nature ecology & evolution

ARTICLES

https://doi.org/10.1038/s41559-022-01710-5

Check for updates

Biological trade-offs underpin coral reef ecosystem functioning

Nina M. D. Schiettekatte[1,2,3]✉, Simon J. Brandl[4], Jordan M. Casey[4], Nicholas A. J. Graham[5], Diego R. Barneche[6,7], Deron E. Burkepile[8,9], Jacob E. Allgeier[10], Jesús E. Arias-Gonzaléz[11], Graham J. Edgar[12], Carlos E. L. Ferreira[13], Sergio R. Floeter[14], Alan M. Friedlander[15], Alison L. Green[16], Michel Kulbicki[2,17], Yves Letourneur[2,18], Osmar J. Luiz[19], Alexandre Mercière[1,2], Fabien Morat[1,2], Katrina S. Munsterman[10], Enrico L. Rezende[20], Fabian A. Rodríguez-Zaragoza[21], Rick D. Stuart-Smith[12], Laurent Vigliola[2,17], Sébastien Villéger[22,23] and Valeriano Parravicini[1,2,23]

Human impact increasingly alters global ecosystems, often reducing biodiversity and disrupting the provision of essential ecosystem services to humanity. Therefore, preserving ecosystem functioning is a critical challenge of the twenty-first century. Coral reefs are declining worldwide due to the pervasive effects of climate change and intensive fishing, and although research on coral reef ecosystem functioning has gained momentum, most studies rely on simplified proxies, such as fish biomass. This lack of quantitative assessments of multiple process-based ecosystem functions hinders local and regional conservation efforts. Here we combine global coral reef fish community surveys and bioenergetic models to quantify five key ecosystem functions mediated by coral reef fishes. We show that functions exhibit critical trade-offs driven by varying community structures, such that no community can maximize all functions. Furthermore, functions are locally dominated by few species, but the identity of dominant species substantially varies at the global scale. In fact, half of the 1,110 species in our dataset are functionally dominant in at least one location. Our results reinforce the need for a nuanced, locally tailored approach to coral reef conservation that considers multiple ecological functions beyond the effect of standing stock biomass.

The flow of elements through biological communities fuels all ecosystems on Earth, and humans are increasingly threatening biodiversity and the persistence of these fluxes[1,2]. Coral reefs are a prime example of an ecosystem that is severely impacted by anthropogenic activities, and drastic declines in habitat quality and fish biomass have evoked serious concerns about the persistence of coral reefs[3,4]. Maintaining ecosystem functions, defined as fluxes of elements, is a major goal for coral reef conservation[5–7]. However, past evaluations of functions on coral reefs have mostly relied on static proxies such as live coral cover, standing stock biomass of reef fishes or measures of diversity[8–10]. These simplified proxies, although useful, may not properly represent ecological functions because fluxes of elements can scale nonlinearly with variables such as biomass[11]. Therefore, improving the quantification of ecological functions constitutes an important step towards the efficient management of coral reef ecosystem functioning[7].

As a dominant group of consumers, coral reef fishes are essential vectors of carbon (C), nitrogen (N) and phosphorus (P)[11–13]. Ecosystem functions mediated by coral reef fishes include nutrient cycling, biomass production, herbivory and piscivory (secondary consumption)[7]. Although the high diversity of coral reef fishes has inspired many studies that focus on ecosystem functioning, only a handful of studies have attempted to quantify functions as continuous fluxes[7]. Furthermore, studies that have quantified functions as a flow of energy and nutrients have mostly focused on single functions (for example, biomass production[14,15] or fish excretion[13]), covering only a small number of species at local scales. Consequently, trade-offs among multiple functions, their drivers and their vulnerability

[1]PSL Université Paris: EPHE-UPVD-CNRS, USR 3278 CRIOBE, Université de Perpignan, Perpignan, France. [2]Laboratoire d'Excellence 'CORAIL', Perpignan, France. [3]Hawai'i Institute of Marine Biology, University of Hawai'i at Mānoa, Kāne'ohe, HI, USA. [4]The University of Texas at Austin, Marine Science Institute, Port Aransas, TX, USA. [5]Lancaster Environment Centre, Lancaster University, Lancaster, UK. [6]Australian Institute of Marine Science, Crawley, Western Australia, Australia. [7]Oceans Institute, The University of Western Australia, Crawley, Western Australia, Australia. [8]Department of Ecology, Evolution, and Marine Biology, University of California, Santa Barbara, Santa Barbara, CA, USA. [9]Marine Science Institute, University of California, Santa Barbara, Santa Barbara, CA, USA. [10]Department of Ecology and Evolutionary Biology, University of Michigan, Ann Arbor, MI, USA. [11]Laboratorio Ecología de Ecosistemas de Arrecifes Coralinos, Departamento Recursos del Mar, CINVESTAV Unidad Mérida, Mérida, Mexico. [12]Institute for Marine and Antarctic Studies, University of Tasmania, Hobart, Tasmania, Australia. [13]Departamento de Biologia Marinha, Universidade Federal Fluminense (UFF), Niterói, Brazil. [14]Marine Macroecology and Biogeography Lab, Depto. de Ecologia e Zoologia, CCB, Universidade Federal de Santa Catarina, Florianopolis, Santa Catarina, Brazil. [15]Department of Biology, University of Hawaii, Honolulu, HI, USA. [16]Red Sea Research Center, King Abdullah University of Science and Technology, Thuwal, Saudi Arabia. [17]Institut de Recherche pour le Développement, UMR UR-IRD-CNRS-IFREMER-UNC ENTROPIE, Nouméa, New Caledonia, France. [18]Université de la Nouvelle-Calédonie, UMR UR-IRD-CNRS-IFREMER-UNC ENTROPIE, Nouméa, New Caledonia, France. [19]Research Institute for the Environment and Livelihoods, Charles Darwin University, Darwin, Northern Territory, Australia. [20]Departamento de Ecología, Center of Applied Ecology and Sustainability (CAPES), Facultad de Ciencias Biológicas, Pontificia Universidad Católica de Chile, Santiago, Chile. [21]Laboratorio de Ecosistemas Marinos y Acuicultura (LEMA), Departamento de 12 Ecología, CUCBA, Universidad de Guadalajara, Zapopan, Mexico. [22]MARBEC, Université de Montpellier, CNRS, IFREMER, IRD, Montpellier, France. [23]These authors contributed equally: Sébastien Villéger, Valeriano Parravicini. ✉e-mail: nina.schiettekatte@gmail.com

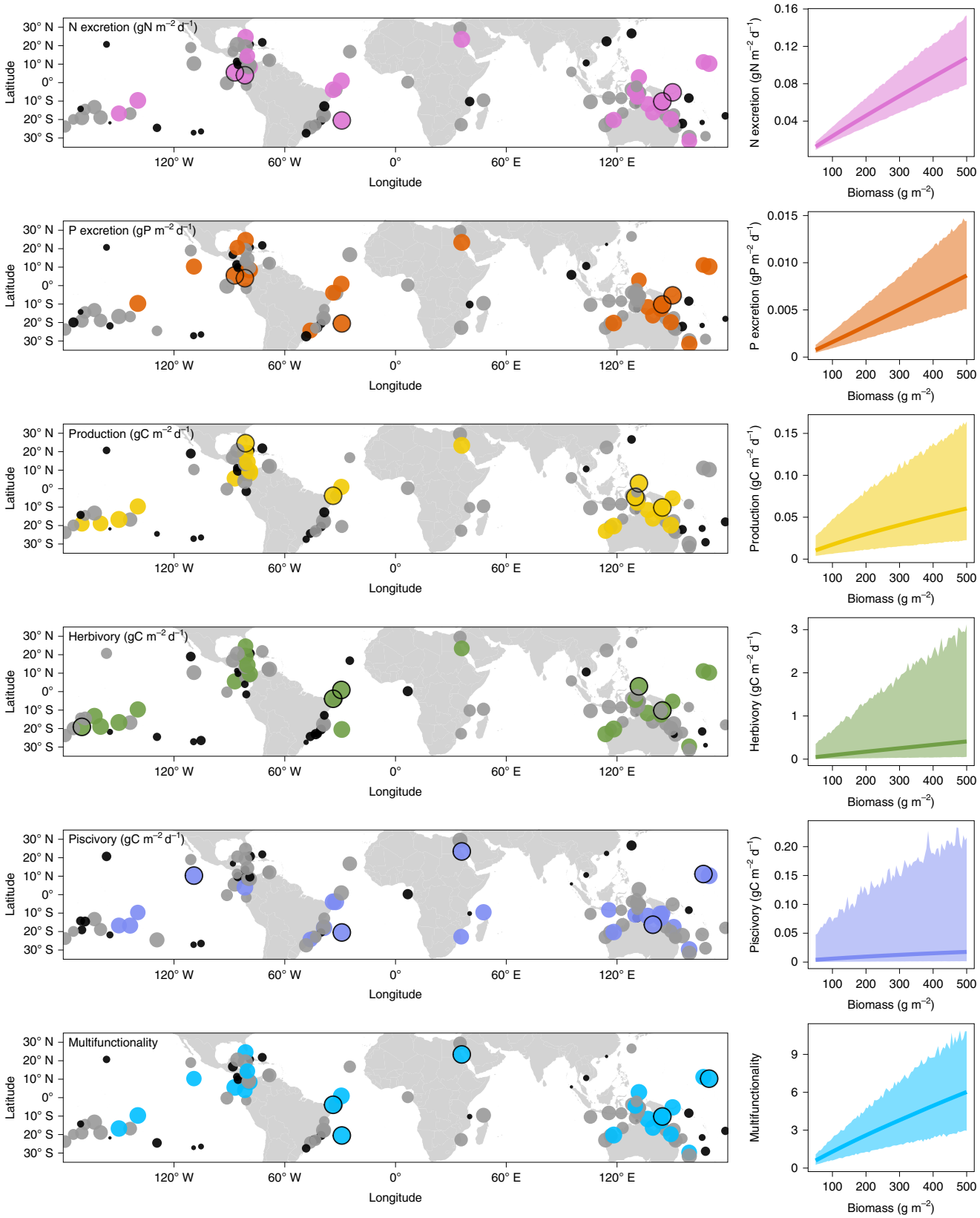


Fig. 1 | Maps of the five key ecosystem functions, multifunctionality and the relationships between the functions and biomass. Left: dots indicate localities of field surveys, with dot sizes representing the ranked values of the locality-level predictions of functions, and colour scales showing categorical assignments (black, ≤25%; grey, 25–75%; colour, ≥75%). Black outlines highlight the five localities with the highest values of each function. Multifunctionality represents the geometric mean of the five normalized functions. Right: the back-transformed predicted values for functions and multifunctionality with increasing biomass. The lines represent the average modelled relationship, and the shaded areas show the 95% CIs of the predictions. All relationships between functions and biomass are nonlinear.

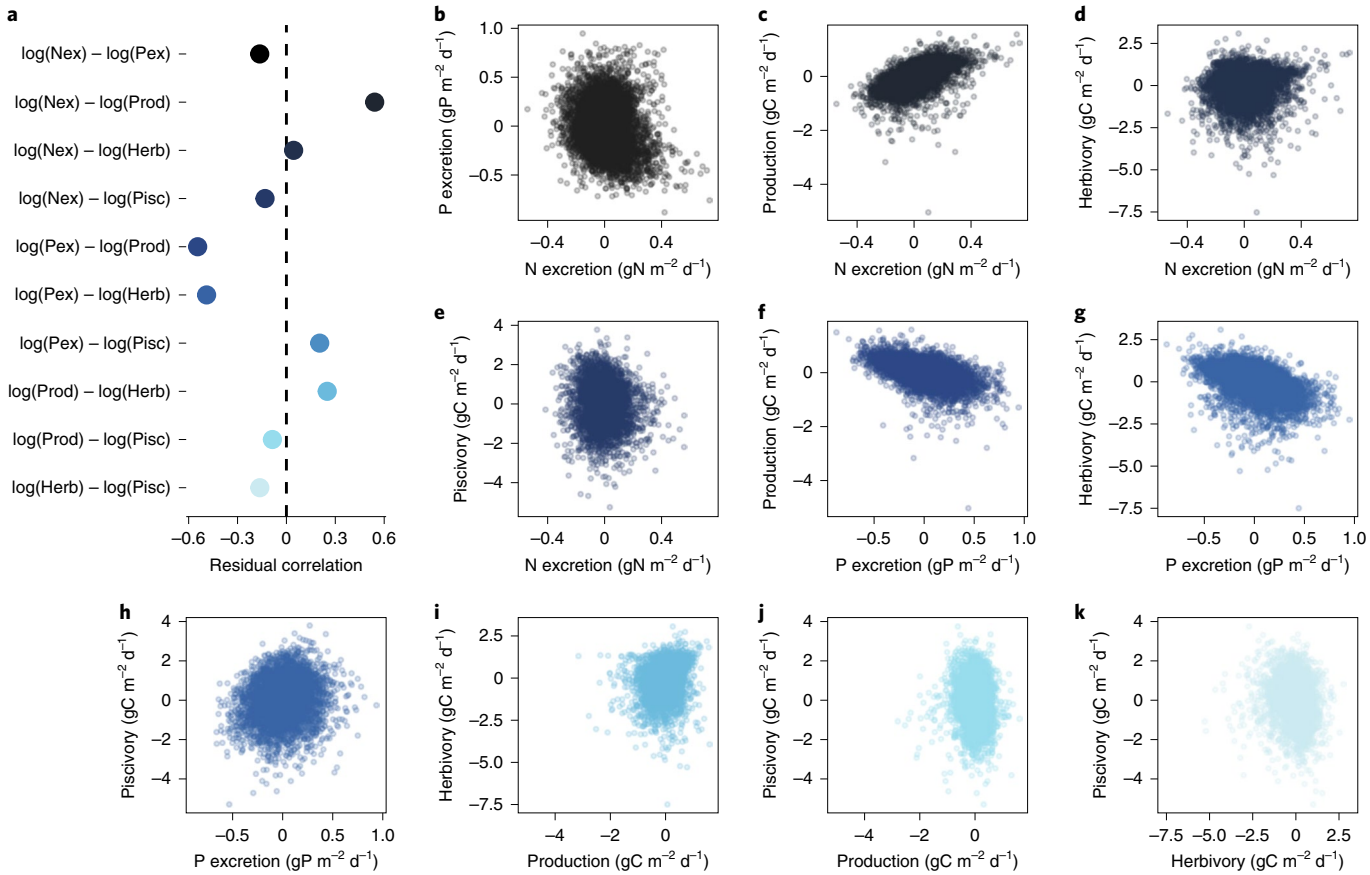


Fig. 2 | Correlations of the five functions, accounting for biomass and SST. a, Modelled correlation coefficients of residual errors. Dots represent the average, and the 95% CI is too narrow to be shown. **b**–**k**, Scatter plots of the mean residual errors of the functions of P excretion (Pex) (**b**), production (Prod) (**c**), herbivory (Herb) (**d**) and piscivory (Pisc) (**e**) as a function of N excretion (Nex); production (**f**), herbivory (**g**) and piscivory (**h**) as a function of P excretion; herbivory (**i**) and piscivory (**j**) as a function of production; and piscivory as a function of herbivory (**k**).

to anthropogenic stressors remain poorly understood in coral reef ecosystems across large spatial scales[7].

In this Article, we integrate biogeochemistry and community ecology to advance understanding of the elemental fluxes that underpin reef fish functioning. Using empirical species-specific data on basic organismal processes and extrapolation with Bayesian phylogenetic models, we parameterize individual-level bioenergetic models to estimate five key ecosystem functions for 1,100 species: N excretion, P excretion, biomass production, herbivory and piscivory. We apply these bioenergetic models to 9,118 reef fish transects across 585 sites at 98 localities (that is, regions encompassing sites that belong to the same biogeographic sub-province) worldwide (Supplementary Table 1) to (1) quantify community-level reef fish functions, (2) investigate trade-offs among functions and (3) extract the community and species-level effects on these functions.

Results

We estimated five key ecosystem functions mediated by coral reef fishes across the globe (Fig. 1). Across localities, all five functions show similar geographical patterns with on average higher values around the Equator. However, at the global scale, no location displayed high levels (that is, top 5%) of functioning across all functions. Therefore, multifunctionality (that is, the geometric mean of the five normalized functions) does not appropriately represent the state of all functions assessed independently. For example, although the northern Coral Sea had the highest multifunctionality value, piscivory in this location was 40% less than its global maximum value.

Biomass is the most commonly employed indicator of coral reef functioning[7,8], and we demonstrate a predictably strong relationship between fish biomass and all five functions (Fig. 1). Specifically, in a multivariate mixed-effects Bayesian model, the slopes of log-transformed biomass were 0.932 (95% credible interval (CI): 0.929, 0.934) for N excretion, 1.051 (CI: 1.047, 1.056) for P excretion, 0.771 (CI: 0.764, 0.780) for production, 0.940 (CI: 0.923, 0.957) for herbivory and 0.668 (CI: 0.635, 0.702) for piscivory. These slopes indicate that the relationships between biomass and functions are all allometric, which demonstrates that biomass is not an appropriate proxy for functioning. We also incorporated sea surface temperature (SST) due to its impact on the metabolism and growth of individual fishes, which scales up to the community[16]. We found positive effects of SST on N excretion, production and herbivory but no effects of SST on P excretion and piscivory (Supplementary Table 3).

Our multivariate model also allowed us to estimate the correlations among functions, independent of the effects of biomass and SST. In particular, we estimated correlations among functions on three levels: locality effects, site effects and residual variation (Fig. 2 and Extended Data Figs. 1 and 2). The correlations displayed similar patterns on each level. We found negative trade-offs between P excretion and N excretion as well as between P excretion and biomass production. Furthermore, we found slightly weaker negative correlations between piscivory and N excretion as well as between

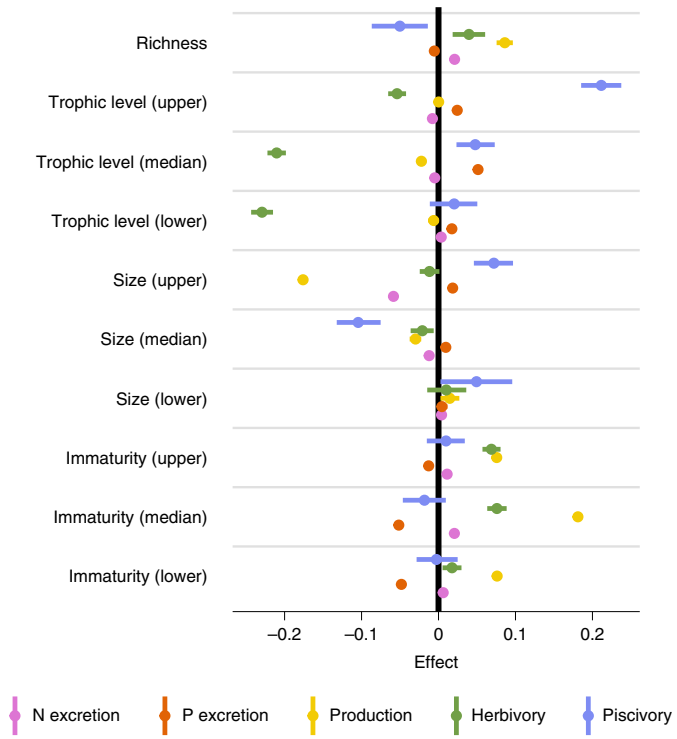


Fig. 3 | Effects of ecological community variables on the five functions. Dots indicate fixed effect values from Bayesian linear regressions that examine the effects of species richness, trophic level, size and immaturity of fishes. To represent both the median and spread of trophic level, size and immaturity across individuals within a community, we included lower and upper 95% quantile values of these three traits as community variables. All data were log transformed and standardized to compare across functions and variables (see Supplementary Table 4 for parameter values on non-standardized data). Dots represent the average effect size estimate, and horizontal lines indicate the 95% CI. Immaturity is defined as the derivative of the von Bertalanffy growth model for a given size; thus, the higher the value, the younger the individual.

piscivory and herbivory. Thus, a reef fish community cannot simultaneously display high values of functioning across all investigated functions.

To determine how community structure affects the variation and trade-offs of functions beyond the effects of biomass and SST, we ran a multivariate Bayesian mixed-effects model with ten variables that describe the structure of each fish assemblage: species richness and the median, lower and upper 95% quantiles of size, immaturity (that is, a measure combining relative size and growth rate; Methods) and trophic level of individuals in a community. Each of these variables has non-zero effects on at least one of the five functions, suggesting that the observed trade-offs may be, at least in part, rooted in the structure of the focal community (Fig. 3 and Supplementary Table 4). Some associations, such as the negative and positive effects of trophic level on herbivory and piscivory, respectively, are intuitive, whereas others, such as the negative effect of immaturity on P excretion, are not immediately obvious (Fig. 3).

Beyond community structure, we examined whether functions are driven by particular species across sites. We quantified the degree of functional dominance (that is, disproportionately large contributions by species to a given function) inside each community at the site level and found that, on average, functions are dominated by a small fraction of species in each community (Fig. 4a). We also calculated the proportion of species that is dominant in at least one site (that is, species with a disproportionately high contribution as compared with a community in which all species contribute equally), and we found that 49% of all species contributed disproportionately to a function in at least one surveyed site (Fig. 4b). However, very few species are dominant throughout their range (Fig. 4c). Thus, functions within communities tend to be driven by few dominant species, but the identity of those dominant species varies across sites.

Discussion

By quantifying five key processes mediated by coral reef fishes, we demonstrate that coral reef ecosystem functioning is shaped by biological trade-offs, local community structure and species identity. Standing biomass is one of the most commonly employed indicators of coral reef functioning[7,8], and our analyses confirm the strong influence of biomass on all other processes. However, our results also show nonlinear relationships between functions and biomass and illustrate a high degree of residual variation, unexplained by biomass. This suggests that biomass alone does not sufficiently characterize functioning; indeed, strong trade-offs occurred among the five functions independent of biomass. Thus, using biomass as a proxy may mask differences in community-level functioning. Furthermore, for a given value of biomass, no reef can yield above-average values across all five functions. Although a reef may stand out as a hotspot for one function, no reef can simultaneously maximize all functions.

The observed trade-offs among functions are driven by fish community structure and the organismal physiology and life-history traits of its constituents[17,18]. For example, we observed a clear trade-off between P excretion and biomass production, which is mostly driven by community age and trophic structure (Fig. 3). Communities dominated by fishes with high trophic levels are characterized by high P excretion rates because predatory fishes have a P-rich diet[13]. In contrast, biomass production is high in communities dominated by fishes that occupy low trophic levels because herbivores tend to exhibit higher growth rates[19]. Moreover, P is retained for skeletal growth in young fishes, thus limiting P excretion rates[17,20]. Metabolic theory predicts that small-bodied individuals have higher mass-specific metabolic rates, leading to elevated consumption rates and disproportionate contributions to functions that rely on rapid energetic turnover, such as herbivory, piscivory, production and N excretion[15,21,22].

Our results reveal that functions consistently rely on a few dominant species, but the identities of local, dominant species strongly vary across sites[23]. Locally, a small number of high-performing taxa may disproportionately impact rates of functioning at the community level due to high biomass or abundance[24], which may have led to their designation as functionally dominant 'key species' in various locations[25]. However, our results also reveal that no species dominated throughout its geographical range, and half of all species contributed disproportionately to a specific function in at least one site. Thus, it is not possible to pinpoint widespread key species that dominate functioning throughout their range and can be placed at the centre of conservation guidelines; rather, identifying local species dominance across functions may be the best approach for small-scale conservation efforts, and the preservation of regional reef fish biodiversity should be prioritized based on broad-scale policy.

Our global analysis of multiple functions suggests pathways by which human-induced shifts in reef fish community structure may impact coral reef ecosystems. Fishing and climate-induced coral loss have caused declines in reef fish biomass and shifts in community structure[26,27], and we suggest that these changes will differentially affect ecosystem functioning. Intensive fishing and associated reductions in the biomass of large fishes, for example, alters the size, age and trophic structure of fish communities[27]. When accounting

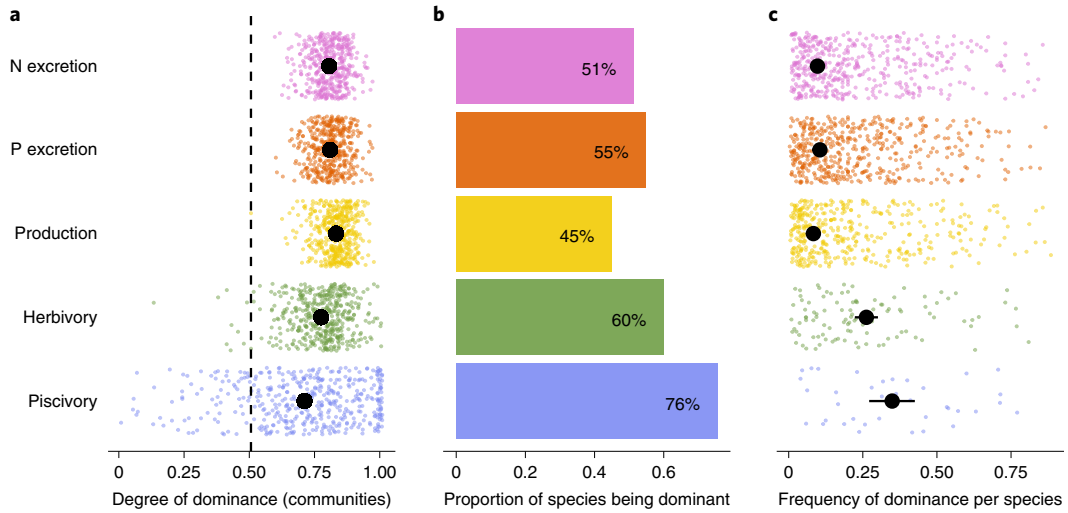


Fig. 4 | Local dominance in species contributions to five ecosystem functions on coral reefs. a, The degree of dominance for each function at the site level. The degree of dominance of a community ranges between zero (all species contribute equally to the function) and one (a single species is the sole contributor to a given function). Coloured dots represent the raw values, and the black dots and lines display the mean and 95% CIs of degree of dominance among all sites. In some cases, the CI is too small to be visible. The vertical dashed line shows the average degree of dominance of 1,000 randomly simulated communities. **b**, Bar plot of the proportion of species that are dominant in at least one site relative to the total number of species or, for herbivory and piscivory, the total number of herbivores and piscivores, respectively. **c**, Species-specific frequencies of dominance in each function across all sites, ranging from zero (species are never dominant) to one (dominant wherever present). The black dots and lines display the mean and 95% CIs of the frequency of dominance among species. A species is categorized as dominant in a community if its contribution to a function is higher than a scenario in which all species are equal (that is, one divided by the number of species that contribute to the function).

for the effect of biomass, these community shifts can enhance N excretion and production[15], but they will negatively impact P excretion, herbivory and piscivory. Furthermore, declines in coral cover related to climate change and warming seas are often associated with shifts towards herbivores[28,29]. Herbivores generally contribute little to P excretion[13,17], so a shift to herbivore dominance and the subsequent decline of community-level P excretion may change the balance of nutrient cycling on coral reefs, potentially favouring algal growth over corals[30].

Sustaining biomass, diversity and ecosystem functioning are central objectives of most conservation initiatives[8]. Although safeguarding fish biomass enhances functioning, the trade-offs among key functions reveal a critical challenge for coral reef conservation, where actions to enhance one function may negatively impact another. For example, the establishment of marine protected areas, which are one of the primary conservation strategies for coral reefs[31], may protect herbivorous species. However, marine protected areas do not protect reefs from the pervasive effects of climate change[31], and community shifts towards herbivore domination may result in the decline of P excretion. Thus, measuring conservation success with biomass or solely one function (for example, herbivory) can mask the collapse of other essential functions. It is necessary to gauge the state of reef ecosystems based on multiple, complementary, process-based functions. However, understanding of process-based functioning or the definition of a 'functional' coral reef are still lacking[7]. Establishing functional baselines for global coral reefs is a critical challenge for future studies. Until then, our results suggest that coral reef fish functions can be managed by enhancing standing stock biomass, protecting local key species and vulnerable constituents of the community (for example, large carnivores) and promoting regional biodiversity.

We demonstrate that the variability in processes that govern elemental cycling in complex ecosystems, such as tropical coral reefs, represents an unrecognized challenge for protecting ecosystem functioning. Management strategies that call for the enhancement of ecosystem functioning via an economic mindset (that is, where higher functioning is better) are not feasible. Instead, conserving coral reef ecosystem functioning will require a more nuanced approach that considers processes that vary beyond the effect of standing stock biomass and are subject to local trade-offs, drivers and anthropogenic threats.

Methods

Underwater visual census database. We used a published global database of reef fish abundances and sizes collected along belt transects[16]. This database encompasses 9,118 transects across 585 sites (within 98 localities) in the central Indo-Pacific, central Pacific, eastern Pacific, western Indian, eastern Atlantic and western Atlantic oceans. Sites are defined as small islands or stretches of continuous reefs in larger coastlines, and localities encompass sites that belong to the same biogeographic sub-provinces[16]. The database only includes transects on the outer reef slope and with a hard reef bottom. Transects were carried out at a constant depth, parallel to the reef crest. We discarded the species inside families for which we did not have body stoichiometry data, individuals that were smaller than 7 cm (to minimize the bias related to the identification of small individuals) and rare species for which fewer than 20 individuals were recorded across all transects. The dataset then included 1,110 species belonging to 25 families (Acanthuridae, Balistidae, Bothidae, Chaetodontidae, Cirrhitidae, Fistulariidae, Haemulidae, Holocentridae, Kyphosidae, Labridae, Lethrinidae, Lutjanidae, Monacanthidae, Mugilidae, Mullidae, Ostraciidae, Pempheridae, Pomacanthidae, Pomacentridae, Sciaenidae, Scorpaenidae, Serranidae, Siganidae, Tetraodontidae and Zanclidae). SST for each site was obtained from daily time-series data from the National Oceanic and Atmospheric Administration covering a five-year period (°C, 0.25° resolution)[32] (available from https://psl.noaa.gov/data/gridded/tables/sst.html). Furthermore, for each transect, we calculated species richness and estimated total standing stock biomass of fishes by using Bayesian length–weight relationships available from FishBase[33,34]. All data processing and analyses were performed in the software program R (version 4.0.2, R Core Team 2020).

Quantification of functions. For each transect, we estimated five key process-based functions mediated by fishes: N excretion rate ($gN\,m^{-2}\,d^{-1}$), P excretion rate ($gP\,m^{-2}\,d^{-1}$), production of biomass through growth ($gC\,m^{-2}\,d^{-1}$), herbivory (that is, ingestion rate of macrophytes ($gC\,m^{-2}\,d^{-1}$)) and piscivory (that is, ingestion rate of fishes ($gC\,m^{-2}\,d^{-1}$))[7]. These five functions were estimated for each transect using individual-based bioenergetic models predicting fluxes of C, N and P (for example, daily C intake rates, N and P excretion rates and growth rates)[17]. This bioenergetic model framework integrates elements of metabolic theory, stoichiometry and flexible elemental limitation[17]. We estimated the input parameters, including elements of metabolism, growth and diet and body

stoichiometry, for all 1,110 species through the integration of empirical data, data synthesis and extrapolation based on Bayesian phylogenetic models (Supplementary Methods). We then ran a unique bioenergetic model for each combination of species identity, body size and SST ($n = 30{,}668$) to obtain the contribution of each individual to each function in each transect. Finally, we summarized functions at the community level by summing up all individual contributions inside a transect and dividing the sum by the surface area. Each function is, thus, expressed as dry mass (of C, N or P) per day per square metre. We note that N excretion, P excretion and biomass production include contributions of all fishes, whereas herbivory and piscivory are carried out by a subset of the community, with respect to their trophic guild as defined by ref. [35]. To reduce the occurrence of misclassification of herbivores and piscivores, we categorized a species as herbivorous or piscivorous if it had both the highest probability to be classified in that trophic group and this probability was more than 0.5, based on the probability scores of trophic guilds presented by Parravicini et al.[35]. Furthermore, as a comparison, we estimated herbivory and piscivory rates using two alternative trophic guild classifications based on expert opinion[9,35]. Both the herbivory and piscivory rates match the expert opinion trophic guild classifications. Finally, we estimated multifunctionality—that is, one measure that combines all five functions by taking the geometric average of the five functions (normalized to a range between 0 and 100). We used the geometric mean because functions are dependent on each other and vary by several orders of magnitude.

Community structure variables. We quantified a set of variables that characterize fish community structure. These variables describe the size, age and trophic distribution of the community, as these may all affect functions[17]. Specifically, we calculated the 2.5%, 50% and 97.5% quantiles of the total length, immaturity and trophic level of all individuals per transect. We included the 2.5% and 97.5% quantiles to account for the spread of these traits within communities while avoiding the effect of outliers. The total length is based on visual estimations by divers. The immaturity is quantified using the following formula:

$$\text{Immaturity}_i = \kappa\,(l_\infty - l_i)\,,$$

where κ is the species-specific growth rate parameter, l_∞ is the species-specific asymptotic adult length and l_i is the total length of individual i. Essentially, this is the derivative of the von Bertalanffy growth model for a certain length, and the higher this value is, the younger the individual. Finally, trophic level was extracted from FishBase[34].

Multivariate regression models. We fitted three multivariate Bayesian models with all five functions to: (1) predict functions on the locality level to create a map of functions, (2) investigate the effects of biomass and SST as well as the correlations among functions independent of biomass and SST and (3) estimate the effects of the community structure on each function. For each model, functions were log-transformed to ensure the normal distribution of residuals and an allometric relationship with biomass, which is hypothesized by metabolic theory[36]. In the underwater visual transect database, 291 transects (3%) did not contain herbivores and 4,467 transects (49%) did not contain piscivores, yielding zeros for herbivory and piscivory, respectively. We considered that these absences of herbivores or piscivores are likely an underestimation of their actual abundance at the surveyed reef site, as all reefs typically host a few herbivores and piscivores (that is, they are likely false zeros). To avoid removing all transects with missing values for herbivory or piscivory ($n = 4{,}620$) from our database when running multivariate analyses, we imputed these zeros as missing values, and they were eventually set as parameters in the multivariate models.

First, we performed a multivariate intercept-only regression model with the five log-transformed functions to estimate the functions per locality. The model structure includes random effects for localities and sites:

$$\begin{bmatrix} y_{E_{\mathrm{N}},i} \\ y_{E_{\mathrm{P}},i} \\ y_{\mathrm{B},i} \\ y_{\mathrm{H},i} \\ y_{\mathrm{P},i} \end{bmatrix} \approx \text{MVNormal}\left(\begin{bmatrix} \mu_{E_{\mathrm{N}},i} \\ \mu_{E_{\mathrm{P}},i} \\ \mu_{\mathrm{B},i} \\ \mu_{\mathrm{H},i} \\ \mu_{\mathrm{P},i} \end{bmatrix}, S \right),$$

$$\mu_{E_{\mathrm{N}},i} = \left(\beta 0_{E_{\mathrm{N}}} + \delta_{E_{\mathrm{N,loc}}} + \delta_{E_{\mathrm{N,site}}}\right)$$
$$\mu_{E_{\mathrm{P}},i} = \left(\beta 0_{E_{\mathrm{P}}} + \delta_{E_{\mathrm{P,loc}}} + \delta_{E_{\mathrm{P,site}}}\right)$$
$$\mu_{\mathrm{B},i} = \left(\beta 0_{\mathrm{B}} + \delta_{\mathrm{B,loc}} + \delta_{\mathrm{B,site}}\right)$$
$$\mu_{\mathrm{H},i} = \left(\beta 0_{\mathrm{H}} + \delta_{\mathrm{H,loc}} + \delta_{\mathrm{H,site}}\right)$$
$$\mu_{\mathrm{P},i} = \left(\beta 0_{\mathrm{P}} + \delta_{\mathrm{P,loc}} + \delta_{\mathrm{P,site}}\right),$$

$$S = \begin{bmatrix} \sigma_{E_{\mathrm{N}}} & 0 & 0 & 0 & 0 \\ 0 & \sigma_{E_{\mathrm{P}}} & 0 & 0 & 0 \\ 0 & 0 & \sigma_{\mathrm{B},i} & 0 & 0 \\ 0 & 0 & 0 & \sigma_{\mathrm{H}} & 0 \\ 0 & 0 & 0 & 0 & \sigma_{\mathrm{P}} \end{bmatrix} R \begin{bmatrix} \sigma_{E_{\mathrm{N}}} & 0 & 0 & 0 & 0 \\ 0 & \sigma_{E_{\mathrm{P}}} & 0 & 0 & 0 \\ 0 & 0 & \sigma_{\mathrm{B},i} & 0 & 0 \\ 0 & 0 & 0 & \sigma_{\mathrm{H}} & 0 \\ 0 & 0 & 0 & 0 & \sigma_{\mathrm{P}} \end{bmatrix},$$

where i is the index of the transect, μ represents the average prediction per function, $y_{E_{\mathrm{N}},i}$ is the N excretion rate of transect i, $y_{E_{\mathrm{P}},i}$ is the P excretion rate, $y_{\mathrm{B},i}$ is the biomass production rate, $y_{\mathrm{H},i}$ is the herbivory rate, $y_{E_{\mathrm{N}},i}$ is the piscivory rate, σ represents the residual error of each function (E_{N}, E_{P}, B, H and P), R is the correlation matrix of the residuals, $\beta 0$ is the intercept for each function, and $\delta_{\text{function,loc}}$ and $\delta_{\text{function,site}}$ represent the random effects of locality and sites, respectively. Locality-level and site-level effects are also structured including covariation among functions. There are, thus, three correlation matrices in total, meaning that the model will estimate the correlation among functions on three levels: locality, site and transect.

We used non-centred parameterization for site and location effects and all standard deviations had the following prior: $\sigma \approx$ student (3, 0, 2.5). We used a prior (lkj_{corr}, where lkj is Lewandowski–Kurowicka–Joe) for each of the three correlation matrices ($R \approx lkj_{\text{corr}}$ (1)).

Second, we ran a mixed-effect model to investigate the effects of biomass and SST on all functions and the correlations among functions (independent of biomass and SST). The standing stock biomass of communities is positively related to all functions because of the additive nature of the quantification and metabolic theory[36]. Furthermore, because of the known relationship between temperature and parameters related to growth and respiration (Supplementary Methods), functions are expected to be affected by temperature. We thus fitted a multivariate Bayesian mixed-effect model using transect-level, log-transformed functions that included random effects for sites and localities:

$$\begin{bmatrix} y_{E_{\mathrm{N}},i} \\ y_{E_{\mathrm{P}},i} \\ y_{\mathrm{B},i} \\ y_{\mathrm{H},i} \\ y_{\mathrm{P},i} \end{bmatrix} \approx \text{MVNormal}\left(\begin{bmatrix} \mu_{E_{\mathrm{N}},i} \\ \mu_{E_{\mathrm{P}},i} \\ \mu_{\mathrm{B},i} \\ \mu_{\mathrm{H},i} \\ \mu_{\mathrm{P},i} \end{bmatrix}, S \right),$$

$$S = \begin{bmatrix} \sigma_{E_{\mathrm{N}}} & 0 & 0 & 0 & 0 \\ 0 & \sigma_{E_{\mathrm{P}}} & 0 & 0 & 0 \\ 0 & 0 & \sigma_{\mathrm{B},i} & 0 & 0 \\ 0 & 0 & 0 & \sigma_{\mathrm{H}} & 0 \\ 0 & 0 & 0 & 0 & \sigma_{\mathrm{P}} \end{bmatrix} R \begin{bmatrix} \sigma_{E_{\mathrm{N}}} & 0 & 0 & 0 & 0 \\ 0 & \sigma_{E_{\mathrm{P}}} & 0 & 0 & 0 \\ 0 & 0 & \sigma_{\mathrm{B},i} & 0 & 0 \\ 0 & 0 & 0 & \sigma_{\mathrm{H}} & 0 \\ 0 & 0 & 0 & 0 & \sigma_{\mathrm{P}} \end{bmatrix}$$

$$\mu_{E_{\mathrm{N}},i} = \left(\beta 0_{E_{\mathrm{N}}} + \delta_{E_{\mathrm{N,loc}}} + \delta_{E_{\mathrm{N,site}}}\right) + \beta 1_{E_{\mathrm{N}}} \log\left(\text{biomass}\right), i + \beta 2_{E_{\mathrm{N}}} \text{SST}, i$$
$$\mu_{E_{\mathrm{P}},i} = \left(\beta 0_{E_{\mathrm{P}}} + \delta_{E_{\mathrm{P,loc}}} + \delta_{E_{\mathrm{P,site}}}\right) + \beta 1_{E_{\mathrm{P}}} \log\left(\text{biomass}\right), i + \beta 2_{E_{\mathrm{P}}} \text{SST}, i$$
$$\mu_{\mathrm{B},i} = \left(\beta 0_{\mathrm{B}} + \delta_{\mathrm{B,loc}} + \delta_{\mathrm{B,site}}\right) + \beta 1_{\mathrm{B}} \log\left(\text{biomass}\right), i + \beta 2_{\mathrm{B}} \text{SST}, i$$
$$\mu_{\mathrm{H},i} = \left(\beta 0_{\mathrm{H}} + \delta_{\mathrm{H,loc}} + \delta_{\mathrm{H,site}}\right) + \beta 1_{\mathrm{H}} \log\left(\text{biomass}\right), i + \beta 2_{\mathrm{H}} \text{SST}, i$$
$$\mu_{\mathrm{P},i} = \left(\beta 0_{\mathrm{P}} + \delta_{\mathrm{P,loc}} + \delta_{\mathrm{P,site}}\right) + \beta 1_{\mathrm{P}} \log\left(\text{biomass}\right), i + \beta 2_{\mathrm{P}} \text{SST}, i$$

where $\beta 1_{E_{\mathrm{N}}}$, $\beta 1_{E_{\mathrm{P}}}$, $\beta 1_{\mathrm{B}}$, $\beta 1_{\mathrm{H}}$, $\beta 1_{\mathrm{P}}$ are the fixed effects of the log-transformed biomass and $\beta 2_{E_{\mathrm{N}}}$, $\beta 2_{E_{\mathrm{P}}}$, $\beta 2_{\mathrm{B}}$, $\beta 2_{\mathrm{H}}$, $\beta 2_{\mathrm{P}}$ are the fixed effects of SST. Locality-level and site-level effects are, thus, structured including covariation among functions, independent of biomass and SST. Similarly, the residual variation of functions incorporates the correlations between functions, without the effect of biomass and SST. We used similar priors as described above, and we used weakly informative normal priors for the model slopes ($\beta 1 \approx$ normal (1, 1), $\beta 2 \approx$ normal (0, 1)).

Finally, to investigate the effect of community structure while still accounting for the effects of standing biomass and SST, we fitted a mixed-effect multivariate model similar to the model specified above, but we added all community structure variables:

$$\begin{aligned} \mu_{\text{function,i}} = {} & \beta 0_{\text{function}} + \beta 1_{\text{function}} \log\left(\text{biomass}\right), i + \beta 2_{\text{function}} \text{SST}, i \\ & + \beta 3_{\text{function}} \text{richness}, i + \beta 4_{\text{function}} \text{size}_m, i + \beta 5_{\text{function}} \text{size}_{2.5\%}, i \\ & + \beta 6_{\text{function}} \text{size}_{97.5\%}, i + \beta 7_{\text{function}} \text{troph}_m, i + \beta 8_{\text{function}} \text{troph}_{2.5\%}, i \\ & + \beta 9_{\text{function}} \text{troph}_{97.5\%}, i + \beta 10_{\text{function}} \text{Im}_m, i + \beta 11_{\text{function}} \text{Im}_{2.5\%}, i \\ & + \beta 12_{\text{function}} \text{Im}_{97.5\%}, i \end{aligned}$$

where richness is the species richness, size is the total length, troph is the trophic level, Im is the immaturity and m, 2.5% and 97.5% represent the 50%, 2.5% and 97.5% quantiles across the fish community, respectively. For these models, we used weakly informative priors for the fixed effect parameters ($\beta 3 - \beta 12 \approx$ normal (0, 1)) and the same priors as described above for other parameters.

All Bayesian models were fitted using the R package brms[37], which uses Stan, a C++ package to perform full Bayesian inference[38]. The posterior distributions of model parameters were estimated using Hamiltonian Monte Carlo methods by using four chains of 2,000 samples, including 1,000 samples as a warm-up. Thus, a total of 4,000 draws were used to estimate posterior distributions. The convergence and fit of the models were verified by examining the Rhat, parameter trace plots and posterior prediction plots (Extended Data Fig. 2).

Species dominance and contributions to functions. We estimated the relative contribution of each species to each function for all sites as follows:

$$\text{Contribution}_{f,i,j} = \frac{F_{f,i,j}}{\sum_{f,j}^{F}},$$

where i is a certain species, j is a site and F is the value of function f.

Then, we quantified the degree of species dominance per function for each site. We first ranked species according to their contribution to function, and then we quantified the cumulative contributions of species to functions. Finally, we used the area under the species accumulation curve as a measure for the degree of dominance (DD). Specifically, the DD for a function performed by R species was calculated as follows:

$$\text{DD} = \frac{A - A_{\min}}{A_{\max} - A_{\min}},$$

where A is the area under the curve, $A_{\min}$ is the theoretical area under the curve where each species has an equal contribution to a certain function and $A_{\max}$ is the theoretical area under the curve where one species performs the entire function. They are quantified as:

$$A_{\min} = \frac{R^2 - 1}{2R},$$

$$A_{\max} = R - 1,$$

$$A = \sum_{i=2}^{R} \frac{C_i + C_{i-1}}{2},$$

where C_i is the contribution of a certain species and R equals the species richness in the case of N excretion, P excretion and production. For herbivory and piscivory, R represents the number of herbivores and piscivores, respectively. The DD, thus, ranges between 0 and 1, where 0 means that each species contributes equally and 1 means that a single species performs the entire function.

Finally, we quantified the frequency of dominance per species (that is, the number of sites in which a species is dominant for a given function divided by the total number of sites in which that species is observed). A species is considered dominant for a certain function in a given site if its contribution is higher than $1/R$ (that is, they contribute more than the situation in which each species contributes equally to a certain function).

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

All data needed to reproduce the figures are available on GitHub (https://github.com/nschiett/global_proc) and figshare (https://doi.org/10.6084/m9.figshare.13285901.v1). All empirical data that were used to estimate parameters for bioenergetic modelling (Supplementary Information) will be available on figshare (https://doi.org/10.6084/m9.figshare.19134446.v1) after a two-year embargo.

Code availability

All code to reproduce the figures are available on GitHub (https://github.com/nschiett/global_proc) and figshare (https://doi.org/10.6084/m9.figshare.13285901.v1).

Received: 17 February 2021; Accepted: 20 February 2022;
Published online: 04 April 2022

References

1. Welti, N. et al. Bridging food webs, ecosystem metabolism, and biogeochemistry using ecological stoichiometry theory. *Front. Microbiol.* **8**, 1298 (2017).
2. Ceballos, G. et al. Accelerated modern human-induced species losses: entering the sixth mass extinction. *Sci. Adv.* **1**, e14002 (2015).
3. Hughes, T. P. et al. Global warming and recurrent mass bleaching of corals. *Nature* **543**, 373–377 (2017).
4. Pauly, D. et al. Towards sustainability in world fisheries. *Nature* **418**, 689–695 (2002).
5. Bellwood, D. R., Streit, R. P., Brandl, S. J. & Tebbett, S. B. The meaning of the term 'function' in ecology: a coral reef perspective. *Funct. Ecol.* **33**, 948–961 (2019).
6. Williams, G. J. et al. Coral reef ecology in the Anthropocene. *Funct. Ecol.* **33**, 1014–1022 (2019).
7. Brandl, S. J. et al. Coral reef ecosystem functioning: eight core processes and the role of biodiversity. *Front. Ecol. Environ.* **17**, 445–454 (2019).
8. Cinner, J. E. et al. Meeting fisheries, ecosystem function, and biodiversity goals in a human-dominated world. *Science* **368**, 307–311 (2020).
9. Mouillot, D. et al. Functional over-redundancy and high functional vulnerability in global fish faunas on tropical reefs. *Proc. Natl Acad. Sci. USA* **111**, 13757–13762 (2014).
10. Mora, C. et al. Global human footprint on the linkage between biodiversity and ecosystem functioning in reef fishes. *PLoS Biol.* **9**, e1000606 (2011).
11. Barneche, D. R. et al. Scaling metabolism from individuals to reef-fish communities at broad spatial scales. *Ecol. Lett.* **17**, 1067–1076 (2014).
12. McIntyre, P. B. et al. Fish distributions and nutrient cycling in streams: can fish create biogeochemical hotspots? *Ecology* **89**, 2335–2346 (2008).
13. Allgeier, J. E., Layman, C. A., Mumby, P. J. & Rosemond, A. D. Consistent nutrient storage and supply mediated by diverse fish communities in coral reef ecosystems. *Glob. Change Biol.* **20**, 2459–2472 (2014).
14. Morais, R. A. & Bellwood, D. R. Pelagic subsidies underpin fish productivity on a degraded coral reef. *Curr. Biol.* **29**, 1521–1527.e6 (2019).
15. Morais, R. A., Connolly, S. R. & Bellwood, D. R. Human exploitation shapes productivity–biomass relationships on coral reefs. *Glob. Change Biol.* **26**, 1295–1305 (2020).
16. Barneche, D. R. et al. Body size, reef area and temperature predict global reef-fish species richness across spatial scales. *Glob. Ecol. Biogeogr.* **28**, 315–327 (2019).
17. Schiettekatte, N. M. D. et al. Nutrient limitation, bioenergetics and stoichiometry: a new model to predict elemental fluxes mediated by fishes. *Funct. Ecol.* **34**, 1857–1869 (2020).
18. Schramski, J. R., Dell, A. I., Grady, J. M., Sibly, R. M. & Brown, J. H. Metabolic theory predicts whole-ecosystem properties. *Proc. Natl Acad. Sci. USA* **112**, 2617–2622 (2015).
19. Morais, R. A. & Bellwood, D. R. Global drivers of reef fish growth. *Fish Fish.* **19**, 874–889 (2018).
20. Hood, J. M., Vanni, M. J. & Flecker, A. S. Nutrient recycling by two phosphorus-rich grazing catfish: the potential for phosphorus-limitation of fish growth. *Oecologia* **146**, 247–257 (2005).
21. Barneche, D. R. & Allen, A. P. The energetics of fish growth and how it constrains food-web trophic structure. *Ecol. Lett.* **21**, 836–844 (2018).
22. Brandl, S. J. et al. Demographic dynamics of the smallest marine vertebrates fuel coral reef ecosystem functioning. *Science* **364**, 1189–1192 (2019).
23. Lefcheck, J. S. et al. Tropical fish diversity enhances coral reef functioning across multiple scales. *Sci. Adv.* **5**, eaav6420 (2019).
24. Topor, Z. M., Rasher, D. B., Duffy, J. E. & Brandl, S. J. Marine protected areas enhance coral reef functioning by promoting fish biodiversity. *Conserv. Lett.* **12**, e12638 (2019).
25. Bellwood, D. R., Hughes, T. P. & Hoey, A. S. Sleeping functional group drives coral-reef recovery. *Curr. Biol.* **16**, 2434–2439 (2006).
26. Darling, E. S. & D'agata, S. Coral reefs: fishing for sustainability. *Curr. Biol.* **27**, R65–R68 (2017).
27. Graham, N. A. J. et al. Human disruption of coral reef trophic structure. *Curr. Biol.* **27**, 231–236 (2017).
28. Graham, N. A. J. et al. Dynamic fragility of oceanic coral reef ecosystems. *Proc. Natl Acad. Sci. USA* **103**, 8425–8429 (2006).
29. Stuart-Smith, R. D., Brown, C. J., Ceccarelli, D. M. & Edgar, G. J. Ecosystem restructuring along the great barrier reef following mass coral bleaching. *Nature* **560**, 92–96 (2018).
30. Burkepile, D. E. et al. Nutrient supply from fishes facilitates macroalgae and suppresses corals in a Caribbean coral reef ecosystem. *Sci. Rep.* **3**, 1493 (2013).
31. Graham, N. A. J. et al. Changing role of coral reef marine reserves in a warming climate. *Nat. Commun.* **11**, 2000 (2020).
32. Reynolds, R. W. et al. Daily high-resolution-blended analyses for sea surface temperature. *J. Clim.* **20**, 5473–5496 (2007).
33. Froese, R., Thorson, J. T. & Reyes, R. B. A Bayesian approach for estimating length–weight relationships in fishes. *J. Appl. Ichthyol.* **30**, 78–85 (2014).
34. Froese, R. & Pauly, D. *FishBase* (2018); https://www.fishbase.in/home.htm
35. Parravicini, V. et al. Delineating reef fish trophic guilds with global gut content data synthesis and phylogeny. *PLoS Biol.* **18**, e3000702 (2020).

36. Brown, J. H., Gillooly, J. F., Allen, A. P., Savage, V. M. & West, G. B. Toward a metabolic theory of ecology. *Ecology* **85**, 1771–1789 (2004).
37. Bürkner, P.-C. brms: an R package for Bayesian multilevel models using Stan. *J. Stat. Softw.* **80**, 1–28 (2017).
38. Carpenter, B. et al. Stan: a probabilistic programming language. *J. Stat. Softw.* **76**, 1–31 (2017).

Acknowledgements

We thank the staff at CRIOBE, Moorea for field support. We also thank J. Carlot, S. Degregori, B. French, T. Roncin, Y. Lacube, C. Gache, G. Martineau, K. Bissell, B. Espiau, C. Quigley, K. Landfield and T. Norin for their help in the field, G. de Sinéty and J. Wicquart for their contribution to otolith analysis, and S. Schiettekatte for proofreading the manuscript. This research was funded by the BNP Paribas Foundation (Reef Services Project) and the French National Agency for Scientific Research (ANR, REEFLUX Project, ANR-17-CE32-0006). This research is the product of the SCORE-REEF group funded by the Centre de Synthèse et d'Analyse sur la Biodiversité of the Foundation pour la Recherche sur la Biodiversité and the Office Francais de la Biodiversité. V.P. was supported by the Institut Universitaire de France, and J.M.C. was supported by a Make Our Planet Great Again Postdoctoral Grant (mopga-pdf-0000000144).

Author contributions

N.M.D.S. and V.P. conceived the idea. N.M.D.S., V.P., S.J.B., J.M.C. and S.V. designed the methodology. N.M.D.S., J.M.C., S.J.B., A.M., F.M., V.P., K.S.M., J.E.A. and D.E.B. collected the data. N.A.J.G., D.R.B., D.E.B., J.E.A., J.E.A.-G., G.J.E., C.E.L.F., S.R.F., A.M.F., A.L.G., M.K., Y.L., O.J.L., F.M., E.L.R., F.A.R.-Z., R.D.S.-S. and L.V. shared existing data. N.M.D.S. analysed the data and led the writing of the manuscript. All authors contributed significantly to the drafts and approved the final version for publication.

Competing interests

The authors declare no competing interests.

Additional information

Extended data is available for this paper at https://doi.org/10.1038/s41559-022-01710-5.

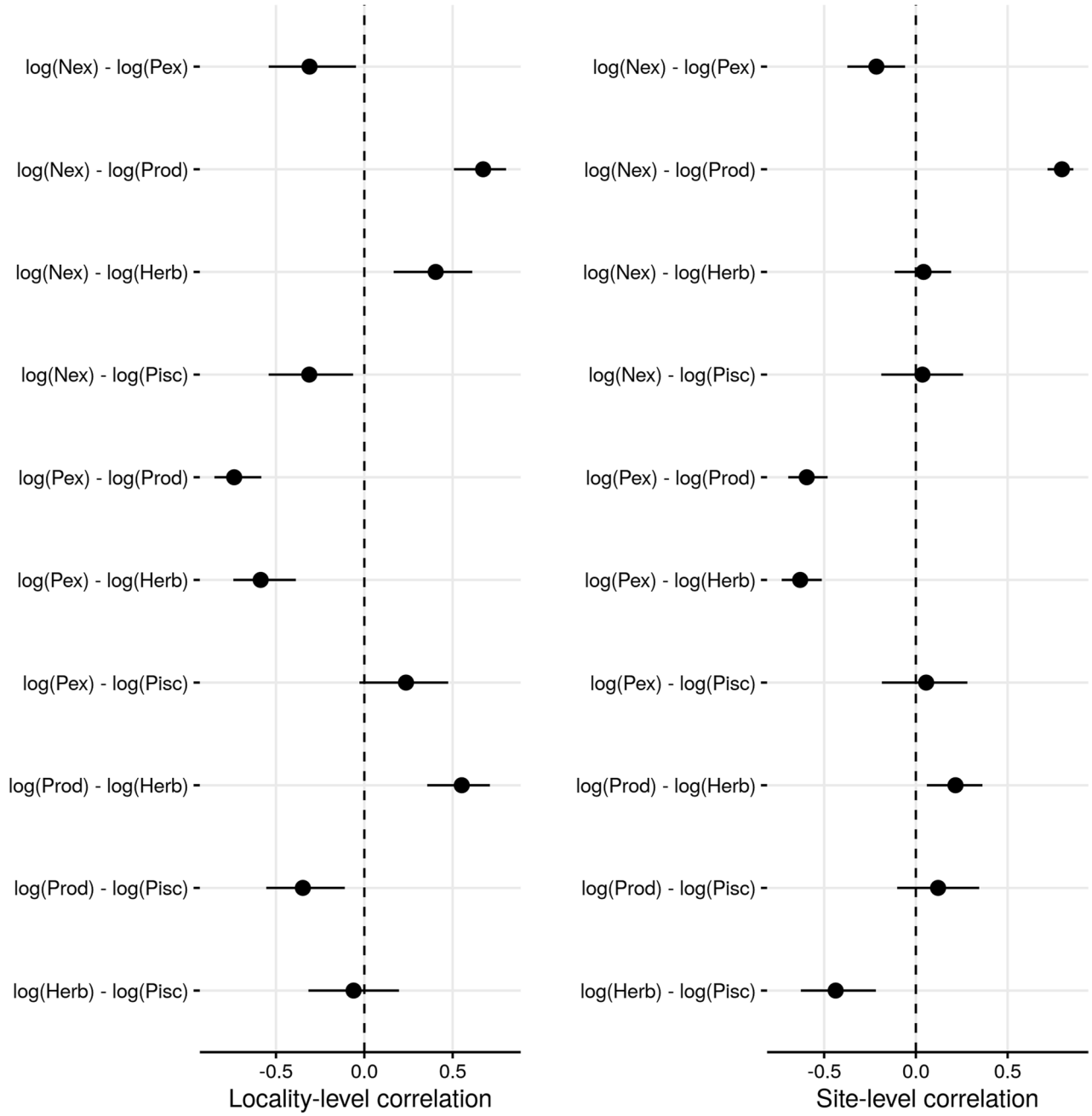
Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41559-022-01710-5.

Correspondence and requests for materials should be addressed to Nina M. D. Schiettekatte.

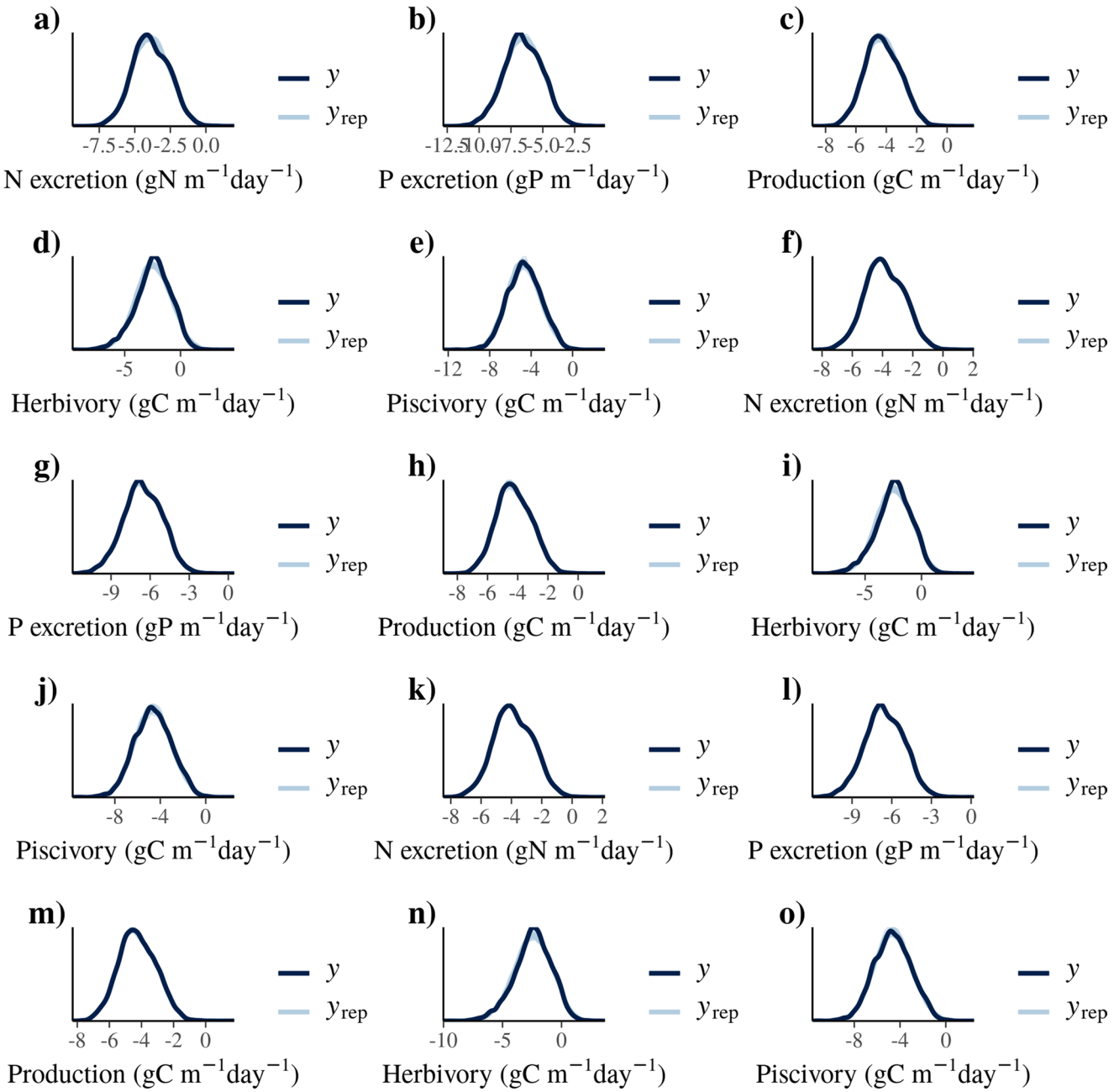
Peer Review File *Nature Ecology & Evolution* thanks Matthew McLean, Marc Hensel and Gareth Williams for their contribution to the peer review of this work. Peer reviewer reports are available.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Extended Data Fig. 1 | Correlations among functions. Correlations, independent of biomass and sea surface temperature, at the locality and site levels. Dotes and lines indicate the mean estimated values and 95% credible intervals, respectively.



Extended Data Fig. 2 | Posterior predictive checks of multivariate models. a-e: Intercept-only model, **f-j**: model with biomass and sea surface temperature, **k-o**: model with all community variables.

nature research

Corresponding author(s): Nina M. D. Schiettekatte

Last updated by author(s): Feb 7, 2022

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☒	☐	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☒	☐	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒	A description of all covariates tested
☒	☐	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☒	☐	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☐	☒	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	No data were collected using computer code or software.
Data analysis	All data and code to reproduce analyses and figures are available online through GitHub (https://github.com/nschiett/global_proc) and figshare (https://doi.org/10.6084/m9.figshare.13285901.v1, https://doi.org/10.6084/m9.figshare.19134446.v1).

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

All data and code to reproduce analyses and figures are available online through GitHub (https://github.com/nschiett/global_proc) or figshare (https://doi.org/10.6084/m9.figshare.13285901.v1). All empirical data used to estimate parameters for bioenergetic modeling are available on figshare (https://doi.org/10.6084/m9.figshare.19134446.v1) after an embargo period of two years.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☐ Life sciences ☐ Behavioural & social sciences ☒ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Ecological, evolutionary & environmental sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description	We used a combination of fish community data and bioenergetic modelling to quantify 5 key ecological functions (i.e. herbivory, piscivory, nitrogen cycling, phosphorous cycling and biomass production) for global coral reefs. Our study encompasses 9118 transect replicates across 585 sites (98 localities) in the Central Indo-Pacific, Central Pacific, Eastern Pacific, Western Indian, Eastern Atlantic, Western Atlantic. Our surveys are referred to 1110 reef fish species. For each transect we quantified the 5 functions. Then we used Bayesian modeling to identify major predictors of the ecological functions, to test for the existence of thresholds among functions and to assess the contribution of each species to the five processes.
Research sample	This is a study conducted at community level. The sampling unit we used to assess community data is the underwater visual census transect. This is the standard technique for fish community in coastal marine ecosystems. This dataset has been already successfully used in the literature and comparable datasets have been already used in several articles published by renowned journals, including Nature.
Sampling strategy	Our community dataset encompasses all major biogeographical regions of the world. Our main goal was to evaluate ecosystem functioning across the largest range of conditions possible. We are confident that sample size is appropriate for several reasons: 1) Our dataset is one of the most comprehensive datasets available for coral reefs; 2) Our experimental design does not include experimental treatments that are extremely sensitive to sample size. Our goal is to quantify ecosystem functioning for global coral reefs and our conclusions are robust to changes in sample size; 3) In order to quantify ecosystem functioning we used bioenergetic models implemented in a Bayesian framework, which quantitatively accounts for uncertainty.
Data collection	We collected data available from the literature and empirical data available to co-authors to inform bioenergetic models. Bioenergetic models relies on a set of input parameters including elements of metabolism, growth, and diet and body stoichiometry. We quantified the input parameters for all species through the integration of empirical data, data synthesis, and Bayesian phylogenetic models. Our work includes empirical and literature data on Carbon, Nitrogen and Phosphorous content CNP content in fish body (data referred to 1633 individuals belonging to 108 species – available from NMDS, DEB, JEA, VP), metabolic rate (data referred to 1393 individuals belonging to 61 species – available from NMDS, SJB, VP), CNP content in diet (data referred to 571 individuals belonging to 51 species – available from NMDS, JMC, VP), growth curves based on otolith readings (data referred to 710 individuals belonging to 45 species – available from NMDS, FM, VP).
Timing and spatial scale	Our community data are referred to the period between 1998 and 2002 and are collected at a global scale.
Data exclusions	We removed certain species or individuals from the underwater visual census database to reduce bias. We selected the species inside families for which we have body stoichiometric data, that were at least 7cm to minimize the bias related to the identification of small individuals, and finally we discarded rare species, for which less than 20 individuals were ever recorded across all transects.
Reproducibility	Our study does not consist in an experiment that can or not be reproduced. All the data and code needed to reproduce our study are made available.
Randomization	Our study design does not consist in an experiment that requires randomization test. Our study is based on data on fish traits that are used to estimate community level functions. We then used a correlative approach to explore the relative importance of correlates.
Blinding	Our study design does not consist in an experiment that requires blinding during data acquisition. Our study is based on data on fish traits that are used to estimate community level functions. We then used a correlative approach to explore the relative importance of correlates.

Did the study involve field work? ☒ Yes ☐ No

Field work, collection and transport

Field conditions	Water temperature ranged between 27 and 30 degrees Celsius for all sample collection and in the case of collections at the outer reef, the wave height did not exceed 1 meter.
Location	Our study uses a combination of published and collected data. All novel sampling was done in Moorea, French Polynesia (-17.53, -149.83). Published data used came from tropical coral reefs across the world and all exact locations are accessible in the data repository.
Access & import/export	Access, extraction, and transport of samples were approved by the government of French Polynesia.

Disturbance	All data collected in the field had minor disturbance through boat noise and diver presence. Disturbance was minimized by respecting the habitat and only affecting targeted individuals.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☐	☒ Animals and other organisms
☒	☐ Human research participants
☒	☐ Clinical data
☒	☐ Dual use research of concern

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Animals and other organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research

Laboratory animals	The study did not involve laboratory animals.
Wild animals	For this study, fishes were observed in their natural environment, collected alive for respirometry trials, and gut contents, and perform body CNP analysis. Fishes were observed to record abundance and size during underwater visual census through scuba diving, a widely applied and standard technique which has minimal disturbance on the environment. All fishes were collected selectively through spear fishing, nets, and clove oil, making while making sure to only affect the targeted individual, and minimizing impact on the environment. Fishes targeted for respirometry were immediately placed in a aerated container with seawater for transport to the laboratory, and were then kept in tanks with a continuous flow of filtered seawater at a constant temperature of 28 degrees Celsius. Fishes targeted for extraction of otoliths, gut content or full body CNP analysis were pitted immediately after capture and placed in a cooler with ice, thus minimizing animal suffering.
Field-collected samples	Field collected samples were freeze-dried stored at room temperature for transport to France and the United States for further analysis.
Ethics oversight	All protocols related to the capture and handling of fish complied to the ethical standards of CRIOBE and EPHE, and the University of California Santa Barbara's Institutional Animal Care and Use Committee (IACUC #915 2016-2019). Extraction and transport of samples were approved by the government of French Polynesia.

Note that full information on the approval of the study protocol must also be provided in the manuscript.