Biological trade-offs underlie coral reef ecosystem functioning

Preserving coral reef functioning is a critical challenge of the 21st century. However, a lack of quantitative assessments of multiple functions across large spatial scales has hindered local and regional conservation efforts. We integrate empirically-parameterized bioenergetic models and global community surveys to quantify five key functions associated with nutrient cycling and energy flux 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. Further, functions are locally dominated by few species, but the identity of dominant species substantially varies at the global scale. In fact, roughly half of the 1,110 species in our dataset are functionally dominant somewhere on a local scale. 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[1]. Humans increasingly threaten biodiversity and ecosystem functioning[2]. Coral reefs are a prime example of an ecosystem severely impacted by anthropogenic activities. 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 conservation of coral reefs[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 the fluxes of elements can scale non-linearly 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]. While 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]. Further, studies that have quantified functions as a flow of matter mostly focused on single functions (e.g. biomass production[14,15] or fish excretion[13]) and covered only a small number of species at local scales. Consequently, trade-offs among multiple functions, their drivers, and their vulnerability to anthropogenic stressors remain poorly understood in coral reef ecosystems across large spatial scales[7].

Here, we integrate biogeochemistry and community ecology to advance our understanding of the elemental fluxes that underpin reef fish functioning. Using empirical species-specific data on basic organismal processes and Bayesian phylogenetic models, we parameterize individual-level bioenergetic models to estimate five key ecosystem functions: N excretion, P excretion, biomass production, herbivory, and piscivory for 1,100 species. We apply these bioenergetic models to 9,118 reef fish transects across 585 sites worldwide (Supplementary Table 1) to: (1) quantify community-level reef fish functions and their trade-offs, (2) extract the community- and species-level effects on these functions, and (3) gauge the vulnerability of reef fish functioning in the Anthropocene.

Results

We quantified five key ecosystem functions mediated by coral reef fishes across the globe (Fig. 1). Functions across localities show a similar geographical pattern. However, at the global scale we could not find a single location with high levels of functioning across all functions. A measure of multifunctionality thus does not appropriately represent the state of all functions assessed independently.

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 the log-transformed biomass were 0.932 (95%CI: 0.929, 0.934) for N excretion, 1.051 (1.047, 1.056) for P excretion, 0.771 (0.764, 0.780) for production, 0.940 (0.923, 0.957) for herbivory, and 0.668 (0.635, 0.702) for piscivory. These slopes indicate that the relationships between biomass and functions are all non-linear, which demonstrates that biomass is not an appropriate proxy for function. We also incorporated sea surface temperature (SST) because it affects the metabolic and growth rates of fishes, which scales up to the community[16]. We found positive effects of SST on N excretion, production, and herbivory, and no effects of SST on P excretion and piscivory (Supplementary Table 3).

Our multivariate model allowed us to estimate the correlations between functions, independent of the effects of biomass and SST. In particular, we estimated correlations between functions on three levels: the locality effects, site effects, and residual variations (Fig. 2, Fig. S1). The correlations displayed comparable patterns on each level. We found strong negative trade-offs between P excretion and N excretion as well as P excretion and biomass production. Further, we found slightly weaker negative correlations between piscivory and N excretion as well as piscivory and herbivory. Thus, a reef fish community does not simultaneously display high values of functioning across all investigated functions.

To explore what drives the variation and trade-offs of functions beyond the effects of biomass and SST, we ran a multivariate Bayesian mixed effects model by adding ten variables that describe the structure of each fish assemblage: species richness and the median, lower, and upper 95% quantiles of size, immaturity, and trophic level of individuals inside a community. Each of these components have 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, Supplementary Table 4). Some associations such as the negative and positive effects of trophic level on herbivory and piscivory, respectively, are expected from first principles, whereas others, such as the positive effect of immaturity on biomass production and negative 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 dominance inside each community at the site-level and found that, on average, functions are dominated by a small fraction of the species in each community (Fig. 4a). We also calculated the proportion of species that is dominant in at least one transect (i.e. species with a disproportionally high contribution as compared to a community in which all species contribute equally), and many species play a dominant role somewhere in the surveyed locations. In particular, 49% of all species contributed disproportionally to a function in at least one surveyed community (Fig. 4b). Yet, very few species are dominant throughout their range (Fig. 4c). Thus, functions within communities tend to be driven by few dominant species, and 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 pervasive influence of biomass on all other processes. Yet, our results also show non-linear relationships between functions and biomass and suggest that biomass alone does not sufficiently characterize functioning; strong trade-offs occurred among the five functions independent of biomass. Using biomass as a proxy might, therefore, mask fundamental differences in community-level functioning. Furthermore, for a given value of biomass, no reef can yield above average values across all five functions. While a reef may stand out as a hotspot for one function, no reef can maximize all functions simultaneously.

The observed trade-offs among functions are driven by fish community structure and the organismal physiology and life-history traits of its consituents[17,18]. For example, we observed a clear trade-off between P excretion and biomass production which is mostly driven by the age and trophic structure (Fig. 3). Communities dominated by fishes with high trophic levels are charachterized 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] and less energy is lost with each step up the food chain ('trophic transfer efficiency'; ref). Moreover, P is retained for skeletal growth in young fishes, thus limiting P excretion rates[17,20]. Furthermore, 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 also 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 revealed that no species dominated throughout their geographical range, and more than half of all species contributed disproportionally to a specific function at a at least one site. Thus, there are no widespread key species to target for coral reef conservation aimed at preserving these five functions; rather, local species dominance across functions can guide local conservation efforts, and the preservation of regional reef fish biodiversity should be prioritized in broad-scale policy.

Our global analysis of multiple functions suggests pathways in which human-induced shifts in reef fish community structure may impact coral reef ecosystems (Fig. 5). 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 for the effect of biomass, these community shifts can enhance N excretion and production (e.g. [15]) but they will negatively impact P excretion, herbivory, and piscivory. Further, declines in coral cover related to climate change, and warming seas at higher latitude coral reefs are often associated with shifts toward herbivores[28]. 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 favoring algal growth over corals[29].

Sustaining biomass, diversity, and ecosystem functioning are important objectives of most conservation initiatives[8]. While safeguarding fish biomass enhances functioning, the trade-offs between 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[30], may protect herbivorous species. However, marine protected areas do not protect reefs from the pervasive effects of climate change[30], and community shifts towards herbivore domination may result in the decline of P excretion. Thus, measuring conservation success with biomass or solely one function (e.g. 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. Yet, our comprehension of process-based functioning or the definition a “functional” coral reef is still poorly understood[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 (e.g. large carnivores), and promoting regional biodiversity.

We demonstrate that the variability in processes that govern the 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 (i.e. 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

1. 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 7cm (to minimize the bias related to the identification of small individuals), and rare species for which less 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, Zanclidae). Sea surface temperature (SST) for each site was obtained from daily time-series data from the National Oceanicand Atmospheric Administration (NOAA) covering a 5-year period (°C; 0.25° resolution) (available from https://www.esrl.noaa.gov/psd/data/gridded/data.noaa.oiSST.v2.highres.html)[31];. Further, for each transect, we calculated species richness and estimated total standing stock biomass of fishes by using Bayesian length-weight relationships available from Fishbase[32]. All data processing and analyses were performed in the software program R (version 4.0.2; R core team 2020).

2. Quantification of functions

For each transect, we estimated five key process-based functions mediated by fishes: nitrogen excretion rate (gN m^{-2} day^{-1}), phosphorus excretion rate (gP m^{-2} day^{-1}), production of biomass through growth (gC m^{-2} day^{-1}), herbivory, (i.e. ingestion rate of macrophytes (gC m^{-2} day^{-1})), and piscivory (i.e. ingestion rate of fishes (m^{-2} day^{-1}))[7]. These five functions were estimated for each transect using individual-based bioenergetic models predicting fluxes of carbon (C), nitrogen (N), and phosphorus (P) (e.g. 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 quantified the input parameters, including elements of metabolism, growth, and diet and body stoichiometry, for all 1110 species through the integration of empirical data, data synthesis, and Bayesian phylogenetic models (see supplementary methods). We then ran a unique bioenergetic model for each combination of species identity, body size, and sea surface temperature ($n = 30668$) 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 deviding the sum by the surface area. Each function is thus Epressed as dry mass (of C, N, or P) per day per square meter. 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[33]. To reduce the occurrence of misclassification of herbivores and piscivores, we categorized a species as a herbivore or piscivore 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. (2020)[33]. Further, as a comparison, we quantified herbivory and piscivory rates using two alternative trophic guild classifications based on Expert opinion (Extended Data Fig. 3)[9,33]. Both the herbivory and piscivory rates match the Expert opinion trophic guild classifications. Finally, we estimated multifunction, i.e. one measure that combines all five functions by taking the geometric average of the five functions (normalized to a range between zero and 100). We used the geometric mean because functions are dependent on each other and vary by several orders of magnitude.

3. 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. The total length is based on visual estimations by divers. The immaturity is quantified using the following formula:

$$\mathrm{immaturity_i} = \kappa(\mathrm{l_\infty} - \mathrm{l_i}),$$

where κ is the species-specific growth rate parameter and 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].

4. Multivariate regression models

We fitted three multivariate Bayesian models with all five functions to (1) predict functions on the locality level to create a maps of functions, (2) investigate the effects of biomass and SST, and 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[35]. In the underwater visual transect database, 291 transects (3%) did not contain herbivores and 4467 transects (49%) did not contain piscivores yielding zeros for herbivory and piscivory, respectively. We considered that these absence 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 (i.e. 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 intercepts and random effects for localities and sites:

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

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

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

where i is the index of the transect, $y_{E_{\mathrm{N}},i}$ is the N excretion rate of transect i, $y_{E_{\mathrm{P}},i}$ is the P excretion rate, $y_{B,i}$ is the biomass production rate, $y_{H,i}$ 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. Locality- 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 between functions (independent of biomass and SST) on three levels: locality, site, and transect.

We used non-centered parameterization for site and location effects and all standard deviations had the following prior: $\sigma \sim student(3,0,2.5)$. We used a prior (lkj_{corr}) for each of the three correlation matrices ($R \sim lkj_{c}orr(1)$).

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[35]. Furthermore, because of the known relationship between temperature and parameters related to growth and respiration (see 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:

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

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

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

where $\beta1_{E_{\mathrm{N}}}, \beta1_{E_{\mathrm{P}}}, \beta1_{B}, \beta1_{H}, \beta1_{P}$ are the fixed effects of the log-transformed biomass, and $\beta2_{E_{\mathrm{N}}}, \beta2_{E_{\mathrm{P}}}, \beta2_{B}, \beta2_{H}, \beta2_{P}$ are the fixed effects of SST. Locality- 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 ($\beta1 \sim normal(1,1)$, $\beta2 \sim normal(0,1)$).

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 adding all community structure variables:

$$\mu_{\mathrm{function,i}} = \beta0_{function} + \beta1_{function} log(biomass), i + \beta2_{function} SST, i + \beta3_{function} richness, i + \beta4_{funct}$$

where $richness$ is the species richness, $size$ is the total length, $troph$ is the trophic level, imm 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 ($\beta3 - \beta12 \sim normal(0,1)$) and the same priors as described above for other parameters.

All Bayesian models were fitted using the R package *brms*[36], which uses Stan, a C++ package to perform full Bayesian inference[37]. The posterior distributions of model parameters were estimated using Hamiltonian Monte Carlo (HMC) 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).

5. Species dominance and contributions to functions

We quantified the relative contribution of each species to each function for all sites as follows:

$$\mathrm{contribution}_{f,i,j} = \frac{F_{f,i,j}}{\sum F_{f,j}},$$

where i is a certain species, j is a site, 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, 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. Specifically, the degree of dominance (DD) for a function performed by R species was calculated as follows:

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

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:

$$\mathrm{A}_{min} = \frac{R^2 - 1}{2R},$$

$$\mathrm{A}_{max} = R - 1,$$

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

where C_i is the contribution of a certain species and R 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 degree of dominance 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 (i.e. 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 their contribution is higher than 1/R (i.e. they contribute more than the situation in which each species contributes equally to a certain function).

References

1. Welti, N. *et al.* Bridging food webs, ecosystem metabolism, and biogeochemistry using ecological stoichiometry theory. *Frontiers in Microbiology* **8**, 1298 (2017).
2. Ceballos, G. *et al.* Accelerated modern human-induced species losses: Entering the sixth mass extinction. *Science Advances* **1**, 1–5 (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. vol. 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. vol. 33 948–961 (2019).
6. Williams, G. J. *et al.* Coral reef ecology in the Anthropocene. *Functional Ecology* **33**, 1014–1022 (2019).
7. Brandl, S. J. *et al.* Coral reef ecosystem functioning: Eight core processes and the role of biodiversity. *Frontiers in Ecology and the Environment* **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. *Proceedings of the National Academy of Sciences of the United States of America* **111**, 13757–62 (2014).
10. Mora, C. *et al.* Global Human Footprint on the Linkage between Biodiversity and Ecosystem Functioning in Reef Fishes. *PLoS Biology* **9**, e1000606 (2011).
11. Barneche, D. R. *et al.* Scaling metabolism from individuals to reef-fish communities at broad spatial scales. *Ecology Letters* **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. *Global Change Biology* **20**, 2459–2472 (2014).
14. Morais, R. A. & Bellwood, D. R. Pelagic Subsidies Underpin Fish Productivity on a Degraded Coral Reef. *Current Biology* **29**, 1521–1527.e6 (2019).
15. Morais, R. A., Connolly, S. R. & Bellwood, D. R. Human exploitation shapes productivity–biomass relationships on coral reefs. *Global Change Biology* **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. *Global Ecology and Biogeography* **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. *Functional Ecology* **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. *Proceedings of the National Academy of Sciences* **112**, 2617–2622 (2015).

19. Morais, R. A. & Bellwood, D. R. Global drivers of reef fish growth. *Fish and Fisheries* **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. *Ecology Letters* **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. *Science Advances* **5**, (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. *Conservation Letters* **12**, e12638 (2019).

25. Bellwood, D. R., Hughes, T. P. & Hoey, A. S. Sleeping Functional Group Drives Coral-Reef Recovery. *Current Biology* **16**, 2434–2439 (2006).

26. Darling, E. S. & D'agata, S. Coral Reefs: Fishing for Sustainability. vol. 27 R65–R68 (2017).

27. Graham, N. A. J. *et al.* Human Disruption of Coral Reef Trophic Structure. *Current Biology* **27**, 231–236 (2017).

28. Graham, N. A. J. *et al.* Dynamic fragility of oceanic coral reef ecosystems. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 8425–8429 (2006).

29. Burkepile, D. E. *et al.* Nutrient supply from fishes facilitates macroalgae and suppresses corals in a Caribbean coral reef ecosystem. *Scientific Reports* **3**, 1493 (2013).

30. Graham, N. A. J. *et al.* Changing role of coral reef marine reserves in a warming climate. *Nature Communications* **11**, 1–8 (2020).

31. Reynolds, R. W. *et al.* Daily High-Resolution-Blended Analyses for Sea Surface Temperature. J. Climate, 20, 5473-5496. (2007).

32. Froese, R., Thorson, J. T. & Reyes, R. B. A Bayesian approach for estimating length-weight relationships in fishes. *Journal of Applied Ichthyology* **30**, 78–85 (2014).

33. Parravicini, V. *et al.* Delineating reef fish trophic guilds with global gut content data synthesis and phylogeny. *PLoS Biology* **18**, e3000702 (2020).

34. Froese, R. & Pauly, D. FishBase. *World Wide Web electronic publication.* (2018).

35. 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).

36. Bürkner, P.-C. brms : An R Package for Bayesian Multilevel Models using Stan. *Journal of Statistical Software* **80**, 1–28 (2017).

37. Carpenter, B. *et al.* Stan : A Probabilistic Programming Language. *Journal of Statistical Software* **76**, 1–31 (2017).

Acknowledgements: We thank the staff at CRIOBE, Moorea for field support. We would also like to thank Jérémy Carlot, Beverly French, Titouan Roncin, Yann Lacube, Camille Gache, Gabrielle Martineau, Kailey Bissell, Benoit Espiau, Calvin Quigley, Kaitlyn Landfield and Tommy Norin for their help in the field, Guillemette de Sinéty and Jérémy Wicquart for their contribution to otolith analysis, and Sophie Schiettekatte for proof-reading 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 product of the SCORE-REEF group funded by the Centre de Synthèse et d'Analyse sur la Biodiversité (CESAB) of the Foundation pour la Recherche sur la Biodiversité (FRB) and the Agence Nationale de la Biodiversité (AFB). VP was supported by the Institut Universitaire de France (IUF) and JMC was supported by a Make Our Planet Great Again Postdoctoral Grant (mopga-pdf-0000000144). **Author contributions:** NMDS and VP conceived the idea and NMDS, VP, SJB, and JMC designed methodology; NMDS, JMC, SJB, AM, FM, VP, KSM, JEA and DEB collected the data; All authors shared existing data. NMDS analyzed 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:** None declared. **Data and materials availability:** All data and code to reproduce the figures are available on GitHub (https://github.com/nschiett/global_proc) and figshare (https://figshare.com/s/f789aec2c20492c4f0f9). All data on individual empirical measurements are available from the corresponding author upon request.

Figure legends

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 color scales showing categorical assignments (black = < 25%, grey = 25-75%, color = >75%). Black outlines highlight the five localities with the highest values of each function. Multifunctionality represents the weighted average of the five standardized functions. Right: The predicted values for functions and multifunctionality with increasing biomass. The lines represent the average modeled relationship and the shaded areas show the 95% credible intervals of the predictions.

Fig. 2: Correlations of the five functions, accounting for biomass and sea surface temperature. a: Modeled correlation coefficients of residual errors. Dots represent the average and lines represent the 95%CI. b-k: Scatter plots of the mean residual errors of the functions.

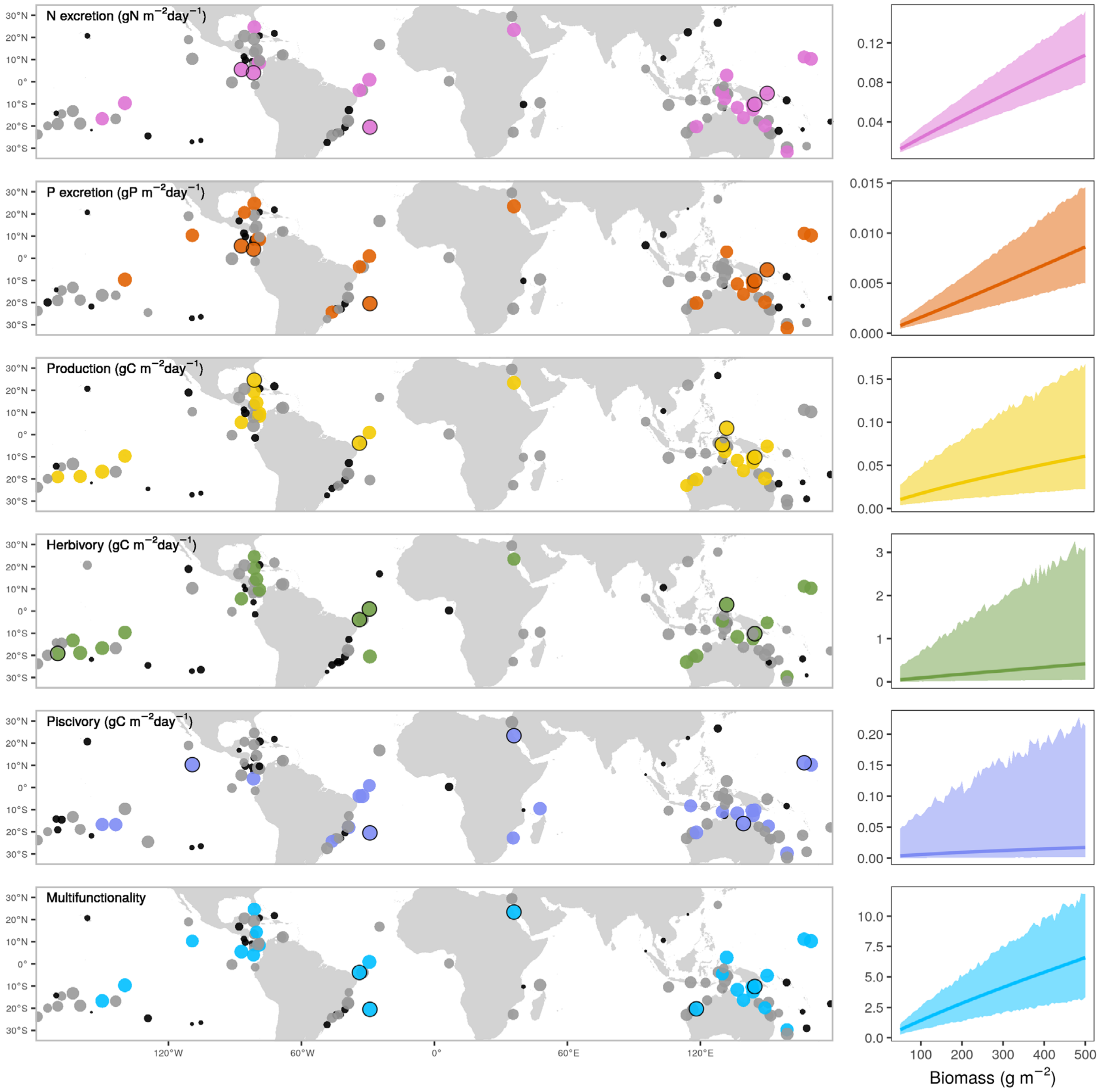
Fig. 3. Effects of ecological community variables on the five functions. Dots indicate fixed effect values from Bayesian linear regressions that examine the effects of species richness, trophic level, size, and immaturity of fishes. To represent both the median and spread of trophic level, size, and 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 2 for parameter values on non-standardized data). Dots represent the average effect size estimate, and horizontal lines indicate the 95% credible interval. Immaturity is defined as the derivative of the von Bertalanffy growth model for a given size; thus, the higher this value, the younger the individual.

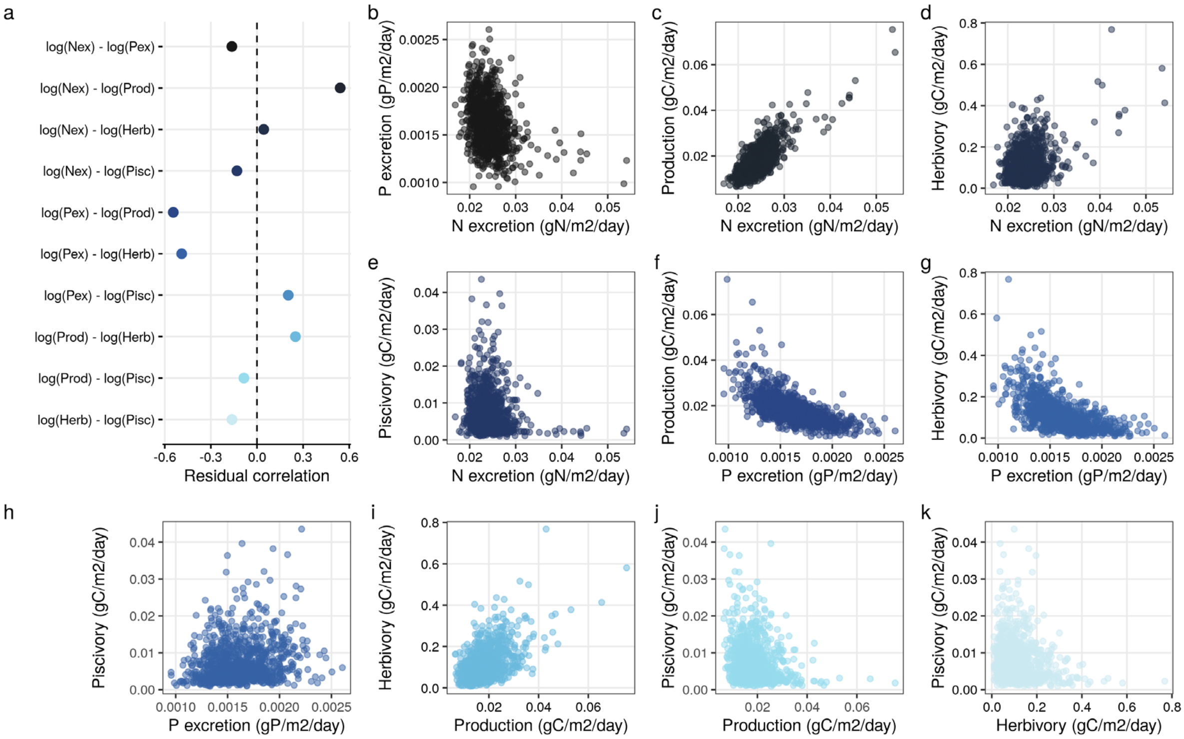
Fig. 4: Local dominance in species contributions to five ecosystem functions on coral reefs. a: The degree of dominance for each function at the site level. The degree of dominance of a community ranges between zero (all species contribute equally to the function) and one (a single species is the sole contributor to a given function). Colored dots represent the raw values, and the black dots and lines display the mean and 95% credible intervals of degree of dominance among all sites. In some cases, the credible interval was 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

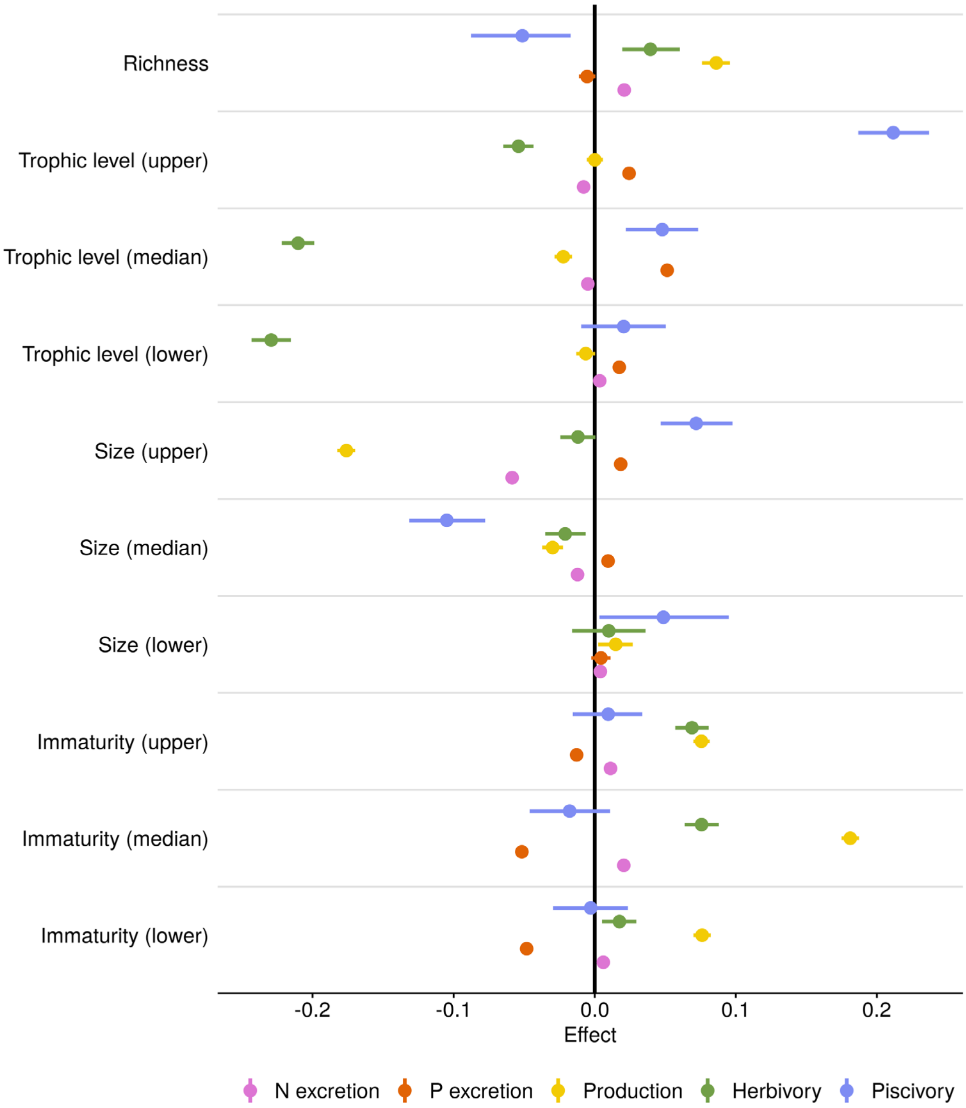
in at least one site relative to the total number of species, or, for herbivory and piscivory, the total number of herbivores and piscivores, respectively. c: Species-specific frequencies of dominance in each function across all sites, ranging from zero (species are never dominant) to one (dominant wherever present). 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 (i.e. one divided by the number of species that contribute to the function).

Fig. 5. Vulnerability of the five functions to fishing and climate change-induced coral loss. Conceptual schematic of the potential ways in which fishing and climate change can affect functions through their known effects on biomass and community structure.

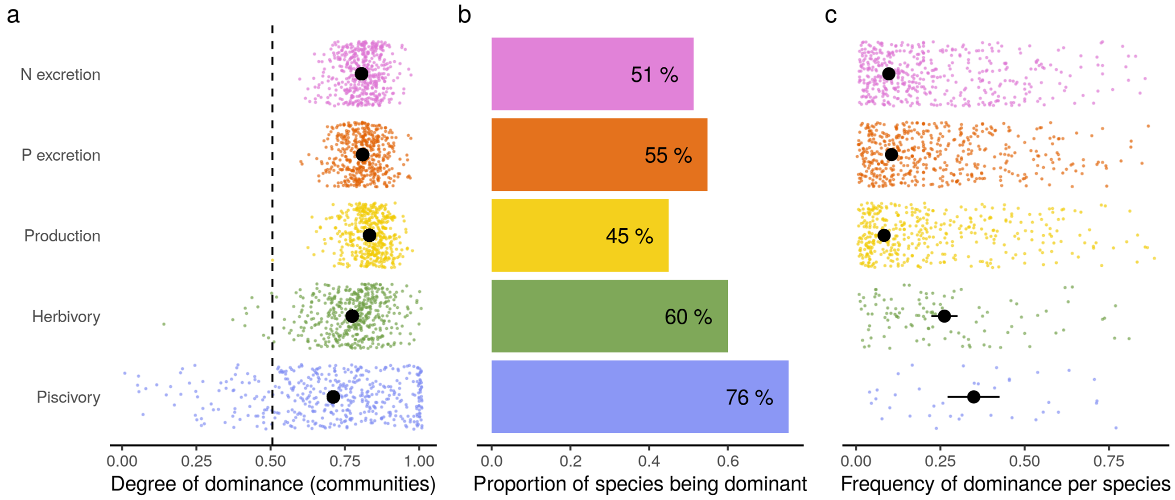
Figures

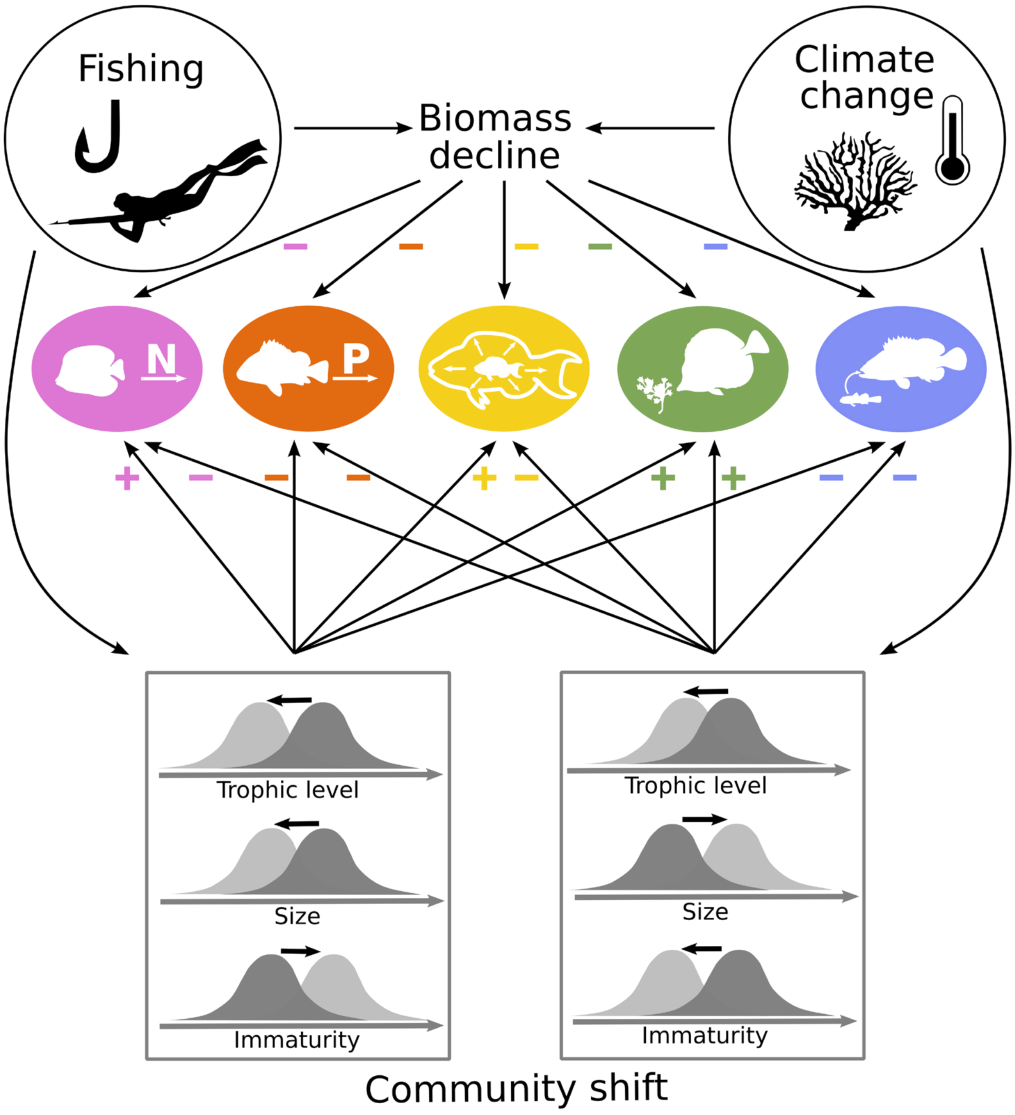

N excretion (gN m^{-2}day^{-1})
P excretion (gP m^{-2}day^{-1})
Production (gC m^{-2}day^{-1})
Herbivory (gC m^{-2}day^{-1})
Piscivory (gC m^{-2}day^{-1})
Multifunctionality
30°N
20°N
10°N
0°
10°S
20°S
30°S
120°W
60°W
0°
60°E
120°E
0.12
0.08
0.04
0.015
0.010
0.005
0.000
0.15
0.10
0.05
0.00
3
2
1
0
0.20
0.15
0.10
0.05
0.00
10.0
7.5
5.0
2.5
0.0
100
200
300
400
500
Biomass (g m^{-2})


a
log(Nex) - log(Pex)
log(Nex) - log(Prod)
log(Nex) - log(Herb)
log(Nex) - log(Pisc)
log(Pex) - log(Prod)
log(Pex) - log(Herb)
log(Pex) - log(Pisc)
log(Prod) - log(Herb)
log(Prod) - log(Pisc)
log(Herb) - log(Pisc)
-0.6
-0.3
0.0
0.3
0.6
Residual correlation
b
P excretion (gP/m2/day)
0.0025
0.0020
0.0015
0.0010
N excretion (gN/m2/day)
0.02
0.03
0.04
0.05
c
Production (gC/m2/day)
0.06
0.04
0.02
N excretion (gN/m2/day)
d
Herbivory (gC/m2/day)
0.8
0.6
0.4
0.2
0.0
N excretion (gN/m2/day)
e
Piscivory (gC/m2/day)
0.04
0.03
0.02
0.01
0.00
N excretion (gN/m2/day)
f
Production (gC/m2/day)
P excretion (gP/m2/day)
0.0010
0.0015
0.0020
0.0025
g
Herbivory (gC/m2/day)
P excretion (gP/m2/day)
h
Piscivory (gC/m2/day)
P excretion (gP/m2/day)
i
Herbivory (gC/m2/day)
Production (gC/m2/day)
0.02
0.04
0.06
j
Piscivory (gC/m2/day)
Production (gC/m2/day)
k
Piscivory (gC/m2/day)
Herbivory (gC/m2/day)
0.0
0.2
0.4
0.6
0.8

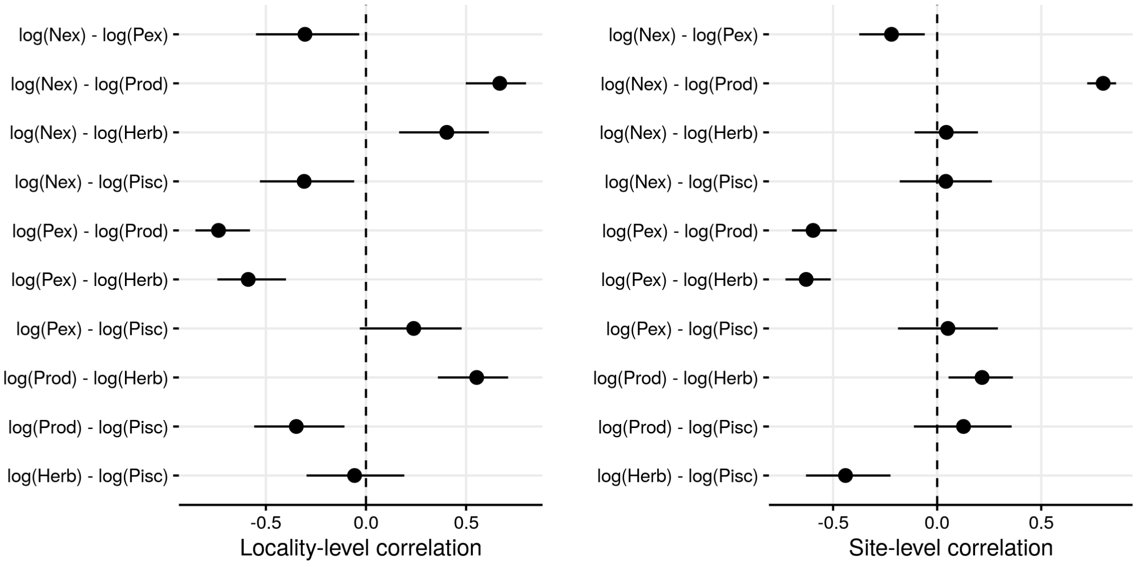

Richness
Trophic level (upper)
Trophic level (median)
Trophic level (lower)
Size (upper)
Size (median)
Size (lower)
Immaturity (upper)
Immaturity (median)
Immaturity (lower)
-0.2
-0.1
0.0
0.1
0.2
Effect
N excretion
P excretion
Production
Herbivory
Piscivory

474

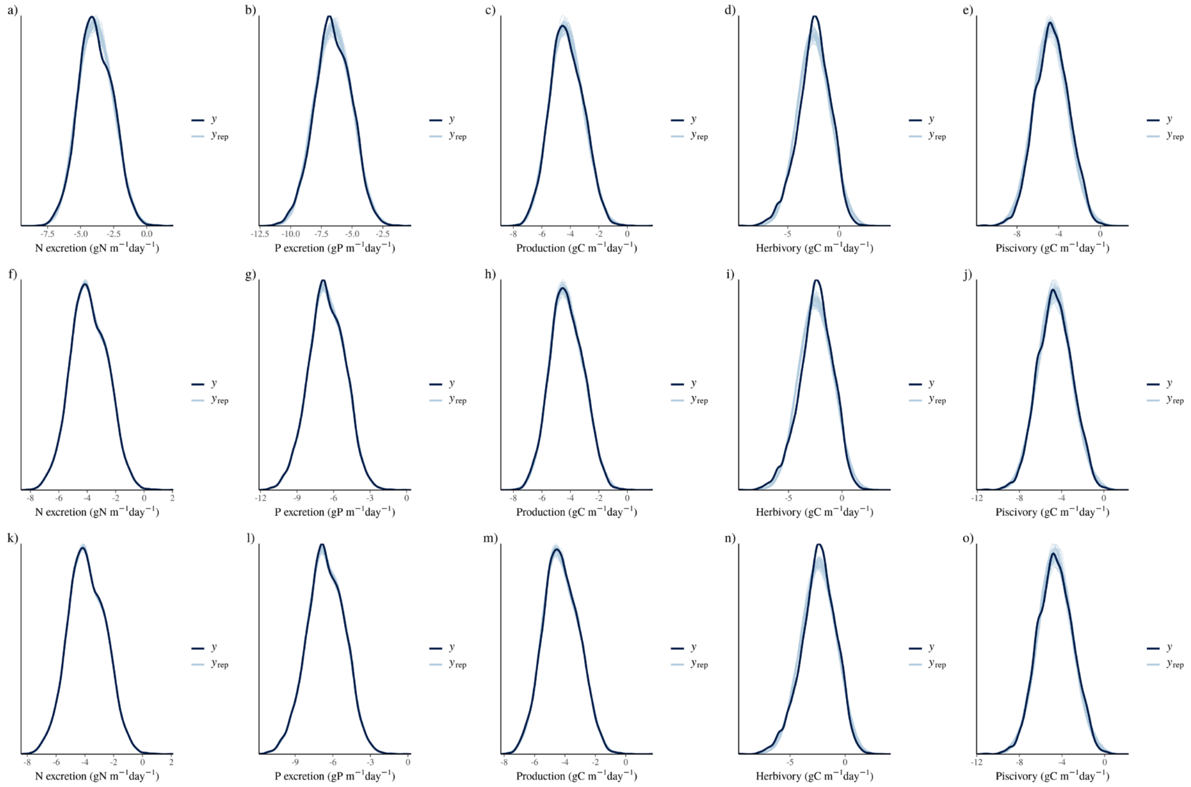

a
N excretion
P excretion
Production
Herbivory
Piscivory
0.00
0.25
0.50
0.75
1.00
Degree of dominance (communities)
b
51 %
55 %
45 %
60 %
76 %
0.0
0.2
0.4
0.6
Proportion of species being dominant
c
0.00
0.25
0.50
0.75
Frequency of dominance per species


Fishing
Biomass decline
Climate change
–
–
–
–
–
N
P
+
–
–
–
+
–
+
+
–
–
Trophic level
Size
Immaturity
Trophic level
Size
Immaturity
Community shift

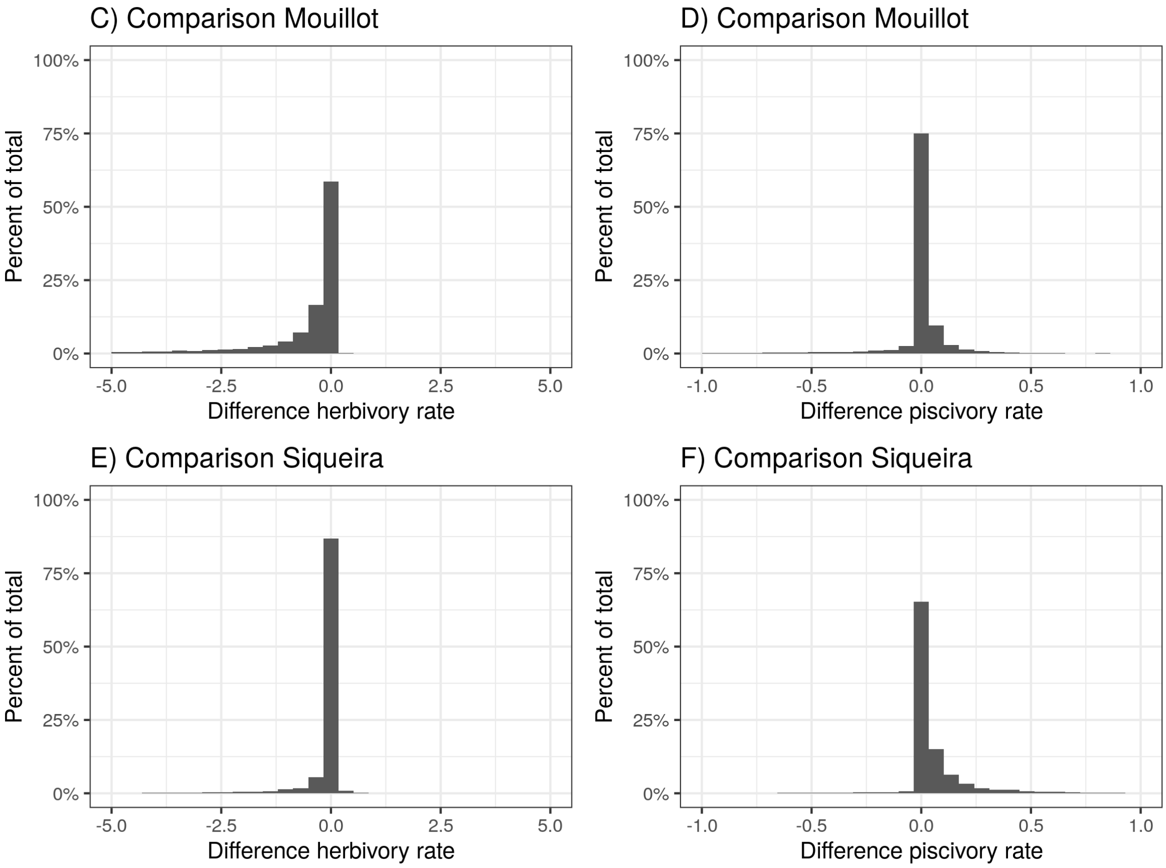
Extended data figures



Extended Data Fig. 1. Correlations among functions, independent of biomass and sea surface temperature, on the locality and site level.



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.



Extended Data Fig. 3: Comparison herbivory and piscivory rates when using alternative diet classifications from Mouillot et al. (2014) and Siqueira et al. (2020).