_i, and the proportion of species at site j that come from group i is q_i^j = n_i^j / n_j. Then, we define the general richness R_j for t taxonomic groups (here t = 6) at site j as

\[
R_j = \sum_{i=1}^{t} p_i^j \cdot q_i^j
\]

Intuitively, R_j provides a measure of richness accounting for the large differences in species numbers observed among taxonomic groups at a given site, p_i^j.

We calculated general richness and singularity for all 42 sites, which led to an overall pattern that was visually consistent in our results (Figure 2): Relatively lower diversity in surveys found across the central core of the island, and higher diversity in surveys found around the main island and across the western sector of the BIG. To further explore these differences, we first classified geographically each of the 42 sites into subregions, namely central channel and northern sector (sites 18–29 and site 42), main island (sites 30–40) and western sector (sites 1–17 and site 41), comprising 13, 11 and 18 sites, respectively. We calculated summary statistics and produced boxplot visualizations to explore differences among the subregions. It was not appropriate to carry out a statistical test of the hypothesis that the three subregions differed in general richness and singularity because this hypothesis was only formulated after observation of the patterns in the data, which increases the chances of finding significance and violates assumptions of most a priori statistical tests, such as ANOVA (Kerr, 1998; Wasserstein & Lazar, 2016). The results of these comparisons are available in supporting information (S4).

### 2.4 Searching for drivers of richness and singularity patterns

We applied Generalized Linear Model (GLM)-based variation partitioning to account for the relative contribution of the selected environmental and spatially structure factors explaining variation in richness and singularity (GLMs with Gaussian error distribution). For explanatory variables, we used the environmental abiotic variables and Principal Coordinates of Neighbour Matrices as descriptors of spatial structure (PCNMs; Dray, Legendre, & Peres-Neto, 2006). We first computed PCNMs as described in Borcard and Legendre (2002), and only those describing positive spatial autocorrelation were retained (Borcard & Legendre, 2002). Briefly explained, the first step is to compute the Principal Coordinates Analysis (PCoA) of a matrix built from geographical distances among all sampling sites and truncated for distances larger than a cut-off set a priori to retain
only neighbouring distances. The eigenvalues of this PCoA describe orthogonal multiscale spatial variables. In other words, PCNMs are distance-based variables capable of describing spatial organization among sites at different spatial scales. For this dataset, 25 orthogonal spatial variables were generated. As explained elsewhere (Borcard & Legendre, 2002; Peres Neto, Legendre, Dray, & Borcard, 2006), larger eigenvalues are associated with broader spatial scale structures, while smaller eigenvalues represent fine-scale spatial structures. Therefore, we classified the PCNMs as broader (PCNMs 1–8), intermediate (PCNMs 9–17) and finer (PCNMs 18–25) spatial scales. Given our relatively high number of explanatory variables, we controlled for over-parameterization by applying a GLM-based variable selection approach, followed by progressive elimination of variables that showed high values of the variance inflation factor (VIF), maintaining only those with VIF < 2 (Table 1). The variable selection and variation partitioning were conducted using the “fields” (Nychka, Furrer, & Paige, 2015) and “vegan” (Oksanen et al., 2016) packages in the R Statistical Environment (R Core Team, 2017).

2.5 | Turnover × nestedness components of beta diversity

Operations on fractions were used to decompose total beta diversity, calculated as Sørensen dissimilarity index $\beta_{SOR}$, into the Simpson index $\beta_{SIM}$ describing spatial turnover without influence of richness gradients, and $\beta_{NES}$ describing variation in composition due to species loss or gain, causing compositions in species-poor sites to be nested within those of the richer sites (i.e., nestedness) (Equation 2)

$$\beta_{SOR} = \beta_{SIM} + \beta_{NES}$$ (2)

These calculations were conducted using the R package “betapart” (Baselga & Orme, 2012). We also calculated the same components for pairwise site comparisons, yielding 861 pairs of sites for the analysis of beta diversity for each group. For corals, only 41 sites were considered (yielding 820 pairs of sites) and for general integrative taxa measures, such as $S_i$ and $R_i$ we considered the number of corals to be zero at the sites where corals were not sampled. Therefore, caution should be taken when interpreting results for this particular sample unit.

3 | RESULTS

3.1 | Biological data collection

Across the 42 sites, 765 taxa (revised at the World Register of Marine Species—WoRMS) were recorded: 108 benthic algae, 26 cnidarians (Anthozoa and Milleporidae), 27 echinoderms from all five classes, 373 molluscs, 61 crustaceans and 170 reef fish (Supporting Information Table S1). For algae, this number is equivalent to one quarter of the whole known diversity of the state of Rio de Janeiro. Almost half (40%) of the crustaceans identified were new records either for BIG or the state of Rio de Janeiro. In Ilha Deserta (site 4), the presence of the fire coral Millepora alcicornis represented a new record for the region and the species’ new southern limit distribution. Species accumulation curves suggested that sampling was sufficient for most taxa, although infraunal groups (molluscs and crustaceans) seemed to be still slightly under-surveyed (Supporting Information Figure S5a-f).

3.2 | An index for site singularity and richness

In general, the western side of the bay and the sites around the main island had higher overall richness and higher singularity values when all taxa were considered together compared to the sites located in the central channel and the northern shore, but it varied considerably among different taxonomic groups (Figure 2; Supporting Information S4). On average, we expect a site chosen at random to have approximately one quarter (mean = 23%, $SD = \pm 4\%$) of the total species found in the bay, and that approximately a third of those species would be considered rare across the bay (30% ± 6%). Tanhangá Island, on the western side (site 14 in Figure 1), had the lowest general richness (less than 10%) but...
**Table 1** Values of selected explanatory variables after stepwise VIF selection

<table>
<thead>
<tr>
<th>(a)</th>
<th>Algal richness</th>
<th></th>
<th>Coral richness</th>
<th></th>
<th>Echinoderm richness</th>
<th></th>
<th>Mollusc richness</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>r²</td>
<td>VIF</td>
<td></td>
<td>r²</td>
<td>VIF</td>
</tr>
<tr>
<td>Chlorophyll a surface</td>
<td>0.11</td>
<td>1.091</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Salinity bottom</td>
<td>0.10</td>
<td>1.06</td>
<td></td>
<td>0.09</td>
<td>1.06</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phosphate bottom</td>
<td>0.08</td>
<td>1.06</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Selection coef. (near)</td>
<td></td>
<td></td>
<td>0.13</td>
<td>1.02</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Oxygen surface</td>
<td></td>
<td></td>
<td>0.07</td>
<td>1.04</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Org. matter (near)</td>
<td></td>
<td></td>
<td></td>
<td>0.17</td>
<td>1.03</td>
<td>0.09</td>
<td>1.28</td>
<td></td>
</tr>
<tr>
<td>Grain diameter (near)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.13</td>
<td>1.03</td>
<td></td>
</tr>
<tr>
<td>Interstitial water (far)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.12</td>
<td>1.18</td>
<td></td>
</tr>
<tr>
<td>Selection coef. (far)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.08</td>
<td>1.22</td>
<td></td>
</tr>
<tr>
<td>Secchi depth (horiz.)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nitrate surface</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inclination</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Oxygen bottom</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>KdV</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rugosity</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| (a) (cont.)                            | Crustacean richness |                | Fish richness |                | Total richness |                |                |                |                |
|                                        | r²       | VIF   |                | r²       | VIF   |                | r²       | VIF   |                | r²       | VIF   |                |
| Chlorophyll (surface)                  |                |                  |                |                  |                |                |                |                  |                  |                |                |                  |            |
| Salinity (bottom)                      | 0.08           | 1.32           |                |                |                  |                |                |                  |                  |                |                |                  |            |
| Phosphate (bottom)                     | 0.18           | 1.19           | 0.13           | 1.04            |                |                |                |                  |                  |                |                |                  |            |
| Selection coef. (near)                 |                |                  |                |                  |                     |                |                |                  |                  |                |                |                  |            |
| Oxygen (surface)                       |                |                  |                |                  |                     |                |                |                  |                  |                |                |                  |            |
| Org. matter (near)                     |                |                  |                |                  |                     | 0.20           | 1.35            |                |                  |                |                |                  |            |
| Grain diameter (near)                  |                |                  |                |                  |                     |                |                  |                |                  |                |                |                  |            |
| Interstitial water (far)               |                |                  |                |                  |                     |                |                  |                |                  |                |                |                  |            |
| Selection coef. (far)                  |                |                  |                |                  |                     | 0.10           | 1.08            |                |                  |                |                |                  |            |
| Secchi depth (horiz.)                  | 0.10           | 1.16           | 0.09           | 1.08            |                | 0.07           | 1.09            |                |                  |                |                |                  |            |
| Nitrile (surface)                      | 0.10           | 1.08           |                |                  | 0.10           | 1.09            |                | 0.10             | 1.08            |                |                |                  |            |
| Inclination                            | 0.07           | 1.09           |                |                  | 0.19           | 1.42            |                |                  |                | 0.09           | 1.46            |                |            |
| Oxygen (bottom)                        | 0.07           | 1.48           | 0.12           | 1.23            |                | 0.09           | 1.46            |                |                  |                |                |                  |            |
| Rugosity                               |                |                  |                |                  |                     |                |                  |                |                  |                |                |                  |            |

| (b)                                    | Algal singularity |                | Coral singularity |                | Echinoderm singularity |                | Mollusc singularity |                |
|                                        | r²     | VIF   |                | r²     | VIF   |                | r²     | VIF   |                | r²     | VIF   |                |
| Salinity (bottom)                      | 0.22   | 1.45           |                | 0.20   | 1.14           |                | 0.20   | 1.14           |                |            |                |            |            |
| Org. Matter (near)                     | 0.20   | 1.14           |                | 0.20   | 1.14           |                | 0.20   | 1.14           |                |            |                |            |            |
| Temperature (bottom)                   | 0.10   | 1.38           |                | 0.10   | 1.38           |                | 0.10   | 1.38           |                |            |                |            |            |
| Inclination                            | 0.09   | 1.18           |                | 0.09   | 1.18           |                | 0.09   | 1.18           |                |            |                |            |            |
| Interst. water (near)                  | 0.10   | 1.16           | 0.16           | 1.04           |                | 0.10   | 1.16           | 0.16           | 1.04           |                |            |                |            |
| Chlorophyll (surface)                  | 0.10   | 1.09           |                | 0.10   | 1.09           |                |            |                | 0.10   | 1.09           |                |            |
| Nitrile (surface)                      | 0.09   | 2.35           |                | 0.09   | 2.35           |                |            |                | 0.09   | 2.35           |                |            |
| Nitrile (bottom)                       | 0.07   | 2.48           | 0.08           | NA             |                | 0.07   | 2.48           | 0.08           | NA             |                |            |                |            |

(Continues)
### TABLE 1 (Continued)

<table>
<thead>
<tr>
<th>(b)</th>
<th>Algal singularity</th>
<th>Coral singularity</th>
<th>Echinoderm singularity</th>
<th>Mollusc singularity</th>
<th>(b) (cont.)</th>
<th>Fish singularity</th>
<th>Total singularity</th>
<th>(cont.)</th>
<th>r²</th>
<th>VIF</th>
<th>r²</th>
<th>VIF</th>
<th>r²</th>
<th>VIF</th>
</tr>
</thead>
<tbody>
<tr>
<td>Silt/Clay (far)</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>0.18</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Secchi depth (vert.)</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>0.14</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Depth</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>0.12</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Chlorophyll (bottom)</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Phosphate (surf)</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Grain diameter far)</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>CaCO₃ (far)</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>(c)</td>
<td>Algal richness</td>
<td>Coral richness</td>
<td>Echinoderm richness</td>
<td>Mollusc richness</td>
<td>Crustacean richness</td>
<td>Fish richness</td>
<td>Total richness</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Environment</td>
<td>0.20</td>
<td>0.30</td>
<td>0.16</td>
<td>0.19</td>
<td>0.09</td>
<td>0.15</td>
<td>0.29</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spatially struct. env</td>
<td>0.04</td>
<td>NA</td>
<td>0.09</td>
<td>0.08</td>
<td>0.08</td>
<td>0.16</td>
<td>0.07</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spatial variables</td>
<td>0.09</td>
<td>NA</td>
<td>0.13</td>
<td>0.01</td>
<td>0.15</td>
<td>0.01</td>
<td>0.15</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unexplained</td>
<td>0.68</td>
<td>0.70</td>
<td>0.62</td>
<td>0.72</td>
<td>0.67</td>
<td>0.68</td>
<td>0.49</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PCNM</td>
<td>25</td>
<td>None selected</td>
<td>10, 25</td>
<td>9</td>
<td>6, 14</td>
<td>4, 8</td>
<td>10, 25</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(c) (cont.)</td>
<td>Algal singularity</td>
<td>Coral singularity</td>
<td>Echinoderm singularity</td>
<td>Mollusc singularity</td>
<td>Fish singularity</td>
<td>Total singularity</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Environment</td>
<td>0.18</td>
<td>0.20</td>
<td>0.01</td>
<td>0.20</td>
<td>0.27</td>
<td>0.21</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spatially struct. env</td>
<td>0.17</td>
<td>0.13</td>
<td>0.07</td>
<td>NA</td>
<td>0.20</td>
<td>0.01</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spatial variables</td>
<td>0.21</td>
<td>0.07</td>
<td>0.15</td>
<td>NA</td>
<td>0.04</td>
<td>0.001</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unexplained</td>
<td>0.44</td>
<td>0.59</td>
<td>0.77</td>
<td>0.80</td>
<td>0.48</td>
<td>0.63</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PCNM</td>
<td>5, 9, 17</td>
<td>9, 10</td>
<td>5, 11</td>
<td>None selected</td>
<td>1, 23</td>
<td>4, 6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. Models used explanatory variables regressed against (a) richness and (b) singularity measures from each of the taxonomic groups and from overall community values. After the Generalized Linear Model, variation partitioning was performed for all models (c) in order to estimate relative contribution of environmental variables, spatially structured environmental variables, spatial autocorrelation (spatial variables) and unexplained variation to variation in richness and singularity. Last row of (c) depicts which Principal Coordinates of Neighbour Matrices (PCNMs) were selected by each taxonomic group. PCNMs are generated in descending order of spatial scale, meaning first PCNMs (e.g., PCNM 1 or PCNM 2) represent broader spatial scales when compared to the last PCNMs (e.g., PCNM 10). Column sums of fractions in (c) might not be exactly one due to rounding.
the highest singularity (53%). At a nearby site (Ponta do Pinto, site 7), proportional richness was 13%, whereas singularity reached 37%. Thus, some sites might not be particularly rich in species but nevertheless have unique species compositions compared to other more-enriched sites. There were also some higher values of singularity on the outer side of Ilha Grande, where sites were usually also species-rich (Figure 3). On the other hand, most sites located in the centre of the region showed relatively low values of singularity, despite varying proportions of richness. General richness had a significant but not strong correlation with singularity values (Spearman RS = 0.33, p = 0.03).

In addition to the general aspects of the marine diversity highlighted above, some taxon-specific attributes could also be distinguished (Figures 3 and 4). First, a high proportion of the richer and most singular sites of each taxonomic group were located on the western side of the bay (Figure 4a), similar to that observed for the general pattern. Second, there was a substantial variation among the different groups in regard to the spatial scale in which they were structured (Figure 4b).

3.3 | Searching for drivers of richness and singularity patterns

Thirty per cent of richness and 21% of singularity were not spatially structured and were associated with environmental differences across the bay (Figure 5). Both were mainly explained by differences in substratum: organic matter availability, sediment characteristics...
and geomorphology of the regions (Table 1a-b). For singularity, most of the environmental variation was structured at broader spatial scales, differentiating the western from the eastern side of the bay. In contrast, variation in richness was mainly driven by environmental factors that were spatially structured at intermediate and finer scales (last row of Table 1c). These fractions and the identity of the significant environmental drivers of variation in richness and singularity varied greatly across the taxonomic groups (Table 1).

### 3.4 Turnover × nestedness components of beta diversity

All six taxonomic groups exhibited high values of total beta diversity (which ranges from 0 to 1), around 0.9. These high values were almost entirely caused by spatial turnover of species (Table 2). The same pattern of dominance of spatial turnover in total beta diversity emerged from the distribution of all pairwise Sørensen dissimilarities (Figure 6) although pairwise comparisons yielded considerably higher variation.

### 4 DISCUSSION

Here, we have used species composition data to propose an integrative framework capable of improving the description of general patterns of richness and rarity and searching for potential drivers of such variation. Coupling this with the knowledge on which type of beta variation these communities present contributes to guide conservation strategies.

#### 4.1 Biological data collection

The RAP approach here described was the most comprehensive assessment of marine biodiversity ever made for the BIG region and one of the more extensive marine assessments to have been carried out in Brazil. The scale of the inventory can be observed in the numbers: 765 species inventoried, including several new records for the area, range expansions for numerous species and three new species discovered (Creed et al., 2007). There were new records for two mollusc genera in the Southwestern Atlantic, *Tornus* and *Eatoniella*, as well as three species being recorded in Brazil for the first time (*Macromphalina apexplanum*, *M. palmalitoris* and *Polygireulima amblytera*, Creed et al., 2007). Two new species of amphipod were recently described: *Puelche irenae* Nascimento & Serejo, 2018 and *Puelche longidactylus* Nascimento & Serejo, 2018. Both are typical burrowers (do Nascimento & Serejo, 2018). The datasets in the Supporting Information, therefore, provide distinctive data on tropical marine rocky reef communities. The singularity measurement here proposed suggested some areas differing in species composition, with the western side of the bay and around the main island comprising less frequently seen species in general and for several of the surveyed groups. Although for most groups the sampling was adequate, it would be productive to implement further expeditions, given that these datasets were collected over 10 years.
ago, especially focusing on species abundances. In this case, our analysis of these data is important to provide a baseline against which to measure recent changes. Further expeditions would be especially beneficial for soft substrate habitats, as these appeared to be slightly under-surveyed.

4.2 An index for site singularity and richness

Our method for computing rarity of taxa (i.e., small spatial range within the studied metacommunity) showed that the marine benthic/fish diversity could be divided into three sectors. The higher general singularity values found in the western side of the bay and around the main island are similar and they are different from the less-singular central core of the region, located between the main island and the continent, including the northern coastline (Supporting information S4). The central channel consists of locations with different levels of richness (structured at a finer scale, presumably due to local variations in habitat conditions), but mainly inhabited by common species. This could be an indication of a more stressed environment, since this region is the one under the most intensive anthropogenic pressures within the region (Creed et al., 2007). The taxa capable of living in the central channel of the bay are generally also the ones ubiquitous to the entire sampled region (Supporting information S6 shows ubiquity of the different species for all taxonomic groups). On the other hand, the western coast sector was characterized by sites with the highest ratio between singularity and richness (shown as small red spots in Figure 2). Therefore, this sector is composed of species not commonly seen elsewhere, showing considerable variation (i.e., high $\beta_{SOH}$) even among its own sites (results not shown here). These western communities also differ from the other highly singular communities found around the main island, comprising deeper locations. At those places, highly singular communities are also richer for several taxonomic groups (Figure 3a–f).

4.3 Searching for drivers of richness and singularity patterns

Variation in species richness and singularity across the BIG was mainly explained by variation in water- and substrate-associated conditions (Table 1). Indeed, the western (more singular) sector of the bay has more rivers and receives more sediments, nutrients and organic matter which may explain the observed changes in community composition. Additionally, variation in richness and singularity responded to geomorphology and sediment aspects of the rocky reefs. More
three-dimensionally complex habitats coincided with the same macro-
division observed for richness and singularity patterns. Therefore,
the combined effects of nutrient and organic matter enrichment and
higher rugosity on the western side of the bay, and along some of the
continental coastline and the outer side of the main island, were as-
associated with both richness and singularity patterns observed (51% and
38%, respectively, Figure 5). Using more restrictive thresholds for rar-
ity (e.g., considering “rare” those species occurring at one to four sites,
results not shown here) produced similar results, but increased the pro-
portional contribution of environmental variables to the explanation
of rarity patterns. This suggests that species rarity in the regional scale for
our system was strongly controlled by environmental filtering.

Both richness and singularity of several taxonomic groups were
spatially structured at different spatial scales, mainly at intermediate
and finer scales, represented by higher PCNMs (e.g., PCNMs 9, 10,
17, 25, see last row of Table 1c and Figure 4). This suggests that man-
agement actions aiming at particular taxonomic groups may require
a careful choice of spatial scale, which could be more complex than
targeting whole community conservation.

### 4.4 Turnover × nestedness components of beta diversity

The analysis of beta diversity in BIG revealed that variation in spe-
cies composition for all groups (Table 2; Figure 6) was high when
compared to other studied systems (e.g., Alsaffar, Cúrdia, Borja,

<table>
<thead>
<tr>
<th>Beta diversity</th>
<th>Total beta</th>
<th>Turnover</th>
<th>Nestedness</th>
</tr>
</thead>
<tbody>
<tr>
<td>Epifauna/flora</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Algae</td>
<td>0.93</td>
<td>0.90</td>
<td>0.03</td>
</tr>
<tr>
<td>Coral</td>
<td>0.90</td>
<td>0.82</td>
<td>0.07</td>
</tr>
<tr>
<td>Echinoderms</td>
<td>0.89</td>
<td>0.81</td>
<td>0.07</td>
</tr>
<tr>
<td>Infauna</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Molluscs</td>
<td>0.94</td>
<td>0.91</td>
<td>0.03</td>
</tr>
<tr>
<td>Crustaceans</td>
<td>0.97</td>
<td>0.95</td>
<td>0.02</td>
</tr>
<tr>
<td>Pelagic</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reef fish</td>
<td>0.93</td>
<td>0.89</td>
<td>0.04</td>
</tr>
</tbody>
</table>

Note. Due to rounding, the sum of the two components might be slightly
different from the total beta result.
Irigoien, & Carvalho, 2017; Magurran, Dornelas, Moyes, Gotelli, & McGill, 2015), including different taxonomic groups from tropical rain forests (e.g. Baselga, Gómez-Rodríguez, & Lobo, 2012; Tonial et al., 2012). In general, around 90% of species composition differs between local sites within the metacommunity, which means it was not possible to predict a site's composition with prior information from a different site. This, as a primary result, suggests tropical rocky reefs have high beta diversity, comparable to values found for plots with high beta diversity in tropical forests measured at a much broader spatial scale (Neves et al., 2017). This also has direct implications for conservation, since the loss of diversity at specific sites is relatively more troublesome, and it is not possible to encompass the whole regional diversity in a few geographically restricted protected areas. Interestingly, almost all variation in species composition is due to spatial replacement of species (turnover), with almost no contribution from species gain or loss (nestedness). This was also generally consistent within individual taxonomic groups, as seen by the centroid values in Figure 6, although it is possible to see a wider variation of values, which is in line with previous criticism on the usage of mean pairwise values for general inferences on multisite analysis (Baselga, 2012, 2013). Higher contributions of turnover to beta diversity have previously been suggested for other low latitude areas (below parallel 37, Baselga, 2012; Bishop, Robertson, Rensburg, & Parr, 2015, but see Neves et al., 2017) and could be related to different causes associated with spatially structured and historical constraints and/or different environmental selection (Baselga, 2010; Qian et al., 2005; Simpson, 1943). Indeed, further investigation revealed that environmental sorting, especially related to depth differences in the bay, is partly responsible for species variation (L.A. Carlos—Junior unpublished data), as well as differences in abundances (M.C. Mantelatto, unpublished data) in BIG. The high value of $\beta_{S\text{OR}}$ and its main component $\beta_{S\text{Sim}}$ in the bay also confirm that, in the marine environment, the gradients driving species variation change abruptly over relatively small spatial scales, revealing the importance of species sorting for community organization in the sea (Heino et al., 2015).

### 4.5 Conservation implications

The singularity and richness patterns, as well as their potential causes, have implications for current and future conservation strategies. Most importantly, marine communities on the west coast and around Ilha Grande (especially the southern side) may be best protected via several distinct yet connected protected areas (or a single large area) to encompass their community distinctiveness. Currently, the Tamboios Federal Ecological Reserve aims to protect a series of islands throughout the western portion of the region together with some specific conservation units, such as the Cairuco Federal Environmental Protected Area (EPA) and Bay of Paraty and Mamanguá Cove County EPA. Although the central channel had in general lower richness and singularity (Supporting information S4.1 and S4.2 panel a), the high values of species turnover observed for the whole area suggest that some portions of the central area should also be included in conservation plans. The observed higher spatial ubiquity (i.e., species with larger spatial ranges, Supporting information S6) of the species present at the central channel suggested these areas could be managed by preservation of smaller portions of its area. As discussed above, it could also suggest that the central channel is under the most intensive anthropogenic pressure, which is consistent with previous studies (Creed et al., 2007). Since the extension of a taxon's adaptation to a broader range of environmental conditions influences its geographical distribution (Holt, 2003; but see Carlos-Junior, Neves, Barbosa, Moulton, & Creed, 2015), the species capable of surviving in this region would also presumably be capable of inhabiting a larger range of environmental conditions across the whole bay.

### 4.6 Concluding remarks

Here, we have shown that although there were distinguishable patterns in both richness and singularity across different taxonomic groups, assemblages were structured by different environmental drivers and, most importantly, at different spatial scales. The contrasting spatial scales in which richness and singularity measures were structured for diverse taxonomic groups highlighted how diversity is organized differently in space for distinct fauna and flora, within the same habitat, such as the rocky reef benthos. Also, although some environmental drivers were found to be important to more than one group, there was a considerable difference in which factors influenced the observed variation in each group of species' richness and singularity (Table 1). Accounting for this plethora of possibilities increases complexity not only for the science of understanding spatial patterns in marine diversity, but also for developing management strategies. Nevertheless, there was a consistent pattern of turnover predominating in community variation, indicating that variability among assemblages is not determined by species loss but rather by substitution of species, which could be related to environmental filtering of different habitats across the bay and/or stochasticity driving immigration/local extinctions. Environmental drivers accounted for a considerable fraction of general variation in richness and singularity, confirming that species sorting in marine systems could be potentially high (Heino et al., 2015).

The method described above for computing rarity was adequate for identifying areas with unique compositions. Besides being consistent with other methods for calculating site endemism (results not shown here), it has the advantage of not being highly sensitive to richness. Independence of richness is a desirable characteristic for an index designed to detect patterns in community composition that are not necessarily the result of mere accumulation of different species. Moreover, the framework proposed here provides numbers that are easily interpretable and meaningful. For example, a site with $S_i = 0.5$ has half of its species considered "rare" for that region and is twice as singular as a site with $S^i = 0.25$. Interpretability and meaning are essential properties of useful diversity measures (Jost, 2006), which can be understood and applied even by non-ecologists, such as most political-decision makers. It is noteworthy that "rare" in this
context is related neither to overall distribution nor to abundance or endangered status. It refers solely to the frequency of the species’ occurrences within the target region. In other words, rare species were regarded as those with small spatial ranges, relative to the largest possible range given our study region. This is similar conceptually to Gaston (1994) and to other studies seeking for rarity in species ranges (see Tables 1.3 and 1.4 in Gaston, 1994). Nevertheless, it should be stressed that testing the abovementioned method under different scenarios and spatial scales could result in improvements. For example, we defined a rare species as one present at fewer than k out of n sites, where k is some number between 2 and the integer part of n/2. We then calculated the mean singularity value over all possible k thresholds as our proxy for rarity. This was the most objective concept of rare we could envision, as well as a general approach to rarity without compromising to a single (and potentially subjective) threshold. Although presumably permissive (considering most communities follow a log-normal distribution where most of the species occur in few sites) it worked well for our system with similar results to other indexes. Also, it worked as a good proxy for most thresholds, especially in the interval between 10% and 40% of the sites (4 and 17 sites, respectively, Supplementary Information S3). However, depending on the studied system, one specific threshold could be chosen as a cut-off for rarity. Another problem may arise in communities with unusually high proportions of rare species, as exemplified by our crustacean dataset. In those systems, singularity values get close (or, in our case, equal) to 1 and become a proxy for general richness (Rj), losing their utility. In summary, through a simple framework using presence/absence data, it was possible to recognize unique patterns that occur in beta diversity of the marine tropical shallow subtidal benthos. Furthermore, it was possible to identify mechanisms driving such patterns of community variation. Understanding better how these drivers operate should be a natural next step. It also remains to be tested whether the high beta diversity values observed here are unusual or are typical for lower-latitude marine systems. The framework and datasets provided here will be useful for answering those and other broader ecological questions.

ACKNOWLEDGEMENTS

We are grateful for insightful comments and suggestions on the singularity index given by H. Rolim from INEA-RJ, and T. Rocha and M. Broglio from UN’s FAQ. Also, we thank R. Villaça, J. Valentin, B. Rosado and E. Zandonà for several considerations on the early version of this manuscript. This study was funded by a Science without Borders fellowship from the National Council for Technological and Scientific Development (CNPq) and by the Brazilian Coordination for the Improvement of Higher Education Personnel (CAPES). LCJ acknowledges R. Athayde, E. Faria-Junior and K. Capel for the help with graphical issues. MCM acknowledges the support of CAPES.

DATA ACCESSIBILITY

Additional accessibility data are provided as supporting information.


**BIOSKETCH**

Lélis A. Carlos-Júnior is a postdoctoral researcher at the Universidade Federal do Rio de Janeiro, Brazil. He is mainly interested in how natural marine communities change over space and time, especially how ecological and biogeographical processes interact to shape distributions.

Authors contribution: LCJ, JCC, MS, SOP, FAC and TPM conceived the ideas. CRRV, CBC, CELF, CSS, DOP, FC, SOP and JCC collected the data. LCJ, JCC, MS and DRM designed the manuscript. LCJ and DRM analysed the data. LCJ, MCM and DRM prepared the figures. LCJ led the writing with valuable input of all contributors. All authors contributed equally to the final version of the paper.

**SUPPORTING INFORMATION**

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

**How to cite this article:** Carlos-Júnior LA, Spencer M, Neves DM, et al. Rarity and beta diversity assessment as tools for guiding conservation strategies in marine tropical subtidal communities. Divers Distrib. 2019;00:1−15. https://doi.org/10.1111/ddi.12896