Body size, reef area and temperature predict global reef-fish species richness across spatial scales

D. R. Barneche1,17 | E. L. Rezende2 | V. Parravicini3 | E. Maire6,7 |
G. J. Edgar8 | R. D. Stuart-Smith8 | J. E. Arias-González9 | C. E. L. Ferreira10 |
A. M. Friedlander11,12 | A. L. Green13 | O. J. Luiz14 | F. A. Rodríguez-Zaragoza15 |
L. Vigliola16 | M. Kulbicki4,5 | S. R. Floeter17

1College of Life and Environmental Sciences, University of Exeter, Penryn TR10 9FE, United Kingdom
2Facultad de Ciencias Biológicas, Departamento de Ecología, Center of Applied Ecology and Sustainability, Pontificia Universidad Católica de Chile, Santiago, Chile
3École Pratique des Hautes Etudes, CRIOBE, USR 3278 CNRS-EPHE-UPVD, Labex Corail, Université de Perpignan, Perpignan, France
4Institut de Recherche pour le Développement, UMR Entropie, Labex Corail, Université de Perpignan, Perpignan, France
5CESAB-FRB, Immeuble Henri Poincaré, Domaine du Petit Arbois, Aix-en-Provence cedex 3, France
6MARBECC, UMR IRD-CNRS-UM-IFREMER 9190, Université Montpellier, Montpellier Cedex, France
7Australian Research Council Centre of Excellence for Coral Reef Studies, James Cook University, Townsville, QLD, Australia
8Institute for Marine and Antarctic Studies, University of Tasmania, Hobart, TAS, Australia
9Departamento de Recursos del Mar, Centro de Investigación y de Estudios Avanzados del Instituto Politécnico Nacional, Mérida, Mexico
10Departamento de Biologia Marinha, Universidade Federal Fluminense, Niterói, Brazil
11Pristine Seas-National Geographic, Washington, District of Columbia
12Department of Biology, University of Hawaii, Honolulu, Hawaii
13The Nature Conservancy, Brisbane, QLD, Australia
14Research Institute for the Environment and Livelihoods, Charles Darwin University, Darwin, NT, Australia
15Departamento de Ecología, CUCBA, Universidad de Guadalajara, Zapopan, México
16Institut de Recherche pour le Développement, UMR Entropie, Labex Laboratoire Excellence Récifs Coralliens, Noumea, France
17Marine Macroecology and Biogeography Lab, Departamento de Ecología e Zoología., Universidade Federal de Santa Catarina, Florianópolis, Brazil

Correspondence
D. R. Barneche, College of Life and Environmental Sciences, University of Exeter, Penryn TR10 9FE, United Kingdom. Email: barnechedr@gmail.com

Abstract
Aim: To investigate biotic and abiotic correlates of reef-fish species richness across multiple spatial scales.
Location: Tropical reefs around the globe, including 485 sites in 109 sub-provinces spread across 14 biogeographic provinces.
Time period: Present.
Major taxa studied: 2,523 species of reef fish.
Methods: We compiled a database encompassing 13,050 visual transects. We used hierarchical linear Bayesian models to investigate whether fish body size, reef area, isolation, temperature, and anthropogenic impacts correlate with reef-fish species richness at each spatial scale (i.e., sites, sub-provinces, provinces). Richness was estimated using coverage-based rarefaction. We also tested whether species packing (i.e., transect-level species richness/m²) is correlated with province-level richness.
Results: Body size had the strongest effect on species richness across all three spatial scales. Reef area and temperature were both positively correlated with richness at all
spatial scales. At the site scale only, richness decreased with reef isolation. Species richness was not correlated with proxies of human impacts. Species packing was correlated with species richness at the province level following a sub-linear power function. Province-level differences in species richness were also mirrored by patterns of body size distribution at the site scale. Species-rich provinces exhibited heterogeneous assemblages of small-bodied species with small range sizes, whereas species-poor provinces encompassed homogeneous assemblages composed by larger species with greater dispersal capacity.

Main conclusions: Our findings suggest that body size distribution, reef area and temperature are major predictors of species richness and accumulation across scales, consistent with recent theories linking home range to species-area relationships as well as metabolic effects on speciation rates. Based on our results, we hypothesize that in less diverse areas, species are larger and likely more dispersive, leading to larger range sizes and less turnover between sites. Our results indicate that changes in province-level (i.e., regional) richness should leave a tractable fingerprint in local assemblages, and that detailed studies on local-scale assemblage composition may be informative of responses occurring at larger scales.

Key words
biogeography, community assembly, local diversity, neutral theory, regional diversity, spatial scale, species energy

1 | Introduction

Determining the proximal and ultimate causes of species richness is fundamental to understand why some regions can accommodate an extraordinary diversity whereas others contain just a few species (Lessard, Belmaker, Myers, Chase, & Rahbek, 2012). Studies have postulated that the dynamics of dispersal, speciation and extinction over large temporal scales shape regional pools of species, whereas, locally, species composition is presumably influenced more strongly by local ecological processes (e.g., mutualism, competition or predation; Ricklefs, 2004; Srivastava, 1999). However, disentangling how regional versus local processes contribute to community composition and species richness remains challenging (e.g., Cornell & Harrison, 2014), and yet it is crucial to understand why different regions exhibit such dramatic differences in biodiversity. In tropical coral reef systems, for instance, species numbers increase rapidly from small transects to whole islands, culminating in a bewildering regional richness; whereas in more temperate rocky reefs new species accumulate at a much slower pace as spatial scale increases, resulting in a substantially lower regional richness (Edgar et al., 2017; Witman, Etter, & Smith, 2004). We postulate that this pattern reflects intrinsic (e.g., body size and life-history strategies) and extrinsic factors (e.g., temperature and geographic barriers), and that their interaction ultimately explains the contrasting levels of species richness and spatial structure across provinces.

The dynamics of populations locally and along the meta-community continuum dictate how species accumulate across spatial scales. Population dynamics across space and time are fundamentally determined by resource demands and life history (Peters, 1983), dispersal capacity (Luiz et al., 2013), and local abundance (Reuman, Gislason, Barnes, Mélin, & Jennings, 2014), all of which are strong correlates of body size. In reef fishes, smaller species are more abundant (Ackerman, Bellwood, & Brown, 2004; Barneche, Kulbicki, Floeter, Friedlander, & Allen, 2016), have smaller home ranges (Nash, Welsh, Graham, & Bellwood, 2014), and have more limited geographic range when compared to larger species (Stier, Hein, Parravicini, & Kulbicki, 2014). Theory predicts, and evidence suggests that size-correlated traits, particularly abundance, home range, dispersal capacity and geographic range, should interact to define how the accumulation of species plays out across spatial scales, from small transects to entire biogeographic provinces (Allen & White, 2003; Belmaker, 2009; Brown & Nicoletto, 1991; Reuman et al., 2014). However, we still lack explicit tests of whether body size contributes to the accumulation of reef-fish species richness across spatial scales around the globe (but see Belmaker, 2009 for existing cross-scale comparisons).

In addition to body size, environmental and geographic factors are often invoked to explain gradients in species richness. For instance, the theory of island biogeography (MacArthur & Wilson, 1967) states that species richness increases with habitat area, and decreases with degree of isolation. In evolutionary time, the degree
of isolation among connected sites (Hubbell, 2001) will directly affect how new species accumulate in space (i.e., from sites to entire provinces), as species’ ability to disperse is expected to affect gene flow, vicariance, and ultimately speciation rates. Studies have shown that, in ectotherms, body size and environmental temperature can directly affect both dispersal capacity (Luiz et al., 2013; O’Connor et al., 2007) and speciation rates, via effects on individual metabolic rates (Allen, Gillooly, Savage, & Brown, 2006). In fact, models that combine thermal effects on speciation rates (Allen et al., 2006) with meta-community dynamics can reproduce realistic latitudinal diversity gradients currently observed in the world’s oceans (Tittensor & Worm, 2016; Worm & Tittensor, 2018).

Externally to “natural” factors, anthropogenic impacts can influence community composition due to over-harvesting, habitat degradation and introduction of non-native species, contributing to multiple processes such as trophic cascades and biotic homogenization (Dornelas et al., 2014; Jackson et al., 2001). For reef ecosystems in particular, recent studies show that population size of human settlements and accessibility to fish markets can have a negative effect on fish biomass (Cinner et al., 2016). Yet, it is still unknown whether such anthropogenic impacts exhibit consistent effects on species richness across spatial scales, which is important because, at present, extinctions have been observed at small scales but rarely within and across large biogeographic provinces (Kulbicki, Parravicini, & Mouillot, 2015).

To understand how these factors might affect species richness at different scales, here we present a global analysis of how reef-fish species richness builds up in space, from local sites to biogeographic provinces. We test whether body size, reef area (proxy for habitat availability) and isolation, human disturbance, and sea surface temperature (SST) correlate with species richness across spatial scales. Reef fishes provide an ideal model for investigating this problem because they are species-rich (Parravicini et al., 2013), globally widespread, and easy to sample with a high level of accuracy. Reef fishes also represent a major food source for millions of people (Teh, Teh, & Sumaila, 2013) and vary considerably in body size (Kulbicki et al., 2015). We show that, after controlling for the sampling-related group effects, and consistent with recent theoretical predictions (Allen & White, 2003; Reuman et al., 2014; Tittensor & Worm, 2016; Worm & Tittensor, 2018), body size, reef area and temperature are systematically correlated with species richness and accumulation across spatial scales. In light of our findings, we hypothesize that in less diverse areas, species are larger and likely more dispersive, leading to larger range sizes and less turnover between sites.

2 METHODS

2.1 Database and field sampling

We compiled a global database that encapsulates several decades of field data collection by several of the authors (Barneche et al., 2018). It encompasses 13,050 belt transects across 485 sites (islands, atolls and coastal contiguous reefs) spread through 14 tropical biogeographic provinces [Tropical Eastern Pacific, Offshore Tropical Eastern Pacific, Easter Island, The Hawaiian archipelago, Polynesia, Central Pacific, SW Pacific, Central Indo-West Pacific, Western Indian Ocean, NW Indian Ocean (Red Sea), Eastern Atlantic, Offshore SW Atlantic, SW Atlantic and the Caribbean; Figure 1, Supporting Information Table S1]. Sites span a 28-fold difference in species richness (Figure 1e).

Reef-fish assemblages were surveyed through belt transects of different areas depending on the data source (Supporting Information Table S1). Data from Cuba, Bahamas and Belize were collected as presence/absence data. Data from some sites along the Pacific coast of Mexico (area of 40 m²) and the Caribbean (area of 100 m²) were collected as total abundance counts. At all remaining sites, divers tallied the numbers and body lengths of all fish simultaneously. All transects started and ended approximately at the same depth (within 3 m of depth variation) and were oriented parallel to the reef. We only utilized transects conducted over hard-reef bottoms in order to make a representative comparison across provinces and reduce methodological bias generated by multi-habitat comparisons (Srivastava, 1999). Sites with fewer than three samples were excluded from the database. A total of 2,523 species was observed across all transects. To quantify overall species richness in these provinces and estimate the potential effects of sampling bias associated with field transects, we contrasted these data against an exhaustive checklist, that is, a compilation of published lists of species occurrences at multiple sites. Each site-specific list was assembled by combining multiple methods of sampling in order to obtain a thorough assessment of species richness, including that of small-bodied species. This checklist encompasses a total of 5,410 species (Parravicini et al., 2013; Supporting Information Table S2). Below we explain how we standardized the different transects for the purposes of estimating species richness at multiple scales.

We also compiled information on body size distribution, reef area and isolation, human disturbance (gravity of human settlement and gravity of human market, based on population size and travel time) and SST for each site, sub-province and biogeographic province. Succinctly, species-level maximum adult body size was obtained from the published literature and online databases, and the modal (i.e., the most frequent) size was calculated for each spatial scale [for visualization purposes, we sometimes divide these estimates into one of six classes following Parravicini et al. (2013): 0–7 cm, 7–15 cm, 15–30 cm, 30–50 cm, 50–80 cm and >80 cm]. Reef area [obtained from the Coral Reef Millennium Census Project and Halpern & Floeter (2008); see “Model predictors” section in the Supporting Information Appendix S1] was estimated in a 12-km radius around each site, and summed across sites within sub-provinces and sub-provinces within provinces to obtain estimates at higher scales, while reef isolation was calculated as the distance to the coast or the distance to the nearest reef. Human disturbance was estimated, following Maire et al. (2016) and Cinner et al. (2016), as the ratio between the population of the nearest major market or the nearest human settlement divided by the squared travel time to reach each site (i.e., gravity), averaged for analyses across localities and then provinces. And finally, SST was obtained from daily time-series data from the National Oceanic and Atmospheric Administration of the USA (NOAA) covering a 5-year period (°C; 0.25° resolution) (Reynolds et al., 2007; available...
FIGURE 1 Reef-fish species richness across scales. For clarity, the world map in (a) was divided into major geographic realms (dashed lines) that accommodate multiple sites (points) within biogeographic provinces (different colours and symbols). In (b–e), we show the species richness at the checklist (i.e., list of species occurrences; $n = 132$ sites, Supporting Information Table S2), province, sub-province and site scales, respectively, with points being ordered on the x axis based on their longitudes in (a). In (c–e), species richness was estimated using coverage-based rarefaction, and the mean coverage among points (±SD) is given at the top right corner [Colour figure can be viewed at wileyonlinelibrary.com]
from https://www.esrl.noaa.gov/psd/data/gridded/data.noaa.oisst.v2.highres.html). We describe in detail how these variables were estimated, the diagnostic analyses performed to ensure that our statistical models are unbiased, and the results robust to different sets of assumptions, in Supporting Information (Appendix S1: Table A1).

2.2 | Species richness at different scales

For the purposes of our study, we adopted three nested spatial scales: sites, sub-provinces and biogeographic provinces (see Supporting Information Table S1 for geographic coordinates). Sites are defined as small islands or stretches of continuous reefs in larger islands or coastlines (e.g., Arvoredo Island in southern Brazil, or reefs around Noumea in New Caledonia, i.e., c. 10s of km). Following Edgar et al. (2017), we aggregated sites if they were closer than 12 km in linear distance from one another. Sub-provinces encompass sites that belong to the same biogeographic sub-provinces (e.g., Cape Verde, Fiji, Meso-American Barrier in Mexico, i.e., c. 100s of km). Finally, biogeographic provinces are well-defined regions (i.e., >1,000s of km) that follow the recent classification proposed by Kulbicki et al. (2013) based on hierarchical analyses of reef-fish species composition. However, we further separated the SW Atlantic into continental (i.e., Brazilian coast) and offshore (oceanic islands) following Floeter et al. (2008).

We estimated species richness at the site, sub-province and province scales using coverage-based rarefaction (i.e., sampling completeness). Sampling coverage represents the proportion of the total number of individuals in a community that belong to the species represented in the sample (Chao & Jost, 2012). In contrast to classic sample-based or individual-based rarefaction methods, this coverage-based sampling technique provides more reliable estimates of species richness across communities. This is because while a fixed area or number of individuals may suffice to represent low-richness communities, it may be insufficient to represent species-rich communities (see details in Chao & Jost, 2012). Calculations were done using the R package iNEXT version 2.0.12 (Hsieh, Ma, & Chao, 2016).

Species richness (i.e., Hill number $D = 0$) was estimated by transforming all transect information into presence/absence data. This procedure, concomitantly with the hierarchical models accounting for differences in sampling method described below (see Statistical analyses), ensured that all species richness estimates are directly comparable. Estimates at the site, sub-province and province scales in these models were based on sampling coverages of 0.83, 0.89 and 0.98, respectively, that correspond to the highest coverage values yielding robust, unbiased estimations across all spatial scales according to the package’s algorithm (Chao & Jost, 2012; Hsieh et al., 2016). In other words, with this approach, we minimize much of the potential variation in species richness estimates due to differences in relative abundance or coverage sampling.

2.3 | Statistical analyses

At the site and sub-province scales, we test whether modal body size, $B$; reef area, $A$; distance to coast, $C$; distance to nearest reef, $R$; gravity of human markets, $M$; gravity of human settlement, $H$; and temperature, $T$, explain reef-fish species richness, $S$, using a hierarchical linear model,

$$\ln S = (\beta_0 + \Delta_g) + \beta_1 \ln B + \beta_2 \ln A + \beta_3 \ln C + \beta_4 \ln R + \beta_5 \ln M + \beta_6 \ln H + \beta_7 T + \epsilon$$

where $\beta_0$ is the estimated model intercept, $\beta_{BACRMHT}$ are fixed-effect slopes, $\Delta_g$ represents random deviations from the model intercept (on log scale) attributable to a grouping random variable that represents a higher spatial scale—that is, sub-provinces nested in provinces for the site scale, and provinces for the sub-province scale—and $\epsilon$ represents the residual variation. Note that by including $\Delta_g$, we are able to control for other unmeasured variables that might otherwise lead to correlated residuals at higher spatial levels (e.g., for spatial autocorrelation or sampling effects given that sub-provinces were generally studied by the same research group). Accordingly, the model residuals obtained at the site scale, controlling for the effects of sub-province nested within provinces, were not spatially autocorrelated according to Moran’s $I$ test ($p = 0.09$). Therefore, our hierarchical model adequately controls for the spatial structure of our sites. We note that species richness is functionally dependent on sampling area, and that perhaps our results could be biased by not including this covariate in the model. Our main findings with regards to the effects of body size, reef area and temperature are maintained even after including total sampling area as a covariate (Supporting Information Appendix S1: Table A4).

At the province scale, due to the small number of independent samples ($n = 14$), we fit separate linear regressions of the form

$$\ln S = \beta_0 + \beta_1 \ln B + \epsilon$$

In each regression, we respectively used modal body size, reef area, and temperature as predictors because they were the only variables with consistently strong effect sizes (i.e., slopes) both at the site and sub-province scales (see Results section below). Given that we only have 14 observations at this scale, adding multiple covariates at once would overfit the model.

We fit Equations 1 and 2 in a Bayesian framework by using the R package brms to derive posterior distributions and associated 95% credible intervals (CIs) for the fitted parameters. Both fixed ($\beta_{BACRMHT}$) and random ($\Delta_g$) effects were assigned normally distributed priors that were vague (i.e., locally uniform over the region supported by the likelihood), with means of zero. Model residuals ($\epsilon$), and standard deviations for both the random effects ($\sigma(\Delta_g)$) and model residuals ($\sigma(\epsilon)$), were also assigned weakly informative priors following a Student’s $t$ distribution. The posterior distributions of model parameters were estimated using Markov chain Monte Carlo (MCMC) methods by constructing three chains of 2,000 steps, including 1,000-step warm-up periods, so a total of 3,000 steps were retained to estimate posterior distributions (i.e., $3 \times (2,000 – 1,000) = 3,000$). Fixed effects were considered statistically significant if their posterior 95% CIs did not overlap zero. We use Bayesian $R^2$ in order to estimate the amount of explained variation of each model (Gelman, Goodman, Gabry, & Ali, 2017). Posterior predictive checks for all three models are provided in Supporting Information Appendix S1: Figure A1.
We also tested whether reef-fish species richness at the transect scale (i.e., species packing = species/m²) was correlated with richness at the province level. To do so, we used a two-step approach. First, we ran a Bayesian hierarchical linear model with richness at the transect scale as a function of transect area on a log-log scale. We do so given the nonlinear nature of the species-area relationship.

**FIGURE 2**

Fitted data based on our Bayesian hierarchical linear model showing the effects of modal body size (a–c), reef area (d–f), and temperature (g–i) on reef-fish species richness (from Figure 1c–e) at the site (left column, Equation 1), sub-province (middle column, Equation 1) and province (right column, Equation 2) scales. Panels at the site (a, d, g) and sub-province (b, e, h) scales depict the partial effects of each variable after correcting species richness for the effects of all remaining fixed-effect covariates. At the province scale, three separate models have been fitted, each for a different predictor. At the site and sub-province scales, species richness has been corrected for the random effects (respectively sub-province nested in province, and province). Bottom left values represent mean (±SD) coverage among data points (Cov.), which are the same for plots within the same column. Bayesian \( R^2 \) are displayed with associated 95% credible intervals (CIs). Panels (d) and (g) are generated from the same model as panel (a), and therefore contain the same Bayesian \( R^2 \) [similarly for panels (b, e, h)]. Panels (c), (f) and (i) are generated from three separate models as explained in the text, and therefore have different Bayesian \( R^2 \) values. Thin dashed lines show 95% Bayesian CIs around the mean model prediction. Colours and symbols as in Figure 1. SST = sea surface temperature. [Colour figure can be viewed at wileyonlinelibrary.com]
Sites nested in sub-provinces nested in provinces were used as a random effect. Model fitting specifications (number of MCMC samples, chains, warm-up periods, and prior distributions) follow the same specifications of the model in Equation 1. Because the random-effect coefficients represent deviations from the estimated model intercept, at each one of the spatial scales we summed each random-effect coefficient with the model intercept in order to obtain a measure of species packing (species/m²). Second, we ran ordinary least squares for each of the 3,000 posterior samples from our random-effect coefficients at the province scale (i.e., species packing on the log scale) against province-level species richness (also on the log scale)—as estimated using coverage-based rarefaction (coverage = 0.98). This approach allowed us to estimate a mean slope between species packing and province-level richness, and $R^2$ values with associated 95% CIs. A slope of 1 would indicate that species packing is directly proportional to province-level richness (i.e., a linear relationship), whereas a slope $>0$ and $<1$ would indicate that species packing scales sub-linearly with province-level richness.

All analyses were repeated removing the smallest size class (<7 cm) to circumvent potential problems of under-detection during sampling (Supporting Information Appendix S1: Table A2). We also fit the model at the site scale separately for the Atlantic and Americas as a group (i.e., including the Tropical Eastern Pacific), and the rest

**FIGURE 3** Violin plots showing the difference in species richness among provinces for each of the six body size classes considered in our study, which are respectively indicated at the top right corner in italics. Each violin represents a distribution of sites within a given province (bottom points as in Figure 1). We filtered our data for each class, and estimated the species richness using coverage-based rarefaction, and the mean coverage among points ($\pm$SD) is given in the top right corner of each plot. Black dashed lines separate major geographic realms [Colour figure can be viewed at wileyonlinelibrary.com]
3 | RESULTS

3.1 | Site scale

Our model explained 73.6% of the variation in reef-fish species richness across sites (Equation 1), which ranged from 10 to 288 species. After controlling for all covariates, reef-fish species richness decreased with modal maximum body size (Figure 2a). After binning species richness based on maximum body sizes, we observe that small-bodied species prevail in species-rich sites, particularly in the Pacific region (Figure 3b,c). We observed that the rate of increase in biodiversity with sampled area (i.e., the number of transects) is substantially higher in species-rich provinces (Figure 4a). However, species packing increased sub-linearly with province-level richness (slope = 0.26, 95% CI = 0.15–0.36; $R^2 = 0.55$, 95% CI = 0.28–0.78; Supporting Information Appendix S1: Figure A2). Moreover, the observed modal body size is consistently smaller (i.e., low mean value and lower variances, Figure 4b,c) within the richest provinces. In particular, the estimated slope for the mode of the body size distribution was the strongest among all estimated slopes ($\beta_m \sim 0.70$; 95% CI: −0.87 to −0.51), implying a 74.0% decrease in richness over the observed range of modal body sizes (from 67 to 9.65 cm).

Species richness increased with reef area ($\beta_A$: 0.05; 95% CI: 0.02–0.08; Figure 2d), and decreased with distance to the closest reef ($\beta_R$: −0.02; 95% CI: −0.03 to −0.00). Our model also indicated that species richness increases with SST ($\beta_T$: 0.05; 95% CI: 0.01–0.09; Figure 2g), implying a 52.4% increase in richness going from 21.3 to 29.73°C.

Variables related to human disturbance (gravity of markets and gravity of nearest human settlement) and distance to coast were not significantly correlated with species richness (respectively $\beta_M$: 0.00; 95% CI: −0.02–0.02; $\beta_H$: 0.00; 95% CI: −0.01–0.01; $\beta_C$: 0.02; 95% CI: −0.00–0.04). The hierarchical structure of our model, with

![Figure 4](https://example.com/figure4.png)

**FIGURE 4** Richness and body size across different spatial scales. In (a), sample-based rarefaction curves (based on 500 permutations) of reef-fish species richness show how reef-fish species richness accumulates as sampling effort increases in different provinces. Mean richness is shown (±95% confidence intervals). Mean modal body size (b) and variance in size (c) of the species pool across different sampling areas, and across sites within our checklist (from Figure 1b; Supporting Information Table S2). Body size means and variances at each province were estimated from randomly permuted subsets of transects based on 1,000 permutations (1, 5, 10 or 15 transects), or across all transects (All), or across all sites for the checklist data (Checklist). Colours and symbols as in Figure 1 [Colour figure can be viewed at wileyonlinelibrary.com]
sub-provinces nested in provinces as our random-effect group, was able to capture considerable variation in species richness at different scales. Particularly, after accounting for the fixed effects, we observed a 1.55-fold variation among sub-provinces on average (i.e., \( \approx e^{0.22 \times 2} \) with \( \sigma_2 = 0.22 \)).

The direction and significance of fixed-effect coefficients were all similar after removing the smallest size class (<7 cm; Supporting Information Appendix S1: Table A2). Results for the Atlantic + Americas, and the rest of the Pacific, show that the mean model coefficients and their directions (i.e., positive or negative) are maintained within the Pacific, but are not significant within the Atlantic + Americas (Supporting Information Appendix S1: Table A6).

### 3.2 | Sub-province scale

At the sub-province scale, reef-fish species richness ranged from 15 to 364 species. Our model (Equation 1) explained 78.8% of the variation in species richness. Consistent with the model at the site scale, after controlling for all covariates, reef-fish species richness decreased with modal maximum body size (Figure 2b). Species turnover from sites to sub-provinces was most pronounced for small-bodied species (Supporting Information Appendix S1: Figure A3). The modal-size estimated slope was again the strongest among all estimated slopes (\( \beta_g = -0.77; 95\% \text{ CI}: -1.11 \text{ to } -0.43 \)). Apart from modal maximum body size, reef area (Figure 2e) and temperature (Figure 2h) were the only other significant variables in our model (\( \beta_A = 0.08; 95\% \text{ CI}: 0.04 \text{ to } 0.12; \beta_T = 0.05; 95\% \text{ CI}: 0.00 \text{ to } 0.11 \)). Our model indicates that after accounting for the fixed effects, species richness varies by 2.41-fold among provinces (i.e., \( \approx e^{0.44 \times 2} \) with \( \sigma_2 = 0.44 \)).

After the removal of the smallest size class, modal body size, reef area and temperature were still significant and in the same direction as the main model. (Supporting Information Appendix S1: Table A2).

### 3.3 | Province scale

Consistent with our findings at the site and sub-province scales, body size (Equation 2) explained 61.2% of the variation in species richness (ranging from 44 to 806 species) at the province scale. Particularly, province-level richness decreased with increasing modal maximum body size (\( \beta_g = -2.79; 95\% \text{ CI}: -4.14 \text{ to } -1.43 \); Figure 2c). Similar effects were detected after removing the smallest size class (\( \beta_g = -2.68; 95\% \text{ CI}: -4.04 \text{ to } -1.35 \); Supporting Information Appendix S1: Table A2). We also analysed the relationship between species richness and area (Figure 2f) or temperature (Figure 2i) in two separate linear regressions (Supporting Information Appendix S1: Table A5), and both variables were positively correlated with species richness. Reef area was positively correlated with richness (\( \beta_A = 0.22; 95\% \text{ CI}: 0.08 \text{ to } 0.37 \); Figure 2f), and explained 46.5% of the variation in species richness. Temperature was also positively correlated with richness (\( \beta_T = 0.31; 95\% \text{ CI}: 0.00 \text{ to } 0.62 \); Figure 2f), and explained 30.7%.

### 4 | DISCUSSION

Our results indicate that reef-fish modal body size, reef area, and environmental temperature are consistently correlated with reef-fish species richness across spatial scales (Figure 2, Supporting Information Appendix S1: Table A5). The fact that body size was negatively correlated with species richness at all spatial scales may suggest that the observed accumulation of species across space is strongly associated with smaller body sizes (Figure 3). This interpretation is consistent with previous studies showing that beta diversity across scales is driven by the accumulation of smaller species in mammals (e.g., Brown & Nicoletto, 1991) and reef fishes (e.g., Belmaker, 2009). Moreover, our empirical results lend support to the "marine diversity spectrum" theory proposed for pelagic marine ecosystems (Reuman et al., 2014), whereby the negative linear relationship between richness and body size arises due to the size dependence of home range and dispersal capacity (Allen & White, 2003; Reuman et al., 2014) that has been previously observed in reef fishes (Luiz et al., 2013; Nash et al., 2014).

We note that sites that are rich in smaller-bodied reef fishes also encompass a range of suitable, heterogeneous habitats (e.g., hard and soft corals, sponges, algae; Messmer et al., 2011), and our results indicate that increasing habitat area will also yield higher species richness across scales. We note that our analysis cannot resolve the causality of these relationships given that shallow-reef organisms (e.g., reef-building corals and reef fishes) show a parallel history of diversification through time (Bellwood, Goatley, & Bellwood, 2017). However, the fact that body size varies with richness even within a gradient of species-rich sites and provinces (e.g., from Indo-West Pacific to Polynesia) suggests that the evolutionary processes associated with the diversification of reef fishes have aided speciation within smaller-bodied species. In particular, the centre of origin and accumulation hypothesis states that the richer Indo-West Pacific is a product of a series of events that facilitated both the speciation ("cradle") as well as the accumulation ("museum") of existing species through time (Bellwood et al., 2017). Our findings could be consistent with this hypothesis considering that Indo-West Pacific reefs have gone through a series of sea-level changes leading to temporarily variable geographic isolation (Bellwood et al., 2017). Therefore, speciation may have been further promoted by vicariance because small-bodied species have, on average, lower realised dispersal capacity (Luiz et al., 2013).

After accounting for the effect of sampling area, species richness at the transect scale scaled sub-linearly with province-level richness (Supporting Information Appendix S1: Figure A2). The sub-linear nature of this relationship corroborates the suggestion of recent studies that species richness estimated at small spatial scales might not necessarily resemble that at the province scale (Dornelas et al., 2014; Vellend et al., 2013). We speculate that this observed phenomenon reflects the interaction between body size, local abundance and geographic range (Supporting Information Appendix S1: Figure A4a), as smaller species are generally more abundant (Ackerman et
al., 2004; Barneche et al., 2016) and often exhibit smaller geographic range sizes than larger species (Belmaker, 2009; Luiz et al., 2013). Consequently, it is expected that species at the extremes of the body size distribution are detected in only a small fraction of transects: for small species this may relate to smaller home range sizes and difficulties in visual detection, whereas for large species low abundance should result in a small representation across those transects within their geographic range (Preston’s veil; Preston, 1948). Our data strongly support this interpretation, with both the smallest and largest size classes being less frequently sampled in nearly all provinces, except in Easter Island (Supporting Information Appendix S1: Figure A4b). Accordingly, while population abundance was negatively correlated with body size across sites in a subset of the data used here (Barneche et al., 2016), the number of provinces in which each species was detected was positively correlated with their maximum size (Supporting Information Appendix S1: Figure A4c), indicating that smaller species tend to have smaller geographic ranges (Supporting Information Figure Appendix S1: A4d).

Because of the interactions between size, abundance and distribution, species of intermediate body size constitute the core of most local assemblages in all provinces (Supporting Information Figure Appendix S1: A4b). However, body size distributions also differ significantly among provinces concomitantly with species richness. The linear regression between total species richness against modal body size at the province scale (Equation 2, Figures 2c and 5a) gives rise to a continuum, with species-poor peripheral provinces of the Atlantic falling at one extreme and highly complex reefs such as those in the Indo-West Pacific, Polynesia and, to a lesser degree, the Caribbean at the other (Figure 5a). Thus, the body size distribution within local assemblages is indicative of the overall reef-fish richness of their respective provinces. This can be readily understood because medium to large fishes are expected to contribute to species richness at the transect level, without a proportional contribution to species richness at higher spatial scales, if they are widely spread (see also Reuman et al., 2014; Soininen, Lennon, & Hillebrand, 2007). As a consequence, in provinces where larger fishes predominate, local assemblages should be more homogeneous in their composition and exhibit body size distributions shifting towards higher values (Figure 5b). The opposite is expected in provinces occupied predominantly by small species, which include a multitude of species with small geographic ranges. Again, this implies that the enormous diversity of species-rich provinces emerges primarily from the accumulation of smaller species with restricted distribution (Figure 5b).

Our models at all scales indicate that temperature was positively correlated with species richness, consistent with the idea that the latitudinal diversity gradient is in part driven by the temperature dependence of ectothermic metabolism and speciation rates (Allen et al., 2006; Tittensor & Worm, 2016). Additionally, our results add to a vast number of studies that support predictions from the theory of island biogeography (MacArthur & Wilson, 1967), whereby species richness increases with increasing habitat area, and decreases with isolation. For instance, it has been shown recently that smaller-bodied reef-fish prey declines more rapidly with increasing isolation when compared to larger predators across Pacific reefs (Stier et al., 2014). We note though that reef isolation was not significant at the sub-province scale. This discrepancy between scales could be due to our choice of measurement for isolation (mean distance to the nearest reef across sites). Given this caveat, our results are consistent with recent theoretical predictions (Tittensor & Worm, 2016; Worm & Tittensor, 2018) that suggest that the combined effects of area,

**Figure 5** The negative relationship between modal body size and species richness across provinces (a). In (b), we propose conceptually that the pattern in (a) arises from the different contribution of large (top) versus small (bottom) species to the observed mean body size and local (but not province-level) richness. Distributions (blue for large species, red for small species) represent hypothetical species-specific local abundances across space. Local richness can be sampled at any point along this continuum, whereas the province-level richness arises from the sum of all species (i.e., distributions). For the purposes of illustration, we represent the low and high richness provinces respectively with large and small species only, but notice that in reality provinces are made up of a multitude of species with varying abundance and geographic ranges. Colours and symbols in (a) as in Figure 1 [Colour figure can be viewed at wileyonlinelibrary.com]
traits in the species pool (e.g., small body size), and warmer temperatures will yield higher beta diversity across spatial scales, perhaps by promoting higher speciation rates over time, and higher species turnover across space (Allen et al., 2006; MacArthur & Wilson, 1967; Tittensor & Worm, 2016; Worm & Tittensor, 2018).

After accounting for the effects of reef-fish body size, reef area and isolation, temperature, sampling artefacts, and other unmeasured potential biogeographic effects, human-related disturbance metrics did not show any significant effect on reef-fish species richness at the site and sub-province scales. This finding is particularly relevant because it suggests that, at least looking at the studied spatial scales, species richness, contrary to assemblage size structure and biomass (e.g., Cinner et al., 2016; Jackson et al., 2001), is not explained by the "gravity" variables, which are based on population size and travel time. This may in large part be a result of the relatively low proportion of species that are exploited, with the rich array of small fishes being less likely to be affected by humans through indirect mechanisms only (such as habitat alteration). It may also be attributable to the fact that our data collection is recent in time, and it is possible that the majority of sampled reefs are already affected by human disturbance to some degree. However, our results might suggest that habitat degradation associated with loss of coral species in species-rich provinces may have profound impacts on species richness (Alvarez-Filip, Dulvy, Gill, Côté, & Watkinson, 2009) via its impact on smaller species. On the other hand, larger fishes might be able to disperse more easily towards other areas, most likely because of their better capacity to survive and establish reproductive populations (Luiz et al., 2013). Comparisons between Pacific provinces, the Tropical Eastern Pacific and the Caribbean support this conjecture. The expansion of a mosaic of reef habitats in the Indo-Australian Archipelago during the Oligocene/Miocene was a significant driver of cladogenesis for coral-reef-fish taxa (Bellwood et al., 2017), with smaller-sized lineages with low mobility and small home ranges radiating in multiple provinces (Munday & Jones, 1998). Before the closure of the Isthmus of Panama, the Caribbean had a reef-building coral fauna that more closely resembled modern Indo-Pacific species. It further underwent a period of extensive faunal loss that has resulted in the modern-day Atlantic and Eastern Pacific fish faunas (Bellwood & Wainwright, 2002; Budd, 2000; Di Martino, Jackson, Taylor, & Johnson, 2018). The predominance of intermediate to large species across both sides of the Isthmus of Panama (Figure 4b) may have resulted both from long-term and more recent loss of habitat (Alvarez-Filip et al., 2009). This provides circumstantial evidence that assemblage structure is altered by habitat loss in a predictable manner (Wilson, Graham, Pratchett, Jones, & Polunin, 2006). Although speculative at present, this hypothesis can readily be tested in future studies. In fact, recent evidence indicates that small species might suffer the most pronounced lags to tracking global changes (Graham et al., 2007).

Here we demonstrate that in reef fishes, fish body size, reef area, and temperature can empirically predict species richness across spatial scales remarkably well. Therefore, our study lends strong support to theories that predict geographic gradients in species richness based on the combination of demographic processes that depend on habitat availability, size-dependent traits such as home range, dispersal and geographic range, and physiological processes such as the temperature-dependence of metabolism and speciation rates (Allen & White, 2003; Allen et al., 2006; Reuman et al., 2014; Tittensor & Worm, 2016). We hypothesize that species richness at small spatial scales builds up into province-level species richness at rates that are inversely related with the geographic range size of its constituent species, everything else being equal; in other words, reefs with larger, better dispersing species should exhibit less beta diversity because these species are found in all localities. In the absence of good proxies of abundance and range size in other taxonomic groups, these patterns that might be quite general across systems may remain unnoticed. Finally, at short temporal scales, activities that impact both small- (e.g., habitat degradation) and large-sized (e.g., overfishing) species might affect gradients in species richness in predictable ways. At longer temporal scales, the above effects combined with novel temperature regimes due to climate change should set the fate of environmental gradients in species richness.

ACKNOWLEDGMENTS

We thank the French Foundation for Research on Biodiversity (FRB) and the Center of Synthesis and Analysis of Biodiversity (CESAB) for funding the GASPAR working group from which this study was generated. We also thank A. P. Allen, L. C. Strotz, J. Belmaker, J. E. Williamson and M. Westoby for comments on earlier drafts. Thanks also to D. Tittensor, F. A. Zapata, J. P. Quimbayo and J. P. Krajewski who shared their data, and all field assistants. ELR was supported by a Bolsa Jovem Talento (CNpq 402053/2012-5), grant FBO02 for the Center for Applied Ecology and Sustainability, line 3, and Fondo de Ciencias y Tecnología, Chile (FONDECYT 1170017). Financial support from Brazil: SISBIOTA-Mar (PI: S. R. Floeter CNpq 563276/2010-0 and FAPESC 6308/2011-8), Pro-Africa (PI: C. E. L. Ferreira CNpq 495031/2007-5), CAPES, Marinha do Brasil, Instituto Laje Viva; France: Institut de Recherche pour le Développement (IRD), PRISTINE project (Fondation Total); Mexico: L. E. Calderón-Aguilera (CICESE), A. L. Cupul-Magaña, C. M. Galván-Villa (Universidad de Guadalajara), H. R. Reyes-Bonilla, A. Ayala-Bocos (UABCS), R. A. López-Perez (UMAR), E. Nuñez (UNACAR), fieldwork supported by CONABIO (PI: L. E. Calderón-Aguilera project HJ026), PROMEP (PI: F. A. Rodríguez-Zaragoza, projects 103.5/08/2919 and 103.5/10/927) and SEP-CONACYT programmes (PI: J. E. Arias-González); Honduras: A. Medina, P. Rico (WWF-CA), M. Aronne (Fundación Coyo Cochinos), fieldwork supported by WWF-CA, USAID, FFEM and Universidad de Guadalajara (PI: F. A. Rodríguez-Zaragoza); Nicaragua: R. Barnuti (INPESCA), fieldwork supported by WWF-CA, USAID and Universidad de Guadalajara (PI: F. A. Rodríguez-Zaragoza); USA: National Geographic Society (PI: S. R. Floeter; A. M. Friedlander).

AUTHOR CONTRIBUTIONS

DRB, SRF and ELR conceived the study. All authors helped to collect and compile data. DRB and ELR analysed the data and wrote the first draft. All authors contributed to substantial revisions.
DATA ACCESSIBILITY


ORCID

D. R. Barneche https://orcid.org/0000-0002-4568-2362
S. R. Floeter https://orcid.org/0000-0002-3201-6504

REFERENCES


