Marine protected areas regulate the structure of fish communities threatened by global warming and human impact

Andrea Tabi[1,*], Luis J. Gilarranz[2], Serguei Saavedra[3]

[1]Institute for Cross-Disciplinary Physics and Complex Systems (IFISC), Consejo Superior de Investigaciones Científicas (CSIC) and University of Balearic Islands, 07122 Palma de Mallorca, Spain

[2]Department of Aquatic Ecology, Eawag (Swiss Federal Institute of Aquatic Science and Technology) Überlandstrasse 133, 8600, Dübendorf, ZH, Switzerland

[3]Department of Civil and Environmental Engineering, MIT
77 Massachusetts Avenue, Cambridge, MA 02139, USA
[*] To whom correspondence should be addressed. E-mail: andrea@ifisc.uib-csic.es

Author contributions: All authors designed research; A.T. performed analyses; A.T and S.S. did calculations; L.J.G. compiled data; all authors interpreted results; A.T. wrote the first draft and all authors edited the manuscript.

Conflict of interest: The authors declare no conflict of interest.

Data deposition: The data and code supporting the results will be archived on Zenodo upon acceptance of the manuscript.

Abstract

Protecting and restoring the structure of marine communities is critical for supporting many biophysical processes including human well-being. Marine Protected Areas (MPAs) can successfully help to protect the abundance of target marine species, however, we know little about the cause-effect relationship between MPAs and community-wide properties of fish communities under the context of climate change and anthropogenic disturbances. Here, we show that MPAs mediate anthropogenic and climatic effects on fish diversity, while they directly increase by 28% the probability that a fish community displays a higher-than-average pyramid-like structure: energetically efficient systems with small basal species as the most abundant and large apex predators as the least abundant. To quantify these causal effects, we integrate a nonparametric causal-inference approach with ecological theory and observational data from 286 geographical sites with more than 1,400 fish species. Our work provides a quantitative platform to study how different factors may protect and restore marine communities.

Introduction

Global environmental change is causing an accelerated rate of biodiversity loss across ecosystems worldwide (1, 2). In this regard, coral reefs are one of the most sensitive marine habitats in decline (3), mainly due to thermal stress and human activities (4). Thermal stress can cause coral disease outbreaks (5, 6) and can disrupt the symbiotic relationships between coral and the photosynthetic algae inhabiting coral tissue ("bleaching"), which in turn can have cascading effects across entire marine communities (7). Additionally, human impacts, including land-based and ocean-based activities that modify, exploit, and pollute habitats have been reported to alter species composition (8). As a response, Marine Protected Areas (MPAs) have been established to protect biological diversity and economic resources (9). Indeed, MPAs can successfully help to protect the abundance of target marine species (10), however, we know little about the cause-effect relationship between MPAs and community-wide properties of marine communities under the context of climate change and anthropogenic disturbances (11). This difficulty mainly resides on the fact that no two communities are subject to the same exact internal (12, 13) (e.g., interspecific effects) and external conditions (14) (e.g., thermal stress), making extremely problematic the interpretation of relationships derived from multivariate regression and meta-analyses that are not structured by a causal hypothesis (15). Currently MPAs only account for 7.7% of the global ocean with a current plea for moving it up to 30% by 2030 (9). To strengthen this plea, it is then paramount to quantify the capacity of MPAs to protect and restore key properties of marine communities (16–19).

One of the most studied properties of marine communities is the distribution of species abundances as a function of body size (12, 20, 21)—commonly referred as *community structure*. Body size is considered a "master trait" that scales with organisms' physiology, regulating metabolic requirements (22), constraining feeding range (23), and shaping the trophic position of species in marine food webs due to energy transfers. It has been shown that interspecific interactions tend to be highly structured within trophic levels and consistent across marine communities (24), often leading to *pyramid-like* (also known as bottom-heavy) community structures, where small basal species are the most abundant and large apex predators the least abundant (20, 21, 25). Theoretically, pyramid-like community structures are expected to emerge and be explained under the energetic equivalence hypothesis (26) (the total energy flow through a population tends to be constant) and under metabolic scaling theory (27) (building on predator-prey mass ratios and transfer inefficiencies (25)). Empirically, however, deviations from pyramid-like community structures have often been observed in marine communities (12, 28, 29), which can be explained by a variety of reasons such as: complex predatory behavior (e.g., large predators feed on lower trophic levels or have wider diet width (12)), foraging of mobile consumers for energy subsidies provided by fish spawning grounds (13, 30), higher transfer efficiency (31), increased rates of trophic energy flux due to warming (27), or simply due to noise in local sampling (32). Therefore, disentangling the impact of internal ecological dynamics and external factors (e.g., climate change, human activity, interventions) on the structure of marine communities is

a non-trivial problem. Yet, this knowledge is central to increase our understanding about the factors regulating the tolerance to perturbations and internal efficiency of marine communities.

Here, to overcome the challenges above, we introduce a nonparametric causal-inference approach (33) to infer the average causal effect (ACE) that MPAs have on the structure of fish communities under the context of climate change and anthropogenic disturbances. We use observational data from coral-reef fish communities, comprising a total of 1,479 fish species together with spatial, temporal, and climatic factors across 286 sampling sites worldwide from the Reef Life Survey database (34) (Fig. 1, Notes). For each sampling location, we compile their external conditions: whether it is associated with a coral reef within a 10-km radius, if it is inside an MPA, how frequent it experienced thermal stress anomalies (TSA–weekly sea surface temperature minus the maximum weekly climatological sea surface temperature) using the sum of all the values of TSA between 1982 and 2019 where the average value of TSA was above 1 °C (35), and the degree of anthropogenic effects using as a proxy human population density (people per km^2) within 25 km radius (36) (Notes). Then, combining big data with ecological theory and a nonparametric causal-inference analysis, we provide evidence of the regulatory effect of MPAs on fish communities under the context of climate change and anthropogenic disturbances.

Methods

To disentangle the cause-effect relationship between MPAs and community structure, we develop our framework following a hypothesized causal graph among our observed variables. Specifically, regarding internal conditions, we hypothesize that MPAs regulate fish communities by directly altering species richness, interspecific interactions, and community structure. Regarding external conditions, we also hypothesize that MPAs counterbalance (mediate) the effect of coral reefs, human density, and thermal stress on species richness (37–39). We assume that human density, the presence of coral reefs, and thermal stress affect the establishment of MPAs. In turn, we also assume that human density and thermal stress anomalies affect the presence of coral reefs (7). Importantly, following previous work (11, 40), we assume that species richness affects the constraints imposed by interspecific interactions on the membership of species in a community—what we call *interaction constraints*. That is, it has been shown that the structure of interspecific interactions is generally strong and consistent across communities (24, 40), and it has been associated with limits to biodiversity (12). As a consequence, we hypothesize that such interaction constraints affect community structure. Figure S1 shows the complete, hypothesized, causal graph.

To quantitatively characterize community structure, we measure the structure of a fish community using the Spearman's rank correlation (ρ) between average species body masses and species abundances (Notes). The perfect pyramid-like structure, reflecting a perfectly monotonic size spectrum, is represented by the correlation value of $\rho = -1$. When the community structure gradually deviates

from pyramid structures, the correlation increases until the perfect inverse pyramid-like community structure (also known as top-heavy and less energetically efficient) reflected in $\rho = 1$. In our data (Fig. S2), communities show a wide spread distribution of structures (measured by ρ), and are centered on negative values of ρ ($median_\rho = -0.23$, $sd_\rho = 0.2$). These results confirm previous studies suggesting that communities are mainly characterized by pyramid-like community structures, but not uniquely.

Assuming that MPAs also regulate interspecific interactions by shifting trophic cascades (11), which in turn constrain the membership of species in a community, we use ecological theory to measure interaction constraints by the *incompatibility* of species coexistence. Following generalized Lotka-Volterra (gLV) dynamics, we quantitatively characterize interaction constraints using the fraction ($C(\mathbf{A})$) of environmental contexts (parameter values $\boldsymbol{\theta}$) *incompatible* with the possible coexistence (*feasibility*) of species in a fish community as a function of its *interaction matrix* $\mathbf{A}$ describing the *direct interspecific effects* (41, 42). To infer such interaction matrix $\mathbf{A}$, first we infer the *total interspecific effect* α_{ij} of a species j on species i using the data on how the abundance of species i is mediated by the presence of species j across different communities (Fig. 2A, Notes). This total interspecific effect is the chain of direct interspecific effects (43), stemming from a variety of interaction types such as feeding relationships, competition for resources and shelter, or heterospecific cueing for predation risk or habitat quality (44). Then, the interaction matrix $\mathbf{A}$ can be approximated by the negative inverse of the *total interaction matrix* ($-\boldsymbol{\alpha}^{-1}$) (43). Next, we approximate the interaction constraints ($C(\mathbf{A})$) by the *complement* of the normalized volume of parameter space of *effective growth rates* ($\boldsymbol{\theta}$) leading to positive solutions under a gLV model (Fig. 2B, Notes). This normalized volume can be interpreted as the expected probability of feasibility of a species in a community and it can be compared across communities of different sizes (45). Importantly, this expected probability of feasibility is valid for a wide family of gLV models with linear and nonlinear functional responses, stochastic dynamics, and as a lower bound for complex polynomial models in species abundances (41). Thus, the larger the interaction constraints ($C(\mathbf{A})$), the less likely it is to find environmental contexts ($\boldsymbol{\theta}$) compatible with the feasibility of a community. Note that $C(\mathbf{A}) = 0.5$ corresponds to the expected value of a community characterized by unstructured interspecific effects (also equivalent to a fully non-interacting community). In our data (Fig. S2), communities display predominantly values of $C(\mathbf{A}) > 0.5$, confirming previous studies suggesting that communities are characterized by consistent structured interspecific interactions constraining the membership of species (24).

Lastly, to formally study the cause-effect relationship between MPAs and community structure under the context of our studied internal and external variables, we follow a nonparametric causal-inference approach using do-calculus (33). This approach allows us to translate (whenever possible) interventional conditional distributions $P(Y = y|do(X = x))$ into observational conditional distributions $P(Y = y|X = x)$ (46). Note that applying the $do(X = x)$ operator to observations involves the removal of all incoming effects from the intervened variable in a causal graph (33). The causal effect of

X on Y can then be quantified as the average causal effect (ACE$_{XY}$: $\frac{\partial}{\partial x}$E[Y|do(X=x)]). To both standardize and simplify our analysis, we transform all quantitative variables into binary variables based on the median values (Fig. S2). That is, values above the median are translated as $X = 1$, otherwise $X = 0$. Hereafter, we refer to $X = 1$ (resp. $X = 0$) to *high* (resp. *low*) values. Note that other variables are already binary by construction, such as the presence or absence of MPAs and coral reefs. Formally, this binarization simplifies the analysis into ACE$_{XY} = P(Y = 1|do(X = 1)) - P(Y = 1|do(X = 0))$. We also assume that the causal graph is Markovian, meaning that measured variables (nodes) are affected by mutually, independent, unknown, random variables. We corroborated our causal graph by confirming all possible d-separations (conditional independencies) emerging from the hypothesized causal graph (Table S1 and Fig. S1). Importantly, this causal graph satisfies the conditions to measure genuine direct causation between MPAs and community structure (33, 46) (Notes and Table S2).

Results

Focusing on the observational patters emerging from our studied data, we found that fish communities that are inside MPAs and those that display *high* interaction constraints have on average lower correlation values (ρ) between species body masses and species abundances (compared to non-MPAs and low interaction constraints)—leading to stronger pyramid-like community structures (Figs. 3A-B). Moreover, we found that these patterns are consistent across any combination of internal factors, such as species richness and being or not near coral reefs (Figs.3C-D).

Importantly, our nonparametric causal-inference analysis revealed that MPAs have a genuine, direct, causal effect on community structure. Specifically, we found that the presence of MPAs directly increases by 28% the probability of observing fish communities with a *high* pyramid-like structure ($ACE_{MPA\rightarrow\rho} = -0.281$, recall that lower values of ρ represent stronger pyramid-like structures). Similarly, our analysis revealed that MPAs have an indirect causal effect on community structure (SI, Fig.4A). That is, first we found that MPAs directly increase by 14% the probability of observing a *high* species richness (S) near coral reefs ($ACE_{MPA\rightarrow S|coral=1} = 0.138$), but they directly reduce this probability by 24% in communities without coral reefs ($ACE_{MPA\rightarrow S|coral=0} = -0.243$). Second, we found that a *high* species richness directly decreases by 39% the probability of observing *high* interaction constraints (C) in an MPA ($ACE_{S\rightarrow C|MPA} = -0.385$), but it decreases by 60% in a non-MPA ($ACE_{S\rightarrow C|non-MPA} = -0.598$), signaling the enhanced requirement for lower constraints in non-protected communities with high biodiversity since they are at a higher risk of being perturbed. This hypothesis can also be corroborated by the direct effect of MPAs on interaction constraints (Fig. 4A), which we found to be negative ($ACE_{MPA\rightarrow C|S=0} = -0.222$) under *low* species richness and slightly positive ($ACE_{MPA\rightarrow C|S=1} = 0.073$) under *high* species richness. Third, we found that fish communities with *high* interaction constraints increase by 20% the probability of observing a *high* pyramid-like structure ($ACE_{C\rightarrow\rho} = -0.2$)—closing the indirect effect of MPAs on promoting a more energetically efficient community structure.

The regulatory effects of MPAs on community structure can be additionally corroborated by studying how MPAs can mediate the effects of anthropogenic and climatic factors on species richness inside and outside coral reefs (Fig. 4B). For example, it is expected that communities near coral reefs tend to have on average higher biodiversity than other communities. In this line, we found that MPAs increase the positive effect of coral reefs on species richness. Specifically, the presence of a coral reef increases by 54% the probability of observing a fish community with a *high* species richness if it is an MPA ($ACE_{Coral \rightarrow S|MPA=1} = 0.535$), but it only increases the probability by 15% if it is a non-MPA ($ACE_{Coral \rightarrow S|MPA=0} = 0.154$). Furthermore, we found that MPAs dampen the negative effect of TSA on species richness inside and outside coral reefs. While a *high* TSA decreases species richness in coral reefs ($ACE_{TSA \rightarrow S|MPA=1,coral=1} = -0.191$) (Fig. 4B), this biodiversity is impacted more strongly in non-protected areas ($ACE_{TSA \rightarrow S|MPA=0,coral=1} = -0.237$)—corroborating previous work (14, 47). Similarly, in areas without coral reefs, we found that a *high* TSA has almost no effect on species richness in MPAs ($ACE_{TSA \rightarrow S|MPA=1,coral=0} = 0.076$), but it decreases species richness by 30% in non-MPAs ($ACE_{TSA \rightarrow S|MPA=0,coral=0} = -0.297$). Note that previous research has shown that coral reefs with thermal stress shifted macroalgal states towards lower species richness, whereas reefs where live coral recovered exceeded pre-bleaching fish richness (14). Additionally, we found that MPAs reduce the detrimental effects of human density on species richness inside and outside coral reefs. Specifically, we found that while a *high* human density (H) decreases by 19% the probability of having a *high* species richness in MPAs near corals ($ACE_{H \rightarrow S|MPA=1,coral=1} = -0.199$), it decreases by 66% in non-MPAs near corals ($ACE_{H \rightarrow S|MPA=0,coral=1} = -0.662$). We observed similar patterns in communities without corals ($ACE_{H \rightarrow SR|MPA=1,coral=0} = -0.169$; $ACE_{MPA \rightarrow S|MPA=0,coral=0} = -0.283$). Depending on the level of protection, it is possible that these differences are due to the limited or lack of extractive activities within MPAs that can contribute to an increase in the abundance of species belonging to upper trophic levels (large predatory fishes) (10, 38).

Finally, it is worth mentioning that fish communities that are near coral reefs have 15% more probability to be inside MPAs (Fig. 4B). However, the distribution of MPAs heavily depends on human density, i.e., MPAs are less likely to be established close to highly-populated areas ($ACE_{H \rightarrow MPA|coral=0} = -0.115$), and this probability is even lower in highly populated areas near coral reefs ($ACE_{H \rightarrow MPA|coral=1} = -0.31$). Moreover, areas under high thermal stress are less likely to be inside MPAs if they are also near coral reefs ($ACE_{TSA \rightarrow MPA|coral=0} = 0.165$ vs. $ACE_{TSA \rightarrow MPA|coral=1} = 0.022$). These patterns highlight the need for an extension of MPAs as their effectiveness can help both to buffer the effects of human population density and climate change.

Conclusions

Our findings have shown that MPAs regulate the structure of marine communities and counterbalance the effects of human disturbances and thermal stress. This regulatory mechanism can

be summarized as follows: First, MPAs directly increase the probability that fish communities are observed with a stronger pyramid-like structure relative to their limits imposed by their own environmental contexts (MPAs can display a wide distribution of community structures as a function of the geographical pressures, which is consistent with empirical observations of top-heavy structures in pristine environments (31)). That is, while strong interaction constraints have a positive effect on pyramid-like structures and species-rich communities require weaker interaction constraints (e.g., near coral reefs), MPAs have an additional positive effect on pyramid-like structures. Second, MPAs have a mediation role in the negative effect caused by thermal and anthropogenic factors on species richness inside and outside coral reefs. Because species richness affects interaction constraints, which in turn affect community structure, these results corroborate the direct and indirect effects of MPAs on community structure.

In addition, our findings have suggested that in non-MPAs, where communities are under higher anthropogenic disturbances, the need for lower interaction constraints (higher tolerance to changes in environmental conditions) can outweigh the energetic efficiency described by the Eltonian pyramid shape (25). In other words, MPAs increase species richness near coral reefs, but reduces it in areas without coral reefs. Moreover, MPAs increase interaction constraints in species-rich communities, but they relax them in species-poor communities. This suggests that MPAs shift communities closer back to their potentially undisturbed state by apparently reducing the need for species mobility in search for food or new shelter, while simultaneously affecting interaction constraints and community structure. We hope future work can expand on this potential trade-offs between tolerance to changing environments and energetic efficiency in marine communities as they can hold the key to a better understanding of ecological dynamics in general.

As human density in coastal areas increases (48), and thermal stress anomalies become more frequent due to climate change (49), it has become necessary to establish successful interventions to protect and restore marine communities. In this line, under a specific set of assumptions (Notes), our results have pointed towards a positive genuine effect between MPAs and the structure of fish communities, as well as a mediation role buffering the negative impact of climate change and anthropogenic disturbances on fish communities. Note that ecological restoration is defined as the process of assisting recovery and management of ecological integrity, which typically requires a reference state (50). While this work has not studied the recovery of communities to a specific restoration baseline, our framework can be used as a quantitative causal platform to study the effect of different potential interventions and reference states that can help us to increase the success of different protection and restoration strategies of marine communities in general.

Notes

Data. We analyzed 462 sampled communities from 286 sites around the planet (Fig. 1) from the Reef Life Survey (RLS) database (34) comprising population data from 1,479 non-benthic marine species with individual body size information. Body size was measured as biomass and data were aggregated by year. We included only sampling sites in our analysis, which were surveyed more than once per year (51). This decision is based on a prior rarefaction analysis that we conducted to assess the impacts of sampling effort from the RLS database due to the noisiness of one sampling event in species richness (see Supporting Information for more details). We collected weekly sea surface temperature (SST) from NOAA's (National Oceanic and Atmospheric Administration) remote sensing database. We used the sum of thermal stress anomalies (TSA) in our analysis that was calculated as the weekly sea surface temperature (SST) minus the maximum weekly climatological SST. The distribution of warm-water coral reef was obtained from UNEP-WCMC World Fish Centre database (52). The information on marine protected areas was obtained from UNEP-WCMC and IUCN Potected Planet database (53). The information on human population density was obtained from Gridded Population of the World (54). The human population density is quantified as $humans/Km^2$ in a 25-km radius around the sampling site (48).

Inferring interspecific effects. We used a phenomenological approach to infer direct interspecific effects under a generalized Lotka-Volterra model: $\frac{dN_i}{dt} = N_i\left(\theta_i + \sum_{j=1}^{S} a_{ij}N_j\right)$, where N_i is the density of species i, θ_i represents the effective growth rate of species i as a function of the environmental context, $\mathbf{A} = (a_{ij})$ is the interaction matrix representing the direct per-capita effect of species j on an individual of species i, and we set $\alpha_{ii} = -1$ without loss of generality (42). To quantify the value of a_{ij}, let x_i^* denote the equilibrium abundance of the species i without species j. Let y_i^* and y_j^* denote the equilibrium abundance of species i and j when both are together. Then, the direct effect can be approximated as (42): $a_{ij} = \frac{y_i^* - x_i^*}{y_j^*}$. To infer a_{ij}, it is necessary to have monocultures and co-cultures (42). Because this is not possible to have those in our data, we estimated the values of x_i^*, y_i^*, and y_j^* for each species pair with an out-of-sample bootstrap method using the whole database filtered by sites where the target species occurred alone (x_i^*) or together (y_i^*). Then, the estimated interaction effects α_{ij} can be interpreted as the total interspecific effects (chains of direct interspecific effects) (43), which are defined as the negative inverse of the interaction matrix: $\boldsymbol{\alpha_i} = -\mathbf{A}_i^{-1}$ (43), where the sub-index corresponds to the ith bootstrap realization. This approach can be used for data not at equilibrium for x_i^* and y_i^*, as long as all inference and predictions are based on absolute data within the same time frame (55).

Calculating interaction constraints. The interaction constraints in a community are approximated by the fraction of environmental contexts incompatible with the possible coexistence of species in a community as a function of the interaction matrix $\mathbf{A}$. We estimate this fraction with the complement of the normalized volume of the parameter space of effective growth rates ($\boldsymbol{\theta}$) leading

to positive solutions under a gLV model (i.e., $\mathbf{N}^* = -\mathbf{A}^{-1}\boldsymbol{\theta} > 0$)—known as feasibility domain. Mathematically, this normalized volume corresponds to $\Omega(\mathbf{A}) = \frac{(D_F(\mathbf{A}) \cap \Theta^S)}{(\Theta^S)} \in [0, 0.5]$, where $D_F(\mathbf{A})$ is the feasibility domain and Θ^S is the $(S-1)$-dimensional closed unit sphere in dimension S (45). This fraction is a decreasing function of community size S. In order to compare $\Omega(\mathbf{A})$ across different community sizes, we re-scaled it to the individual level (45): $\omega(\mathbf{A}) = \Omega^{1/S}(\mathbf{A}) \in [0, 1]$. This can be interpreted as the expected probability of feasibility of a species in a community. Thus, the larger the complement of the expected probability of feasibility ($C(\mathbf{A}) = 1 - \omega(\mathbf{A}) \in [0, 1]$), the larger the interaction constraints (i.e., interactions constrain the membership of species in a community). Note that we inferred $\mathbf{A}$ from a bootstrap method, yielding potentially many matrices $\mathbf{A_k}$ (see Methods above). Hence, to obtain a single value for $C(\mathbf{A})$, we generated $k = 1000$ interaction matrices ($\mathbf{A_k}$) and took the median value of the distribution: $\overline{C(\mathbf{A})} = Median(C(\mathbf{A_k}))$.

Causal inference. We applied a nonparametric probabilistic approach to calculate cause-effect relationships rooted on do-calculus (33, 46). First, we tested our causal graph by confirming all the possible d-separations emerging from the directed acyclic graph (DAG) depicted in Figure S1 (also Fig.4). This test implies finding the conditional independencies (i.e., $X \perp\!\!\!\perp Y|Z$) between all the 8 pairs of non-adjacent nodes: (1) Coral $\perp\!\!\!\perp$ Constraints |{Species richness, MPA}, (2) Coral $\perp\!\!\!\perp$ Structure (ρ) |{Constraints, MPA}, (3) Human density $\perp\!\!\!\perp$ Constraints |{Species richness, MPA}, (4) Human density $\perp\!\!\!\perp$ Structure (ρ) |{Species richness, MPA}, (5) Human density $\perp\!\!\!\perp$ TSA |{$\emptyset$}, (6) TSA $\perp\!\!\!\perp$ Constraints |{Species richness, MPA}, (7) TSA $\perp\!\!\!\perp$ Structure (ρ) |{Constraints, MPA}, (8) Species richness $\perp\!\!\!\perp$ Structure (ρ) |{Constraints, MPA}. We confirmed all these 8 independencies using a G^2 test with 99% confidence level (see Table S2 in SI). Once we empirically corroborated our causal graph, we calculated the direct average causal effects (ACEs) following the rules of do-calculus. This implies working with a modified version of the causal graph ($P(Y|do(X)) = P_m(Y|X)$), where all the incoming links to the node X have been removed (46). We used product decomposition to simplify conditional independencies as much as possible, and to use the observed probability distributions (i.e., $P(Y|X)$) whenever $P_m = P$. That is, we substitute modified probabilities P_m by observational probabilities (P) if and only if they are both equivalent in the modified and original graph. For example, the direct ACE between MPAs and structure (ρ) is computed by: $ACE_{MPA \to \rho} P(\rho = 1|do(MPA) = 1) - P(\rho = 1|do(MPA) = 0) = P_m(\rho = 1|MPA = 1) - P_m(\rho = 1|MPA = 0)$. Recall that we transform all variables into binary values (1: above median, 0: below median). This in turn, can be calculated using observational distributions: $\sum_{constraints} P_m(\rho|MPA) = P(\rho|MPA, constraints)P(constraints)$. Importantly, per definition (33), our causal graph allows us to identify the genuine effect between MPAs and community structure (Table S2 in SI). To identify a genuine effect, one needs a set of variables Z and a context (set of values) T, such that Z is a potential cause of X, Z and Y are dependent given T, and Z and Y are independent given $T \cup X$. In the relationship between X=MPA and $Y = \rho$, $Z = \{TSA\}$, and $T = \{constraints\}$ (Table S2 in SI).

Assumptions. In this work we have established a series of important assumptions (e.g., Markovian graphs, binary variables, Lotka-Volterra dynamics) in order to simplify our framework and to allow us to estimate such causal effects. It would be fruitful in future work to investigate the consequences of relaxing some of these assumptions (in light of relevant ecological interpretations) and incorporating additional ecological variables, interventions, and reference states that can increase our understanding of different protection and restoration processes.

Acknowledgments: A.T. was supported by the Spanish State Research Agency, through the Maria de Maeztu Program for Units of Excellence in R&D (MDM-2017-0711). L.J.G. was supported by the Swiss National Science Foundation Ambizione Fellowship, PZ00P3_185951. S.S. was supported by NSF DEB-2024349 and MIT Sea Grant College Program.

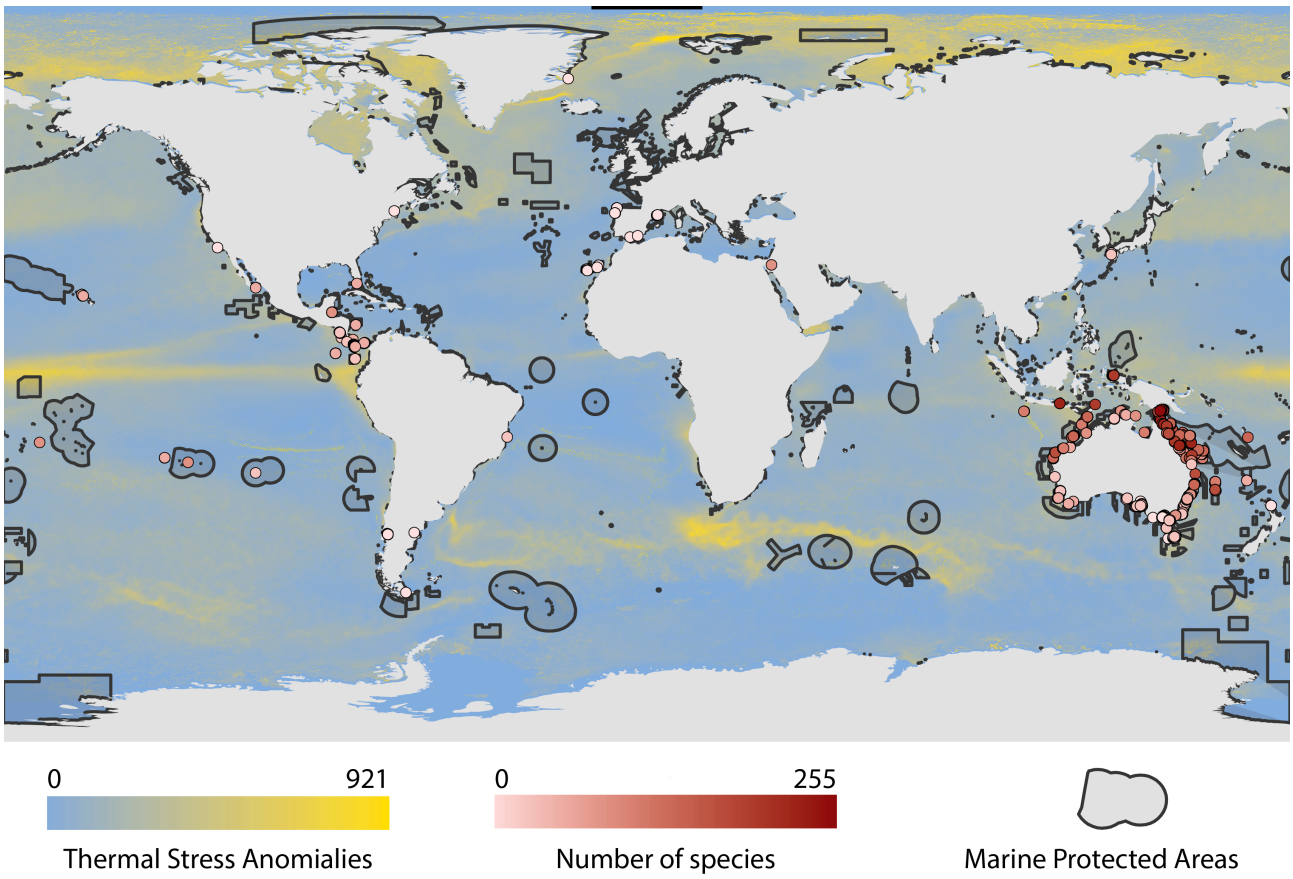

Figure 1. Global distribution of sampling sites and their attributes under our studied data set. We considered only sampling sites (286 sites in total) in our analysis, which were surveyed more than once per year. Data compiled from the Reef Life Survey database (34) (Notes). The color of the circles corresponds to the number of species observed at a given site. The background color corresponds to the thermal stress anomalies (TSA), which are calculated as the sum of all the values of TSA between 1982 and 2019 at which the average value of TSA was above 1 °C. The black lines show the borders of reported Marine Protected Areas (MPAs).

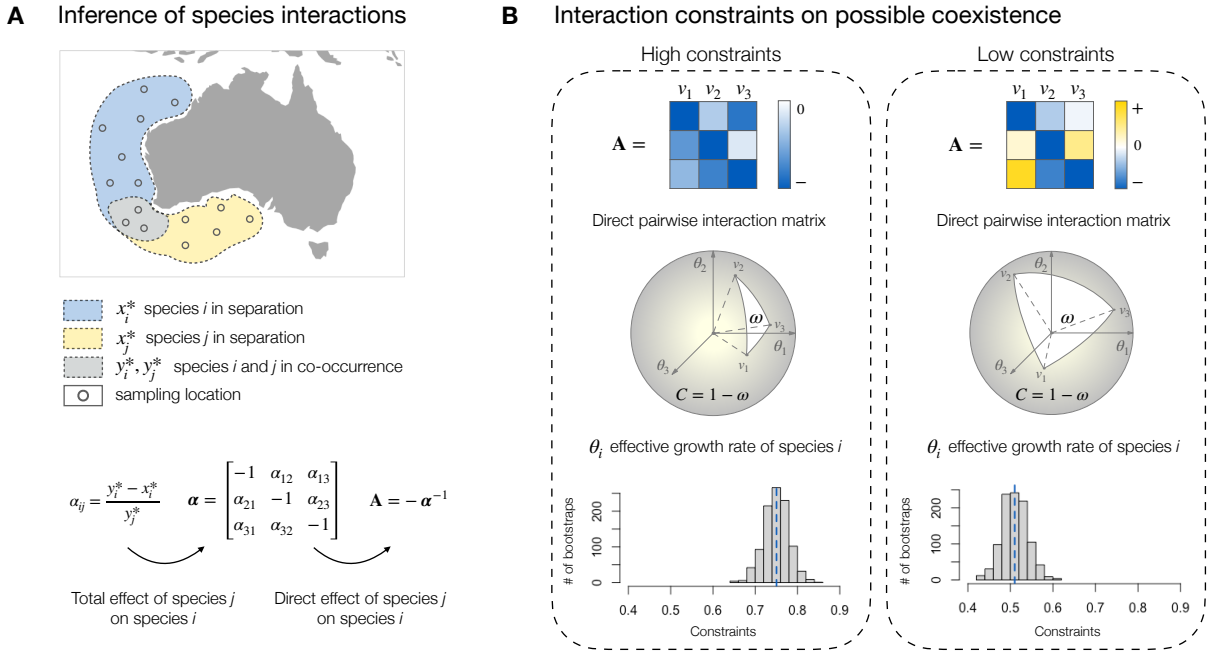

Figure 2. Inferring interspecific effects and interaction constraints. Panel **A**: The interaction matrix of direct interspecific effects ($\mathbf{A}$) was inferred assuming generalized Lotka-Volterra (gLV) dynamics (Notes). Absolute values of species abundances were bootstrapped from geographical regions, where species i and j occurred together (y_i^*, y_j^*) and when the target species i was not found together with species j (x_i^*). This estimation results in the total interspecific effect α_{ij} of species j on species i. The direct interspecific effects can then be approximated as $\mathbf{A}_k = -\boldsymbol{\alpha}_k^{-1}$ (43), where the sub-index k represents the kth bootstrap. Panel **B**: Strong, negative, direct, interspecific effects (blue colors) can be translated into high interaction constraints, whereas strong, positive, direct, interspecific effects (yellow colors) can be translated into low interaction constraints. Interaction constraints ($C(\mathbf{A})$) are quantified as the fraction of environmental contexts (parameter values) incompatible with the feasibility ($\omega(\mathbf{A})$) of a community: $C(\mathbf{A}) = 1 - \omega(\mathbf{A})$. Note that feasibility ($\omega(\mathbf{A})$) is defined by the column vectors ($\mathbf{v}$) of the interaction matrix $\mathbf{A}$, representing the combination of effective growth rates ($\boldsymbol{\theta}$) compatible with a positive equilibrium under gLV dynamics (i.e., $\mathbf{N}^* = -\mathbf{A}^{-1}\boldsymbol{\theta} = \boldsymbol{\alpha}\boldsymbol{\theta} > 0$). The larger $C(\mathbf{A})$ (resp. the smaller $\omega(\mathbf{A})$), the larger (resp. smaller) the interaction constraints. These constraints are calculated as the median value over the bootstrap distribution (dashed blue line), where $k = 1000$ possible realizations of interaction constraints are sampled for each community (Notes).

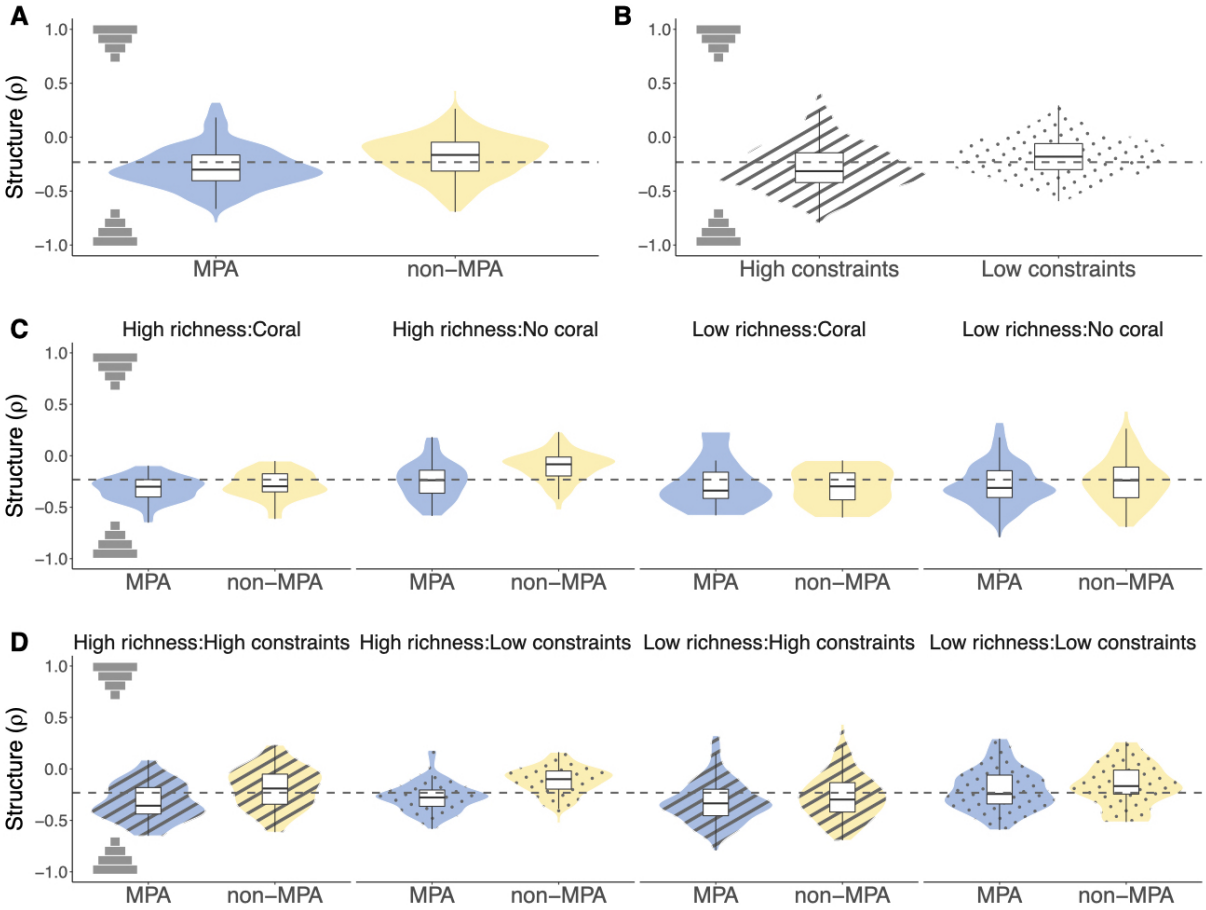

Figure 3. Distribution of community structure across different marine and geographical properties. The panels show the distribution of community structures, measured as the Spearman's rank correlation coefficients (ρ) between average body size and abundance. Panel **A**: distribution between MPAs vs. non-MPAs. Panel **B**: Distribution of high vs. low interaction constraints. Panel **C**: distribution of MPAs vs. non-MPAs divided into the combination of species richness and presence of coral. Panel **D**: Distribution of MPA vs non-MPA divided into the combinations of interaction constraints and species richness. Note that interaction constraints and species richness were converted into binary variables (higher or lower than the median value). The grey dashed lines correspond to the median value of community structure.

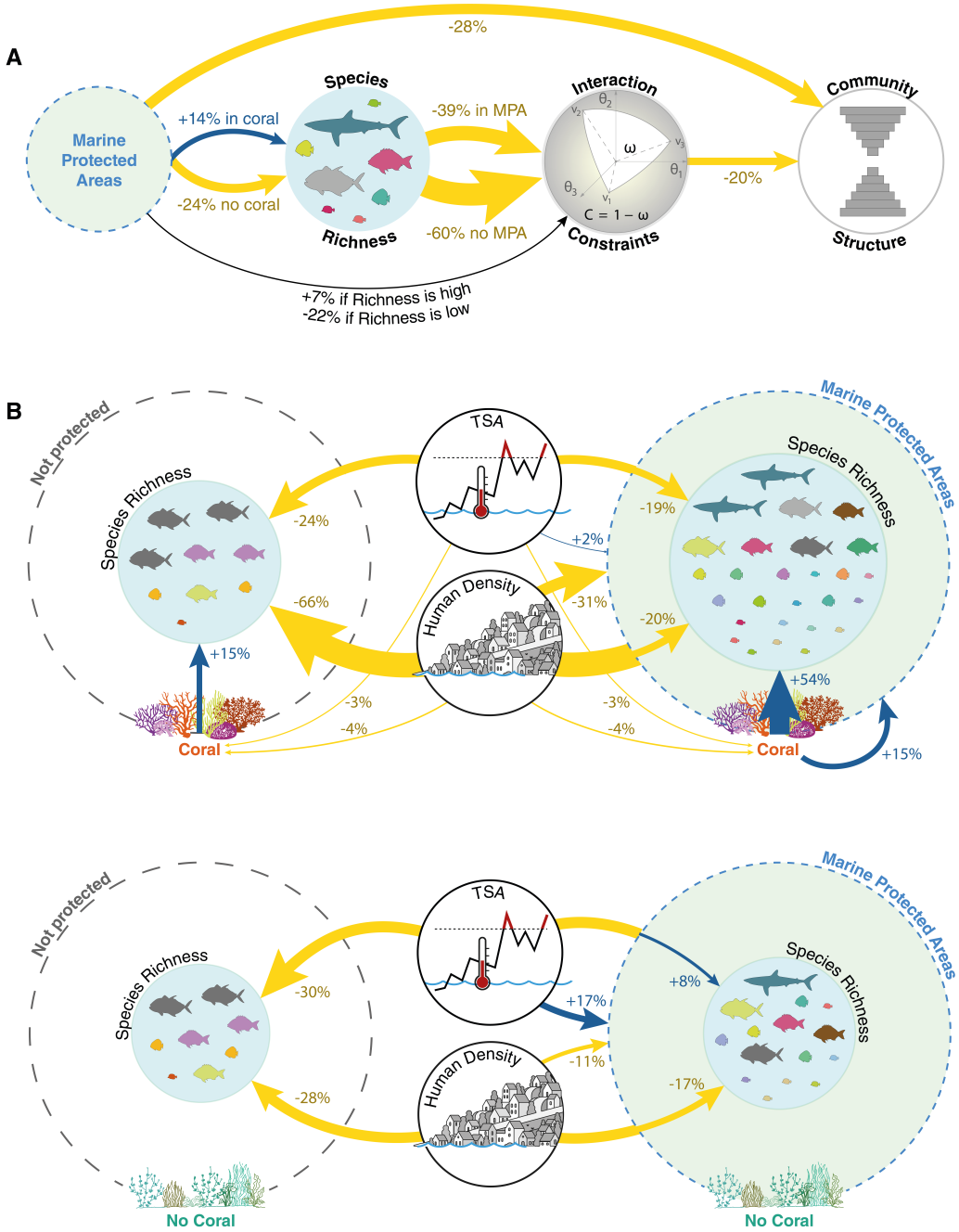

Figure 4. MPAs regulate the structure of marine communities. Panel **A**: This part of the causal graph (the complete is shown in Fig.S1) depicts the average causal effects (ACE) of MPAs on community structure. MPAs increase by 28% the probability of having a pyramid-like structure (i.e., a lower-than-average ρ). Following the rules of do-calculus, we show the direct ACE between MPAs, species richness, interaction constraints, and community structure (ρ). Note that all variables have been transformed to binary values, where a value of 1 is a data point above the median and 0 otherwise. Recall that low values of community structure represent pyramid-like structures. The ACE then corresponds to the probability difference in observing a value of 1 in the target effect when the cause takes a value of 1 compared to when it takes a value of 0. Panel **B**: This other part of the causal graph (the complete is shown in Fig.S1) shows the effect of human density and thermal stress anomalies (TSA) on species richness in MPAs vs. non-MPAs combined with coral vs. no coral. MPAs generally mitigate the detrimental impacts of human density and thermal stress on species richness in both coral and no coral regions.

References

(1) IPCC (2019) Special Report on the Ocean and Cryosphere in a Changing Climate [H.-O. Pörtner, D.C. Roberts, V. Masson-Delmotte, P. Zhai, M. Tignor, E. Poloczanska, K. Mintenbeck, A. Alegría, M. Nicolai, A. Okem, J. Petzold, B. Rama, N.M. Weyer (eds.)]., Technical report.

(2) McCormack SA, et al. (2019) Using stable isotope data to advance marine food web modelling. *Reviews in Fish Biology and Fisheries* 29(2):277–296.

(3) Mora C, et al. (2006) Coral Reefs and the Global Network of Marine Protected Areas. *Science* 312(5781):1750–1751.

(4) Muñiz-Castillo AI, et al. (2019) Three decades of heat stress exposure in Caribbean coral reefs: a new regional delineation to enhance conservation. *Scientific Reports* 9(1):11013.

(5) Garren M, et al. (2014) A bacterial pathogen uses dimethylsulfoniopropionate as a cue to target heat-stressed corals. *The ISME Journal* 8(5):999–1007.

(6) Shapiro OH, Kramarsky-Winter E, Gavish AR, Stocker R, Vardi A (2016) A coral-on-a-chip microfluidic platform enabling live-imaging microscopy of reef-building corals. *Nature Communications* 7(1):10860.

(7) Ainsworth CH, Mumby PJ (2015) Coral–algal phase shifts alter fish communities and reduce fisheries production. *Global Change Biology* 21(1):165–172.

(8) Halpern BS, et al. (2008) A Global Map of Human Impact on Marine Ecosystems. *Science* 319(5865):948–952. Publisher: American Association for the Advancement of Science.

(9) Lewis N, et al. (2017) Large-Scale Marine Protected Areas: Guidelines for design and management., (Gland, Switzerland: IUCN), Technical Report No. 26.

(10) Grorud-Colvert K, et al. (2021) The MPA Guide: A framework to achieve global goals for the ocean. *Science* 373(6560):eabf0861.

(11) Aburto-Oropeza O, et al. (2011) Large Recovery of Fish Biomass in a No-Take Marine Reserve. *PLOS ONE* 6(8):e23601.

(12) Woodson CB, Schramski JR, Joye SB (2018) A unifying theory for top-heavy ecosystem structure in the ocean. *Nature Communications* 9(1):1–8.

(13) Mourier J, et al. (2016) Extreme Inverted Trophic Pyramid of Reef Sharks Supported by Spawning Groupers. *Current Biology* 26(15):2011–2016.

(14) Robinson JPW, Wilson SK, Jennings S, Graham NAJ (2019) Thermal stress induces persistently altered coral reef fish assemblages. *Global Change Biology* 25(8):2739–2750.

(15) Bareinboim E, Pearl J (2016) Causal inference and the data-fusion problem. *Proceedings of the National Academy of Sciences* 113(27):7345–7352.

(16) MacKeracher T, Diedrich A, Simpfendorfer CA (2019) Sharks, rays and marine protected areas: A critical evaluation of current perspectives. *Fish and Fisheries* 20(2):255–267.

(17) Downing AL, Leibold MA (2002) Ecosystem consequences of species richness and composition in pond food webs. *Nature* 416(6883):837–841.

(18) Williams SL, Ambo-Rappe R, Sur C, Abbott JM, Limbong SR (2017) Species richness accelerates marine ecosystem restoration in the Coral Triangle. *Proceedings of the National Academy of Sciences* 114(45):11986–11991.

(19) Maureaud A, et al. (2019) Biodiversity–ecosystem functioning relationships in fish communities: biomass is related to evenness and the environment, not to species richness. *Proceedings of the Royal Society B: Biological Sciences* 286(1906):20191189.

(20) Sheldon RW, Prakash A, Sutcliffe Jr. WH (1972) The Size Distribution of Particles in the Ocean. *Limnology and Oceanography* 17(3):327–340.

(21) Hatton IA, Heneghan RF, Bar-On YM, Galbraith ED (2021) The global ocean size spectrum from bacteria to whales. *Science Advances* 7(46):eabh3732.

(22) Kleiber M (1932) Body size and metabolism. *Hilgardia* 6(11):315–353.

(23) Andersen KH, et al. (2016) Characteristic Sizes of Life in the Oceans, from Bacteria to Whales. *Annual Review of Marine Science* 8:217–241.

(24) Pozas-Schacre C, et al. (2021) Congruent trophic pathways underpin global coral reef food webs. *Proceedings of the National Academy of Sciences* 118(39).

(25) Elton CS (1927) *Animal Ecology.* (University of Chicago Press, Chicago, IL).

(26) Nee S, Read AF, Greenwood JJD, Harvey PH (1991) The relationship between abundance and body size in British birds. *Nature* 351(6324):312–313.

(27) Brown JH, Gillooly JF, Allen AP, Savage VM, West GB (2004) Toward a Metabolic Theory of Ecology. *Ecology* 85(7):1771–1789.

(28) Trebilco R, Baum JK, Salomon AK, Dulvy NK (2013) Ecosystem ecology: size-based constraints on the pyramids of life. *Trends in Ecology & Evolution* 28(7):423–431.

(29) Fick SE, Hijmans RJ (2017) WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. *International Journal of Climatology* 37(12):4302–4315.

(30) McCann KS, Rasmussen JB, Umbanhowar J (2005) The dynamics of spatially coupled food webs. *Ecology Letters* 8(5):513–523.

(31) McCauley DJ, et al. (2018) On the prevalence and dynamics of inverted trophic pyramids and otherwise top-heavy communities. *Ecology Letters* 21(3):439–454.

(32) White EP, Ernest SKM, Kerkhoff AJ, Enquist BJ (2007) Relationships between body size and abundance in ecology. *Trends in Ecology & Evolution* 22(6):323–330.

(33) Pearl J (1995) Causal diagrams for empirical research. *Biometrika* 82(4):669–688.

(34) Edgar GJ, et al. (2020) Establishing the ecological basis for conservation of shallow marine life using Reef Life Survey. *Biological Conservation* 252:108855.

(35) Saha K, et al. (2018) The Coral Reef Temperature Anomaly Database (CoRTAD) Version 6 - Global, 4 km Sea Surface Temperature and Related Thermal Stress Metrics for 1982 to 2019. NOAA National Centers for Environmental Information. Dataset.

(36) Mora C, et al. (2011) Global Human Footprint on the Linkage between Biodiversity and Ecosystem Functioning in Reef Fishes. *PLOS Biology* 9(4):e1000606.

(37) DeVantier L, Turak E, Szava-Kovats R (2020) Species Richness and Abundance of Reef-Building Corals in the Indo-West Pacific: The Local–Regional Relation Revisited. *Frontiers in Marine Science* 7.

(38) Edgar GJ, et al. (2014) Global conservation outcomes depend on marine protected areas with five key features. *Nature* 506(7487):216–220.

(39) Lefcheck JS, et al. (2021) Species richness and identity both determine the biomass of global reef fish communities. *Nature Communications* 12(1):6875. Number: 1 Publisher: Nature Publishing Group.

(40) Bellwood DR, Hughes TP (2001) Regional-Scale Assembly Rules and Biodiversity of Coral Reefs. *Science* 292(5521):1532–1535. Publisher: American Association for the Advancement of Science.

(41) Saavedra S, Medeiros LP, AlAdwani M (2020) Structural forecasting of species persistence under changing environments. *Ecology Letters* 23(10):1511–1521.

(42) Deng J, Angulo MT, Saavedra S (2021) Generalizing game-changing species across microbial communities. *ISME Communications* 1(1):1–8.

(43) Bender EA, Case TJ, Gilpin ME (1984) Perturbation Experiments in Community Ecology: Theory and Practice. *Ecology* 65(1):1–13.

(44) Mitchell MD, Cowman PF, McCormick MI (2012) Chemical Alarm Cues Are Conserved within the Coral Reef Fish Family Pomacentridae. *PLOS ONE* 7(10):e47428.

(45) Song C, Rohr RP, Saavedra S (2018) A guideline to study the feasibility domain of multi-trophic and changing ecological communities. *Journal of Theoretical Biology* 450:30–36.

(46) Saavedra S, Bartomeus I, Godoy O, Rohr RP, Zu P (2022) Towards a system-level causative knowledge of pollinator communities. *Philosophical Transactions of the Royal Society B: Biological Sciences* 377(1853):20210159.

(47) Halford AR, Caley MJ (2009) Towards an understanding of resilience in isolated coral reefs. *Global Change Biology* 15(12):3031–3045.

(48) Small C, Nicholls RJ (2003) A Global Analysis of Human Settlement in Coastal Zones. *Journal of Coastal Research* 19(3):584–599.

(49) Hoegh-Guldberg O (1999) Climate change, coral bleaching and the future of the world's coral reefs. *Marine and Freshwater Research* 50(8):839–866. Publisher: CSIRO PUBLISHING.

(50) Temperton VM, Hobbs RJ, Nuttle T, Halle (Eds.) S (2004) *Assembly rules and restoration ecology.* (Island Press).

(51) Tabi A, Gilarranz LJ, Saavedra S (2022) Data from: Marine protected areas maintain pyramid-like structures of coral-reef fish communities, Dryad, Dataset.

(52) UNEP-WCMC, WorldFish Centre, WRI, TNC (2021) Global distribution of warm-water coral reefs, compiled from multiple sources including the Millennium Coral Reef Mapping Project. Version 4.1. Includes contributions from IMaRS-USF and IRD (2005), IMaRS-USF (2005) and Spalding et al. (2001). Cambridge (UK): UN Environment World Conservation Monitoring Centre.

(53) UNEP-WCMC and IUCN (2022) Protected Planet: The World Database on Protected Areas (WDPA) and World Database on Other Effective Area-based Conservation Measures (WD-OECM), January 2022, Cambridge, UK: UNEP-WCMC and IUCN.

(54) Center for International Earth Science Information Network CIESIN Columbia University (2020) Gridded population of the world, version 4 (gpwv4): Basic characteristics, revision 11. palisades,448 ny: Nasa socioeconomic data and applications center (sedac), (Gland, Switzerland: IUCN). Technical Report No. 26.

(55) Xiao Y, et al. (2017) Mapping the ecological networks of microbial communities. *Nature Communications* 8(1):2042.

(56) Hopf JK, Caselle JE, White JW (2021) Recruitment variability and sampling design interact to influence the detectability of protected area effects. *Ecological Applications.*

(57) Kalisch M, Mächler M, Colombo D, Maathuis MH, Bühlmann P (2012) Causal Inference Using Graphical Models with the R Package pcalg. *Journal of Statistical Software* 47:1–26.

Supporting Information

Causal graph

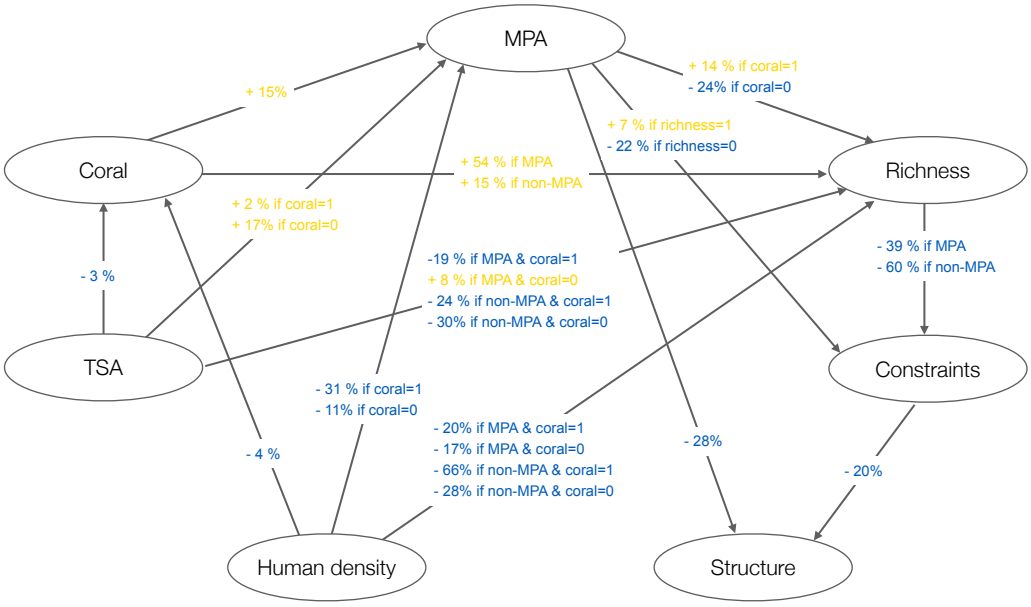

Figure S1. Causal graph. Here we show the directed acyclic causal graph that depicts the causal relationships of the studied variables. The arrows (edges) represent a direct causal effect between two variables (nodes). The percentages on the arrows show the direct average causal effects (ACE). All variables have been transformed to binary values, where a value of 1 is a data point above the median and 0 otherwise.

Distributions of variables

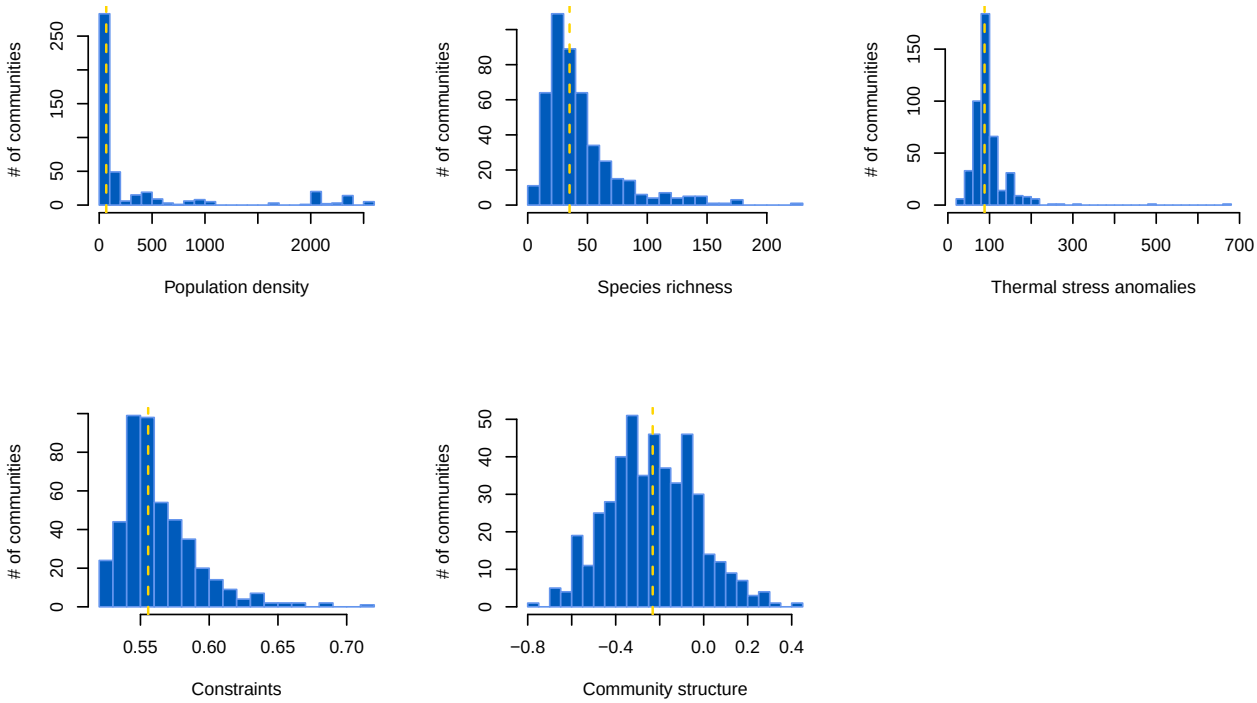

Figure S2. Distribution of the variables. Panels show the histograms of human population density, species richness, the sum of thermal stress anomalies (TSA), constraints and community structure of each communities. The yellow dashed lines show the median values of each variable.

Sampling effort

We also showed that results can be highly dependent on the number of annual sampling events. The seasonal variations in fish populations affects population dynamics, which causes temporal changes in composition (56). Therefore, we conducted a rarefaction analysis on the entire dataset to assess the effect of sampling effort on species richness at a given location and time (year). We found that reliable metrics about the communities can be produced only when the sampling effort was more than one time per year. As our rarefaction analysis concluded that increased sampling effort (> 1 per year) provides more reliable description of the composition of communities.

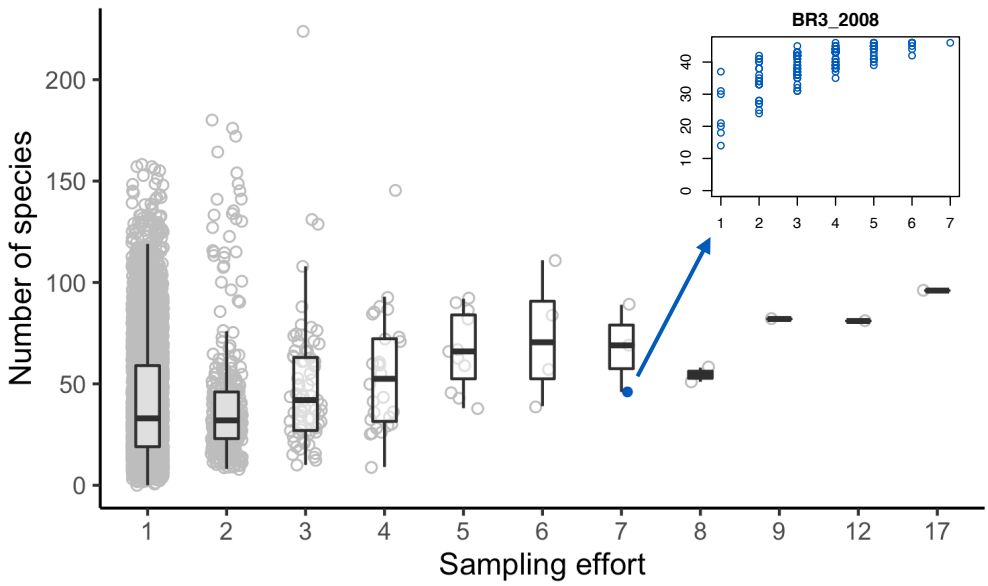

Figure S3. Sampling effort per year and an example of a rarefaction curve. The grey circles represent communities aggregated in a given location across a year. The majority of communities were sampled once per year, only a small fraction of the aggregated communities were sampled more than once in a year (sampling effort > 1). For communities sampled more than once in a given year, we conducted a rarefaction analysis to estimate the effect of sampling effort on species richness by resampling communities and then plotting the number of species in each constructed community against sampling effort (an example shown in the top right panel).

D-separations in our causal model

Testing for conditional independence is essential in causal inference. We used the G^2 test from the pcalg R package (57) for identifying eight d-separation in our causal graph.

	Variables	Conditions	G^2 test (p-value)
1	Coral ↔ Interaction constraints	Species richness, MPA	0.561
2	Coral ↔ Community structure	Interaction constraints, MPA	0.101
3	Human density ↔ Interaction constraints	Species richness, MPA	0.049
4	Human density ↔ Community structure	Species richness, MPA	0.103
5	Human density ↔ TSA	-	0.781
6	TSA ↔ Interaction constraints	Species richness, MPA	0.078
7	TSA ↔ Community structure	Interaction constraints, MPA	0.089
8	Species richness ↔ Community structure	Interaction constraints, MPA	0.308

Table S1: We tested for the conditional independence for the direct paths using G^2 test. The following direct links in our causal model are considered independent of each other (i.e. the p-values $> \alpha$, where we set $\alpha = 0.01$).

We also tested, whether there is a genuine effect between MPA and community structure. Following our causal graph, for the genuine effect: (1) X =MPAs have to be statistically dependent on Y =community structure across all contexts (Table S2, row 1). (2) One can find one Z =potential cause (e.g., TSA) of MPAs. This is true if Z and X are statistically dependent across all contexts (Table S2, row 2). We can find one W =variable (e.g., constraints) and one S_1 =context (e.g. species richness) such that Z and W are statistically independent under the context (Table S2, row 3). (3) Z and Y have to be statistically dependent under one S_2 =context (e.g., interaction constraints, Table S2, row 4), and they are statistically independent under the context of $X \cup S_2$ (Table S2, row 5).

	Variables	Conditions	G^2 test (p-value)
1	MPA ↔ Community structure	Human density, Interaction constraints, Coral, TSA	0
2	TSA ↔ MPA	Human density, Coral	0
	TSA ↔ Interaction constraints	Species richness	0.115
3	TSA ↔ Community structure	Interaction constraints	0.007
	TSA ↔ Community structure	MPA, Interaction constraints	0.089

Table S2: We tested for the genuine effect between MPA and community structure using G^2 test. The following links are considered independent of each other if the p-values > 0.01.