The marine fish food web is globally connected

Camille Albouy [1]*, Philippe Archambault [2], Ward Appeltans[3], Miguel B. Araújo [4,5,6], David Beauchesne[7], Kevin Cazelles [8], Alyssa R. Cirtwill [9,20], Marie-Josée Fortin[10], Nuria Galiana[11], Shawn J. Leroux[12], Loïc Pellissier[13,14], Timothée Poisot[15,16], Daniel B. Stouffer [9], Spencer A. Wood[17,18] and Dominique Gravel[15,19]

The productivity of marine ecosystems and the services they provide to humans are largely dependent on complex interactions between prey and predators. These are embedded in a diverse network of trophic interactions, resulting in a cascade of events following perturbations such as species extinction. The sheer scale of oceans, however, precludes the characterization of marine feeding networks through denovo sampling. This effort ought instead to rely on a combination of extensive data and inference. Here we investigate how the distribution of trophic interactions at the global scale shapes the marine fish food web structure. We hypothesize that the heterogeneous distribution of species ranges in biogeographic regions should concentrate interactions in the warmest areas and within species groups. We find that the inferred global metaweb of marine fish—that is, all possible potential feeding links between co-occurring species—is highly connected geographically with a low degree of spatial modularity. Metrics of network structure correlate with sea surface temperature and tend to peak towards the tropics. In contrast to open-water communities, coastal food webs have greater interaction redundancy, which may confer robustness to species extinction. Our results suggest that marine ecosystems are connected yet display some resistance to perturbations because of high robustness at most locations.

The ~15,000 living marine fish species[1] known globally constitute a major source of protein for billions of humans. Balancing fish conservation with human needs is difficult because a myriad of direct and indirect relationships among species make the dynamics of diverse food webs challenging to understand and predict[2,3]. For example, predators require the occurrence of prey species for their existence and, consequently, the disappearance of a species at the bottom of the food chain can trigger cascades of extinction right up to the top predators[4]. However, disappearance of species at any position in the food web can lead to community disassembly and profound transformation in ecosystem productivity[5]. The consequences of loss of species, due for example to human activities[6], for the functioning of marine ecosystems will depend on the robustness of the network of trophic interactions[7]. As ecological interactions are critical to community functioning, any attempt to understand marine fish communities must account for their structure and spatiotemporal distribution.

Because most species have patchy, uneven and not entirely overlapping distributions, interactions among them may vary across time, regions and ecosystems[8]. Such variation can affect trajectories of network reorganizations after disturbances[9]. Mapping the structure of food webs in space could thus help in anticipating the

consequences of global changes on biodiversity. However, the time and expense required to empirically document the spatial variability of millions of interactions is prohibitive[10]. Consequently, although the ocean covers approximately 71% of the Earth's surface, we do not know how food webs across the global marine realm are organized. Rather than relying solely on empirical methods, we can infer the probability of the occurrence of a trophic interaction based on measurable information on species distributions, body mass and local environmental conditions that are known to influence predation. Given empirical data on the occurrence of interactions among species under varying conditions, it is feasible then to estimate the probability of interaction among species in other locations[11]. The resulting structure is a 'metaweb', describing potential interactions based on extant knowledge[12] (Fig. 1).

In this study, we investigate how the distribution of trophic interactions in marine fishes influences the structure of their ecological networks globally. We expect the latitudinal gradient in species richness[13] and sea surface temperature (SST) to influence the structure of marine fish ecological networks and, consequently, metrics such as connectance and mean trophic level should peak around the tropics. Then, we expect that the heterogeneous spatial distribution of species across biogeographic regions will concentrate interactions

[1]IFREMER, unité Ecologie et Modèles pour l'Halieutique, Nantes, France. [2]Departement de biologie, Faculté des sciences et de génie, Université Laval, Québec, Québec, Canada. [3]Ocean Biogeographic Information System, IODE, Intergovernmental Oceanographic Commission, UNESCO, Ostend, Belgium. [4]Departmento de Biogeografía y Cambio Global, Museo Nacional de Ciencias Naturales, Consejo Superior de Investigaciones Científicas, Madrid, Spain. [5]InBio/Centro de Investigação em Biodiversidade e Recursos Genéticos, Universidade de Évora, Évora, Portugal. [6]Center for Macroecology, Evolution, and Climate, Natural History Museum of Denmark, University of Copenhagen, Copenhagen, Denmark. [7]Institut des sciences de la mer, Université du Québec à Rimouski, Rimouski, Québec, Canada. [8]Department of Integrative Biology, University of Guelph, Guelph, Ontario, Canada. [9]Centre for Integrative Ecology, School of Biological Sciences, University of Canterbury, Christchurch, New Zealand. [10]Department of Ecology and Evolutionary Biology, University of Toronto, Toronto, Ontario, Canada. [11]Centre for Biodiversity Theory and Modelling Station d'Ecologie Théorique et Expérimentale du CNRS, Moulis, France. [12]Department of Biology, Memorial University, St. John's, Newfoundland, Canada. [13]Landscape Ecology, Institute of Terrestrial Ecosystems, Department of Environmental Systems Science, ETH Zürich, Zürich, Switzerland. [14]Swiss Federal Research Institute WSL, Birmensdorf, Switzerland. [15]Département des Sciences Biologiques, Université de Montréal, Montréal, Québec, Canada. [16]Québec Centre for Biodiversity Sciences, McGill University, Montréal, Québec, Canada. [17]EarthLab, University of Washington, Seattle, WA, USA. [18]eScience Institute, University of Washington, Seattle, WA, USA. [19]Département de Biologie, Universite de Sherbrooke, Sherbrooke, Québec, Canada. [20]Present address: Department of Ecology, Evolution, and Plant SciencesPhysics, Chemistry and Biology (IFM), Linköping Stockholm University, Stockholm, Sweden. *e-mail: albouycamille@gmail.com
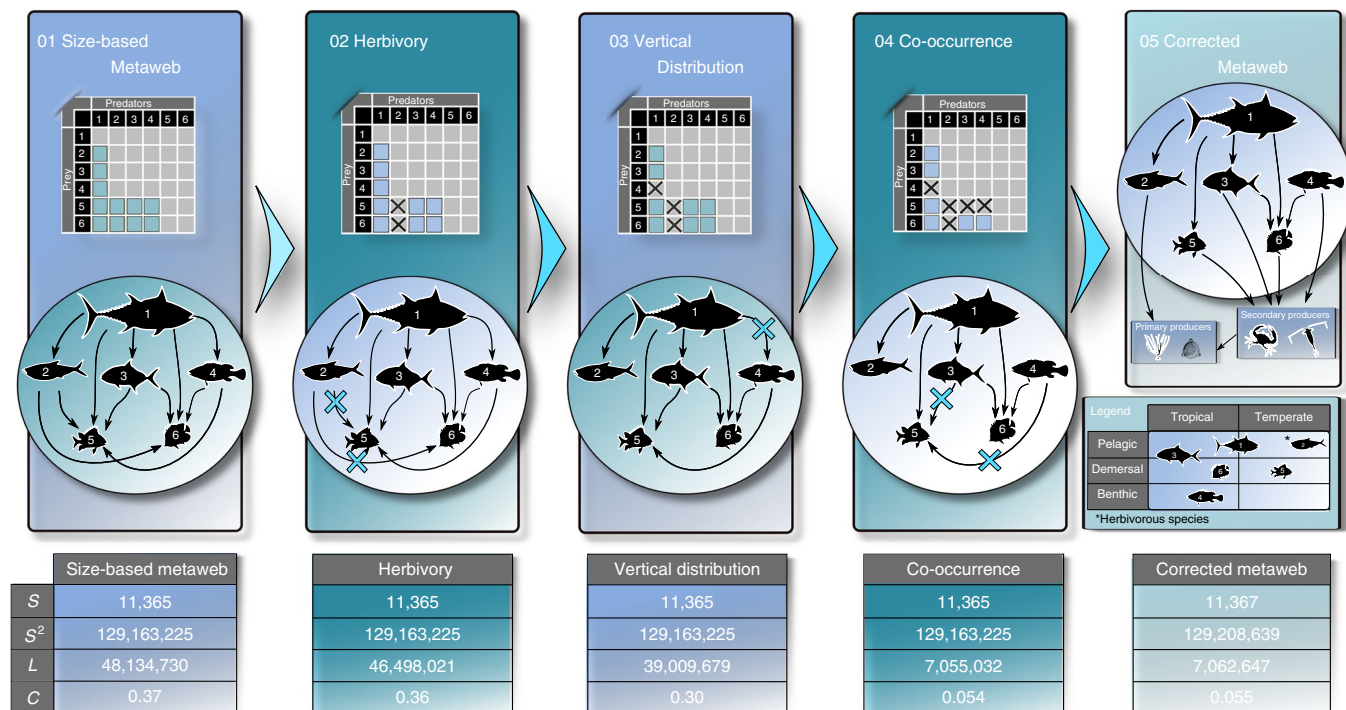
	Size-based metaweb	Herbivory	Vertical distribution	Co-occurrence	Corrected metaweb
S	11,365	11,365	11,365	11,365	11,367
S^2	129,163,225	129,163,225	129,163,225	129,163,225	129,208,639
L	48,134,730	46,498,021	39,009,679	7,055,032	7,062,647
C	0.37	0.36	0.30	0.054	0.055

Fig. 1 | Construction of the metaweb. The size-based metaweb (step 1) is obtained by fitting an empirical allometric niche model. Links from herbivores are subsequently removed (step 2), as well as those between species that are found at different positions in the water column (step 3); for example, a pelagic fish cannot eat a benthic one (species 1 versus species 4). Interactions between species that do not co-occur based on their range are removed (step 4); for example, a tropical fish cannot eat a temperate fish (species 3 versus species 4). Finally, the primary and secondary producers are linked together and to each appropriate fish (step 5). S, observed number of species; S^2, number of potential links in the metaweb; L, number of realized links; C, estimated connectance (L/S^2).

within groups of species and that the fish metaweb will exhibit modules corresponding to these regions. Therefore, modules should limit propagation of perturbations and the robustness of marine food webs to species extinctions should be strongly localized. The robustness of networks may also vary along environmental gradients in correspondence to the global distribution of marine fish diversity (for example, SST)[13], and is also dependent on the internal network structure such as connectance[14]. We investigate these expectations by reconstructing all possible predator–prey interactions among 11,365 marine fish species distributed across the globe. We extracted local food webs across the globe at 1° × 1° resolution and computed network structure metrics and robustness that were related to environmental gradients. We then mapped the structure of ecological networks and identified those areas with the greatest complexity of interactions and those most sensitive to extinctions.

The metaweb was assembled based on life history traits documented in FishBase[15] (body size, bathymetry and diet), and on species co-occurrence data for 75% of the known marine fish species ($n = 11,365$) from the Ocean Biogeographic Information System (OBIS)[16] database (see Methods and Supplementary Fig. 1). These data were assigned to a spatial location in a grid system covering the global oceans at 1° × 1° resolution. Trophic interactions were predicted using a probabilistic model of trait-matching calibrated on a published dataset of >34,000 observed predator–prey interactions[17], and evaluated using 999-fold cross-validation methods by retaining only unique combinations of predator–prey body sizes in the dataset (mean Boyce index = 0.55, s.d. = 0.08, $n = 999$; Supplementary Fig. 2). The metaweb was further trimmed using known information about species habitat requirements (position in the water column) and knowledge of fish diet. The final metaweb has 7.062,647 potential interactions (Fig. 1) and a connectance (realized links/potential

links) of 0.055 (Fig. 2e,f), a value that falls within the range of most reported food webs[14,18]. Moreover, the metaweb is well connected considering the path length distribution, with most species pairs separated by two or fewer links (Fig. 3a).

Results

The occurrence of network subgroups of species can be measured as the difference between observed and expected intra-group interactions, divided by the total number of interactions[19]. This measure is called modularity (Q), positive values of Q indicating that interactions occur predominantly within groups and negative values indicating that interactions are more frequent between than within groups[19]. Generally this modular organization can have stabilizing effects[19] and enhances the persistence of ecological networks[20]. Sub-biogeochemical provinces (Supplementary Table 1) are potentially strong candidate locations for modules in the metaweb, as these are defined by environmental conditions driving specific assemblages, especially among low trophic levels[21,22]. To test this, we extracted all the species present in each pair of biogeochemical provinces (BGCPs) according to the range maps obtained from the OBIS occurrences, then crossed this information with the metaweb to acquire the links between species in each pair of BGCPs. Finally, we summed the number of links between provinces. On this new matrix we calculated (1) the connectance among provinces and (2) the modularity of the network using the Walktrap community algorithm that maximizes the modularity index (Q)

Contrary to that expectation, we found that all 56 BGCPs are equally connected with each other (mean connectance = 0.047, s.d. = 0.005, $n = 56$; Fig. 2). The modularity is null when groups are forced to belong to BGCPs, while the maximum modularity ($Q = 1.42 \times 10^{-7}$) indicates that no other cluster of interactions
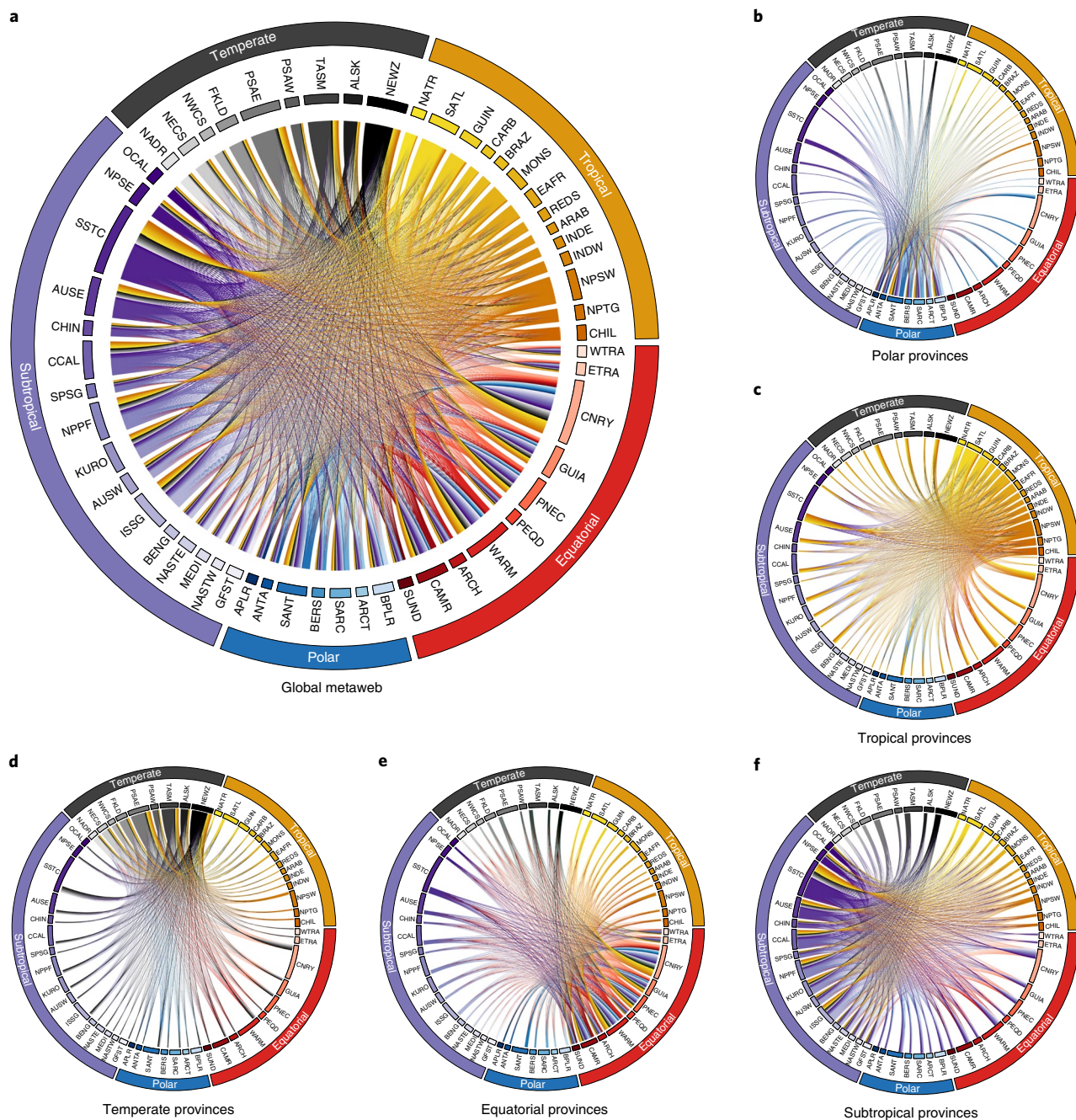
Fig. 2 | Circular representation of inferred trophic interactions among species. a–f, Species are positioned according to their primary areas of occurrence in the climatic biomes and BGCPs. The outer circles represent the climatic biomes while the inner circles represent the 56 BGCPs (see Supplementary Table 2 for details). The size of each province's rectangle is proportional to the number of links between provinces. **a**, Trophic interactions among all provinces for the corrected metaweb (metaweb based on body size, bathymetry, diet and species co-occurrence data for 75% of known marine fish species). **b–f**, Subset of trophic interactions among polar provinces (**b**), tropical provinces (**c**), temperate provinces (**d**), equatorial provinces (**e**), subtropical provinces (**f**) and others.

exists in the metaweb. This modularity value was lower than expected by chance, based on a benchmark distribution of 999 modularity values calculated on randomized metawebs presenting the same connectance and species richness as the observed metaweb (Z-score $= -20.6$, $P < 0.0001$). Many large-ranged species are distributed across several BGCPs, increasing the connectivity of the global network. Species with the largest ranges (up to $423,000\,km^2$; third quartile of the range's distribution) frequently overlap, spanning on average approximately 18 BGCPs (mean $= 17.75$, s.d. $= 9.28$, $n = 56$). In addition, one-third of the

species pairs co-occur in at least one $1° \times 1°$ cell (connectance of the co-occurrence matrix $= 0.29$). The path length distribution and the lack of modularity confer global connectivity to the metaweb and should increase the propagation of disturbances across species and provinces[23]. Annual migrations and movement of individuals within their range will connect species across provinces, even if they do not interact directly together. However, global connectivity may also increase the opportunity for interaction redundancy of species at higher trophic levels and may therefore promote robustness to species extinctions.
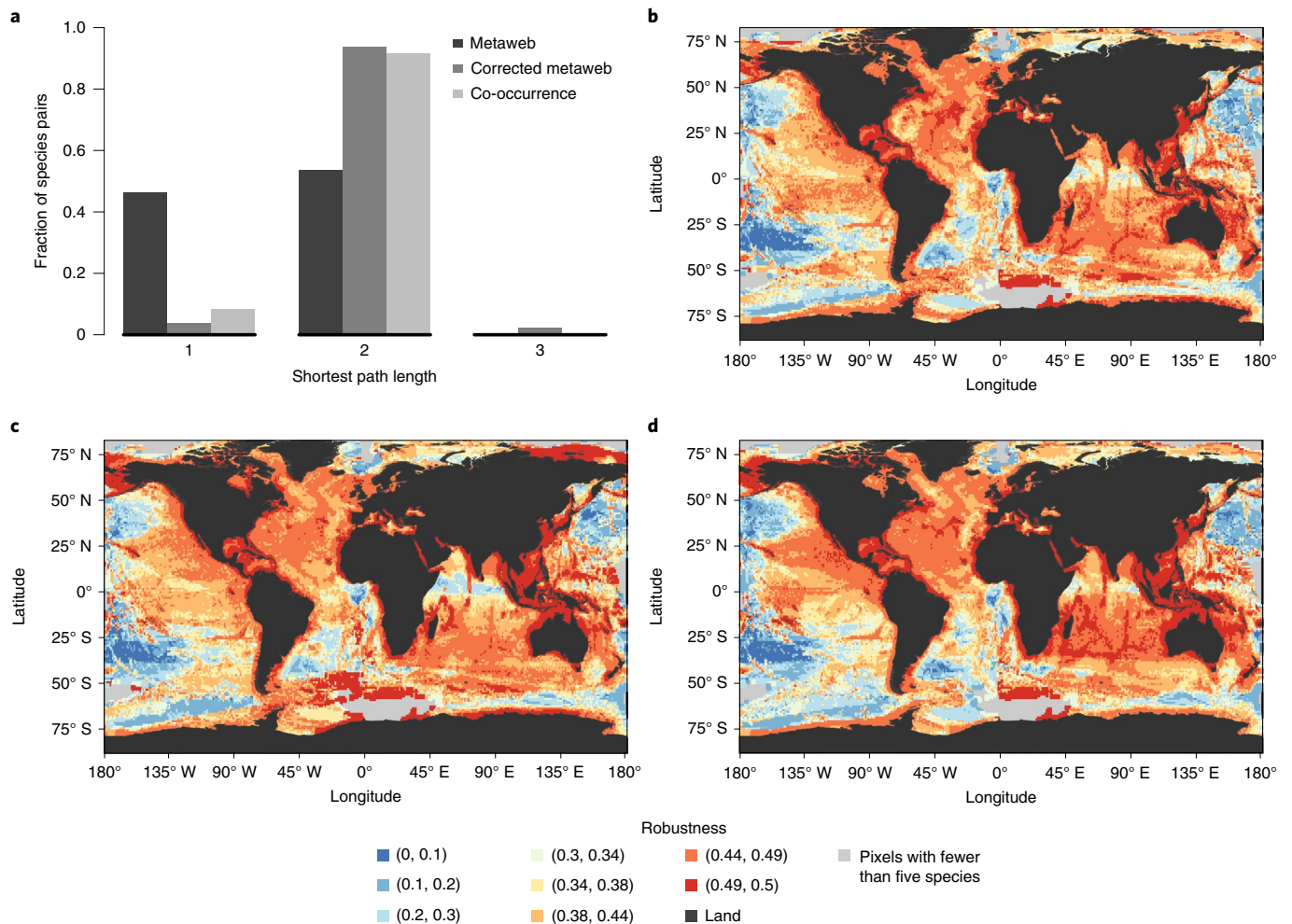
Fig. 3 | The global connectivity of the metaweb and local webs provides robustness against species extinctions. a, Distribution of shortest path length in the metaweb based only on body size (black bars), the corrected metaweb based on body size, bathymetry, diet and species co-occurrence (dark grey bars) and that based only in the matrix of species co-occurrence (light grey bars). **b,** Spatial distribution of local food web robustness to species extinctions ordered by decreasing species body size. **c,d,** Spatial distribution of local food web robustness to species extinctions ordered by decreasing species range size (**c**) and by increasing species range size (**d**). The random scenario is presented in Supplementary Fig. 9, and the correlation among scenarios in Supplementary Table 3.

While the global connectivity of the metaweb suggests that marine food webs are robust, or persistent despite extinctions, we assess whether this is also true of local food webs. Within each of the grid cells, we measured structural properties of food webs and explored their geographical distribution (Fig. 4 and Supplementary Fig. 3). We found that the mean trophic level (Fig. 4b) is correlated positively with species richness (Spearman's rank correlation $\rho_S = 0.94$, $P < 0.001$), and decreased from the equator to the poles (Fig. 4c,d). Connectance was almost constant between $-40°$ and $40°$ of latitude (mean = 0.047, s.d. = 0.048, $n = 21053$; Fig. 4e,f). Outside this latitudinal range, connectance increased sharply because of a drop in species richness ($\rho_S = -0.61$, $P < 0.001$), especially near the poles[24]. Because diversity in fish species usually increases with SST[13], we therefore looked at the relationship between network properties and SST (Supplementary Table 2 and Supplementary Fig. 4). At the local level we found a significant correlation between SST and species richness ($\rho_S = 0.55$, $P < 0.001$), number of herbivorous species (partial Spearman correlation (PSC) controlled by species richness = 0.17, $P < 0.001$), number of links (PSC; $\rho_S = 0.52$, $P < 0.001$), food chain length (PSC; $\rho_S = 0.29$, $P < 0.001$), vulnerability (mean number of predators per prey species; PSC; $\rho_S = -0.42$, $P < 0.001$; Supplementary Figs. 3 and 5) and

connectance (PSC; $\rho_S = -0.36$, $P < 0.001$). These results indicate that more energy at the bottom of the food chain supports food webs with higher species richness, more links and more elongated trophic chains[25]. Overall, spatial variation in food web structure is strongly related to the underlying latitudinal gradients in species richness, the proportion of herbivorous fish species in the local food web (Supplementary Fig. 6) and the average and variance in local body size (Supplementary Fig. 7).

Within each cell, we characterized the network robustness to extinctions by sequentially removing species according to four different scenarios: (1) the largest-bodied species that are fished intensively and may disappear first; (2) species with the widest geographic range and that predominantly connect the local networks spatially; (3) species with the most restricted geographic range, with smaller population size and that are more vulnerable to perturbation; and (4) randomly selected species for comparison to other scenarios (99 random deletion sequences initiated for each web; Supplementary Fig. 8). For each sequential removal, we quantified the number of secondary extinctions occurring after loss by a non-basal species of all of its prey. We then measured robustness as the fraction of species that had to be removed to result in a loss of ≥50% of species[5]. We found that marine fish food webs were
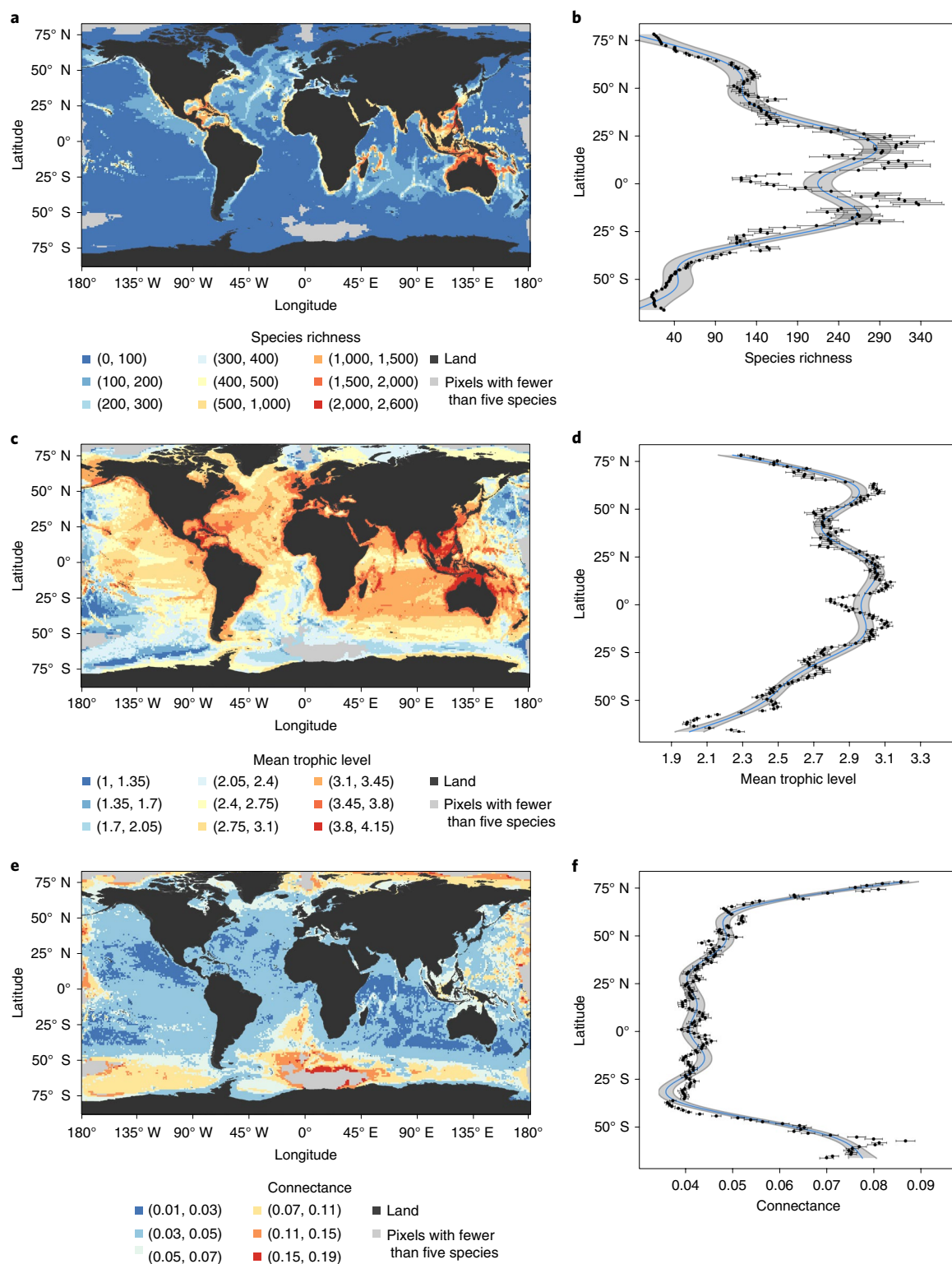
Fig. 4 | Spatial distribution of food web properties globally. a–f, The maps represent the distribution of species richness (**a**), trophic levels (**c**) and connectance (**e**) for 11,365 fish species on a 1° × 1° grid, while the charts represent species richness (**b**), mean trophic level index (**d**) and connectance (**f**) averaged across latitude and smoothed by a generalized addition model function (blue line). These models were fitted between species richness (**b**, adjusted $R^2 = 0.87$, estimated degrees of freedom (EDF) = 8.84), mean trophic level (**d**, adjusted $R^2 = 0.90$, EDF = 8.83), connectance (**f**, adjusted $R^2 = 0.94$, EDF = 9) and latitude, respectively, within the 95% confidence limits (grey shading). The black bars on the charts represent s.e.m.

generally robust to species loss, with no detectable difference in robustness from the tropics to the poles (Fig. 3b–d), and among the four sequential removal scenarios (Supplementary Table 3).

Rather, the strongest difference we observed was between open ocean and coastal areas (Wilcoxon rank test; mean coastal robustness, 0.45; mean open water robustness, 0.38; $P < 0.05$; Fig. 3b,c

and Supplementary Table 4), even when accounting for the effect of species richness (Supplementary Table 5). Coastal area was defined as all cells of depth 0–200 m. This difference was confirmed by the inverse correlation between all robustness scenarios and distance to land (for the body size scenario; PSC; $\rho_S = -0.24$; $P < 0.001$) and the first split observed in the regression tree between body size scenario robustness and environmental data (Supplementary Fig. 9a). This result can be explained by a combination of environmental factors and internal network characteristics. The coastal robustness of networks is mainly associated with species richness and SST (Supplementary Fig. 9b; $R^2 = 0.61$). However, when controlling by species richness, the coastal robustness of networks was related to SST whereas the open water robustness was related to primary productivity (Supplementary Fig. 9b; $R^2 = 0.21$). Moreover, when considering network proprieties, the difference between coastal and open water robustness can be explained by shorter chain length, shortest path length and difference in connectance (Supplementary Fig. 9c), all probably associated with differences in species richness between open ocean and coastal areas (Supplementary Fig. 9d). All of these observations led to the conclusion that robustness relates to SST and connectance, even though species richness also plays a major role. Overall, species are highly connected within local food webs such that in 67% of local food webs, 95% of the species pairs are separated by no more than three links (Supplementary Fig. 3i). That most species pairs are separated by so few links means that, topologically, top predators are close to the bottom of the food web. This implies that, when intermediate or even basal species are lost, there will be alternative pathways for energy to move up marine food chains and that cascading extinctions should be rare.

Discussion

Our simulations of species extinction emphasize the spatial heterogeneity of robustness in marine food webs between coastal and open-water areas. The elevated connectance observed globally (across the metaweb) and locally (in each grid cell) indicates that a local perturbation, such as pollution or overfishing, causing population collapse may impact several trophic levels within the food web relatively quickly[23,26]. Nevertheless, disturbances on well-connected nodes are less likely to have a strong impact on neighbouring nodes than those on nodes with only a few links[26]. Consequently, the structure of marine fish food webs is expected to be able to absorb perturbations up to a critical point. An analysis of how major marine perturbations, such as commercial fishing, may affect the robustness of marine fish local and metawebs is required to better understand the buffering capacity of the food web structure[27]. For example, in ref. [28] it was demonstrated that fishing affects the Caribbean coastal marine food web structure by eroding food web robustness and persistence. Simulated perturbations may represent a simplification of real perturbations observed in nature. Here, under the assumption that because fish species rely equally on all available prey items, we considered only the structure of the food web and not interaction strength. Our interpretation of elevated robustness would be reinforced by integration of invertebrate and vertebrate taxa other than fish in the food web, since these could increase connectivity at the global scale and robustness at the local scale. Other processes such as movement, dispersal or, more broadly, migration, which are fundamental to marine systems, could act to increase robustness against species extinctions. Indeed, by linking and creating more complex food webs, migration could have a stabilizing effect[29] and enhance food web robustness[30]. Migration is implicitly considered in our modelling approach because we looked at the spatial overlap between ranges, thereby connecting species across populations. Range dynamics from colonization–extinction processes is not, however, considered a phenomenon that might contribute to network rewiring and a further increase in robustness.

While there has been an intense research effort in macro-ecology to determine why biodiversity peaks in the tropics, the distribution of biotic interactions along environmental gradients in both terrestrial and marine environments, neglected to date, is now increasingly studied[31,32]. Indeed, a community is more than just a list of species because its dynamics and stability depend critically on the myriad interactions shaping it[10]. The next challenge faced by ecologists, if they wish to explore variation in ecosystem services, is therefore to move beyond description and prediction of the biodiversity structuration to consider the relationships among its components[33]. Here, we propose a trait-based methodology to circumvent data limitations and to estimate the structure of fish food webs. While the role of temperature in driving biological rates and biodiversity distribution is recognized[34], a full understanding of the large-scale variation in trophic regulation[35] requires that we also document how it influences key properties of the network of interactions linking species.

Our approach, estimating local fish-to-fish interaction networks across the globe, highlights new challenges for empirical research in macro-ecology. Food web studies are usually performed at either a single or several locations because of the amount of work required to observe large numbers of species and their interactions. As a consequence, such local data are too scarce for large-scale analyses of the distribution of network properties such as that conducted here. We therefore had to rely exclusively on fish data because of their wide availability, while acknowledging the need to expand eventually to invertebrate and vertebrate taxa other than fish, and on models to simulate where and which species will interact. We also focused on binary interaction matrices to limit model assumptions and to acquire a first approximation of their spatial variation. Network topology has been shown to be more important than distribution of interaction strength in estimation of stability[36], lending support to our approach. Nevertheless, integration of the interaction strength in our approach will facilitate progression from study of the food web structure to ecosystem functioning and the provision of services by ecosystems.

We performed validation at the lowest level of our hierarchical modelling approach, first by cross-validating the allometric model of interactions fitted on the ~34,000 interaction records and second by cross-checking distributions derived from the OBIS occurrences with published range maps and check-lists of well-known and tropical species (see Methods for further details). Both procedures demonstrated that we had good model performance (see Methods), giving us confidence in the reliability of our predictions. Several other marine food web datasets available in different regions of the world (for example, refs. [37,38]) could theoretically be used to test our predictions. Nevertheless, after verification these datasets were deemed unsuitable. They were either incompletely sampled (nodes often representing tropho-species or guilds) or had been collected using different methodologies (from expert knowledge to gut contents). The various maps we provide should also be interpreted carefully: these are not meant to predict the structure of local food webs (although we have used the method successfully for this purpose[39]), but rather the expected global variation in their properties. These maps are interpretations of extensive species richness variation combined with the distribution of body size and herbivory, and it is at this large scale that validation testing should be performed by analysis of the global tendencies of trophic indicators. Our study attempts to reveal large-scale variation in marine food web structure and, as such, to encourage testing of novel predictions in macro-ecology with a new generation of food web data. Recent progress in the development of DNA meta-barcoding techniques and their application to documenting of trophic interactions[40] is highly promising, and suggests that such research will progress in the near future.

We show that the integration of open global databases, new computing tools and statistical approaches facilitates the mapping of trophic interactions at the global level. This new approach to macro-ecological research overcomes the limitations of sampling by reliance on existing data. The sampling effort required to estimate species interactions is orders of magnitude larger than that required to detect species presence and, while not perfect, synthetic datasets provide opportunities to generate testable predictions about the outcomes of future large-scale management scenarios. Fish are an integral component of the marine food web, and some countries are advocating for a holistic, ecosystem-based approach to sustainable ocean biodiversity management and conservation (for example, in Europe[41]). Such a commitment requires not only a deep understanding of the distribution of biodiversity, but also of biotic interactions, in inferring its functioning. The observation that fish food webs are globally connected further illustrates the global dimension of marine biodiversity. Our results should thus strengthen the importance of global negotiations, such as those currently under way aiming towards a new United Nations treaty to conserve and use sustainably marine biological resources of areas beyond national jurisdictions.

Methods

Data preparation. Species data were obtained from OBIS (http://www.iobis.org) on 27 August 2014, representing the best available dataset for a global study of marine fish food webs. We inventoried 16,238,200 occurrence records from 34,883 entries. We cleaned the data by identifying the synonyms, misspellings and rare species (only one occurrence) and by restricting them to species present in the marine environment according to FishBase[15]. Synonyms were converted to accepted names. This resulted in a set of 13,916,517 occurrences for 11,365 fish species around the world. We considered every occurrence available in the OBIS database spanning the period 1826–2013, as occurrences before 1950 represent only 0.28% of the entire dataset. To counteract certain known biases in OBIS data (for example, not all species/regions are equally represented), we reconstructed distribution maps for each species, defined as the convex polygon surrounding the area where each species was observed (for details see Supplementary Fig. 10). The resulting polygon was divided into four parts across the world to integrate possible discontinuity between the two hemispheres and the Atlantic and Pacific Oceans. We then refined each species distribution map by removing areas where maximal depths fell outside the minimum or maximum known depth range of the species[15]. Final distribution maps of 500 species randomly chosen were checked visually and reviewed by the authors according to their expertise. We also compared the distribution of ~3,500 tropical species present in this database to the Gaspar database frequently used in reef fish analyses[42]. Finally, we aggregated fish distributions on a 1°-resolution grid covering all oceans. This scale allows computational feasibility and reproducibility of all analysis conducted here. All the data are freely available (https://doi.org/10.6084/m9.figshare.7034789). We compiled the following environmental descriptors: distance to land, SST, sea surface salinity and primary productivity. Environmental descriptors were compiled from different sources and resolutions, and were fit to our grid system (Supplementary Table 2).

Data analyses. We used a global interaction dataset derived from ref. [17] to calibrate a model of trophic interactions[43] using the log of observed body size for predators (M_{pred}) and prey (M_{prey}). This dataset comprises 34,931 marine predator–prey interactions from 27 locations covering a wide range of environmental conditions, from the tropics to the poles, for 93 predator species of size 0.3–309.69 cm and 174 prey species of size 4.16 μm to 122.66 cm. In the database, 226 predator–prey pairs were identified. Interactions were compiled from published literature and, if predator or prey length was not measured in the original study, that was calculated using length–mass relationships[17]. We inferred the probability of interactions using a trait-matching function based on the niche model for food web structure[44], where the main niche axis is log body size. The log body size of the predator determines its optimum and the range of its niche, while log prey size determines its niche position[45]. We consider a Gaussian function to represent the probability of an interaction given the size of the predator (pred) and prey:

$$P(L_{pred,prey} = 1 | M_{pred}, M_{prey}) = \exp \frac{-(\alpha_0 + \alpha_1 \times M_{pred} - M_{prey})^2}{2(\beta_0 + \beta_1 \times M_{pred})^2}$$

where α_0, α_1, β_0 and β_1 are fitted parameters describing the relationship between predator size, optimal prey size and the range of the feeding niche. The model parameters were estimated by simulated annealing. To evaluate the model, the dataset provided in ref. [17] was cleaned by retaining only unique combinations of

predator/prey body sizes and then splitting it into two parts—70% of the dataset was randomly chosen to calibrate the model while the remaining 30% was used to evaluate it. We calculated the Boyce index to evaluate the quality of the model (mean Boyce index = 0.55, s.d. = 0.08, $n = 999$; Supplementary Fig. 2). This index varies from −1 to +1, with positive values indicating a model whose predictions are consistent with the presence of interactions in the evaluation dataset; values close to zero mean that the model is not different from a null model, while negative values indicate an incorrect model. This cross-validation procedure was repeated 999 times. Once the model was calibrated and evaluated, it was applied to the 11,365 fish species using common body sizes extrapolated from the linear relation between common and maximum body size. Maximum body size data were extracted from FishBase[15]. We calculated the median body size among congeneric species ($n = 1,509$) to fill in missing data. To correct this size-based metaweb, we first removed links when large herbivorous fish eat smaller fish. Second, we removed links between fish that do not co-occur vertically in the water column and then we deleted interactions between fish that do not co-occur spatially (no range overlap). Finally, we considered the primary and secondary producers as two compartments that we combined to each other and to each appropriated fish species that feed on these resources (Supplementary Fig. 1). We hypothesized that both primary and secondary producers were distributed around the global ocean.

We used the 56 BGCPs established in 2013 (ref. [46]) to search for spatial structure in the metaweb. The BGCPs were improved compared to the previous hand-made Longhurst BGCPs[47] by using a non-parametric statistical methodology based on four environmental parameters (bathymetry, chlorophyll-a concentration, SST and sea surface salinity) for the period 1997–2007. These parameters directly affect the species abundance of lower trophic levels[21]. We also tested whether modules of tightly interacting species were associated with the BGCPs. To do this, we extracted all species present in each pair of provinces according to range maps obtained from OBIS occurrences, then crossed this information with the metaweb to determine the links between species in each pair of provinces. Finally we summed the number of links between regions to create a provincial-based metaweb. On this new matrix, we calculated (1) the connectance between provinces as the number of links between pairs of provinces divided by the number of potential links between pairs of provinces, and (2) the modularity of the network using the Walktrap community algorithm that maximizes the modularity index (Q)[48]. This index denotes the proportion of links falling within modules minus the expected proportion based on an equivalent network where links were placed at random. Modularity has a value between 0 and 1, with 1 corresponding to a perfectly modular matrix.

Food web descriptors. We calculated descriptors of local food web structure for fish assemblages[12,44,49] for each cell. We used the number of species (S), the number of actual links (L) and the fraction of all possible links (S²) that are realized in the network, called connectance (L/S²). These indices have been highlighted as important predictors of population stability and community structure[49]. We also calculated the prey-averaged trophic level of each species as one plus the mean trophic level of all of the species resources[44], and the associated omnivory index expressed as the trophic level variance of a consumer's prey. For each food web ($n = 43,041$), we simulated species loss by sequentially removing species using one of four criteria: removal of (1) the largest-bodied species, (2) species with the widest range, (3) species with the most restricted range and (4) randomly chosen species (99 random deletion sequences initiated for each web). We calculated the total trophic connections ('degree') for each species and computed the number of potential secondary extinctions. Secondary extinction occurs when a non-basal species loses all of its prey items. We then quantified the robustness of food webs to species loss by measuring the fraction of species that had to be removed to result in a total loss of ≥50% of the species (that is, primary species removal plus secondary extinctions). Maximum possible robustness is 0.50 and minimum is 1/S (refs. [14,50]). All trophic indicators were mapped and the coastline was defined using the data from ref. [51].

To examine whether the robustness of networks (1) varies along environmental gradients and (2) is dependent on the internal network structure, we performed four different regression trees[52]. Before running the regression trees, we checked for collinearity among predictors using the variance inflation factor (VIF)[53]. A VIF without collinearity has a value of 1, while VIF values >3 are indicative of collinearity issues. We first explored the effect of environmental factors (primary productivity, distance from land, SST, salinity) on the robustness of networks with and without inclusion of species richness in the regression tree models. The variable distance from land was transformed into a categorical variable—that is, cells with a depth of 0–200 m were identified as coastal and the remainder as open ocean. Collinearity among variables was found to be low (VIF values: species richness = 1.27, salinity = 1.38, primary productivity = 1.24, SST = 1.42). In the same way, we explored the effect of internal structure networks. After performing VIF analyses we selected four indicators that present low VIF values: the number of links (1.28), chain length (1.95), the modularity (2.12) and the shortest path length (2.3). We substituted the number of links by species richness, because these variables are highly correlated. To avoid overfitting, we selected the number of splits in each model as a compromise between the best R^2 value and the minimum number of splits.

References

1. Nelson, J. S. *Fishes of the World* 4th edn (John Wiley & Sons, 2006).
2. Allesina, S., Bodini, A. & Pascual, M. Functional links and robustness in food webs. *Phil. Trans. R. Soc. B* **364**, 1701–1709 (2009).
3. Worm, B. et al. Impacts of biodiversity loss on ocean ecosystem services. *Science* **314**, 787–790 (2006).
4. Newman, E. I. *Applied Ecology and Environmental Management* (Blackwell Science, 2000).
5. Dunne, J. A., Williams, R. J. & Martinez, N. D. Network structure and robustness of marine food webs. *Mar. Ecol. Prog. Ser.* **273**, 291–302 (2004).
6. Ruppert, J. L. W. et al. Human activities as a driver of spatial variation in the trophic structure of fish communities on Pacific coral reefs. *Glob. Change Biol.* **24**, e67–e76 (2018).
7. Estes, J. A. et al. Trophic downgrading of planet Earth. *Science* **333**, 301–306 (2011).
8. Gravel, D. et al. Bringing Elton and Grinnell together: a quantitative framework to represent the biogeography of ecological interaction networks. *Ecography* **42**, 401–415 (2019).
9. Woodward, G. et al. in *Advances in Ecological Research: Ecological Networks* Vol. 42 (ed. Woodward, G.) 71–138 (Academic Press, 2010).
10. Morales-Castilla, I., Matias, M. G., Gravel, D. & Araújo, M. B. Inferring biotic interactions from proxies. *Trends Ecol. Evol.* **30**, 347–356 (2015).
11. Albouy, C. et al. From projected species distribution to food-web structure under climate change. *Glob. Change Biol.* **20**, 730–741 (2014).
12. Dunne, J. A. in *Ecological Networks: Linking Structure to Dynamics in Food Webs* (Oxford Univ. Press, 2006).
13. Tittensor, D. P. et al. Global patterns and predictors of marine biodiversity across taxa. *Nature* **466**, 1098–1101 (2010).
14. Dunne, J. A., Williams, R. J. & Martinez, N. D. Food-web structure and network theory: The role of connectance and size. *Proc. Natl Acad. Sci. USA* **99**, 12917–12922 (2002).
15. Froese, R. & Pauly, D. *FishBase* (Fishbase, 2015); https://www.fishbase.org
16. *Ocean Biogeographic Information System Intergovernmental Oceanographic Commission of UNESCO* (OBIS, 2014); http://www.iobis.org
17. Barnes, C. et al. Predator and prey body size in marine food webs. *Ecology* **89**, 881–881 (2008).
18. Yen, J. D. L. et al. Linking structure and function in food webs: maximization of different ecological functions generates distinct food web structures. *J. Anim. Ecol.* **85**, 537–547 (2016).
19. Grilli, J., Rogers, T. & Allesina, S. Modularity and stability in ecological communities. *Nat. Commun.* **7**, 12031 (2016).
20. Stouffer, D. B. & Bascompte, J. Understanding food-web persistence from local to global scales. *Ecol. Lett.* **13**, 154–161 (2010).
21. Reygondeau, G. et al. Biogeography of tuna and billfish communities. *J. Biogeogr.* **39**, 114–129 (2012).
22. Richardson, A. J. & Schoeman, D. S. Climate impact on plankton ecosystems in the Northeast Atlantic. *Science* **305**, 1609–1612 (2004).
23. Hosseini, M., Nabavi, S. M. B. & Parsa, Y. Bioaccumulation of trace mercury in trophic levels of benthic, benthopelagic, pelagic fish species, and sea birds from Arvand River, Iran. *Biol. Trace Elem. Res.* **156**, 175–180 (2013).
24. Macpherson, E. Large-scale species–richness gradients in the Atlantic Ocean. *Proc. R. Soc. Lond. B* **269**, 1715–1720 (2002).
25. Trebilco, R., Baum, J. K., Salomon, A. K. & Dulvy, N. K. Ecosystem ecology: size-based constraints on the pyramids of life. *Trends Ecol. Evol.* **28**, 423–431 (2013).
26. Montoya, J. M., Woodward, G., Emmerson, M. C. & Sole, R. V. Press perturbations and indirect effects in real food webs. *Ecology* **90**, 2426–2433 (2009).
27. Bascompte, J., Melián, C. J. & Sala, E. Interaction strength combinations and the overfishing of a marine food web. *Proc. Natl Acad. Sci. USA* **102**, 5443–5447 (2005).
28. Gilarranz, L. J., Mora, C. & Bascompte, J. Anthropogenic effects are associated with a lower persistence of marine food webs. *Nat. Commun.* **7**, 10737 (2016).
29. Mougi, A. & Kondoh, M. Food-web complexity, meta-community complexity and community stability. *Sci. Rep.* **6**, 24478 (2016).
30. Plitzko, S. J. & Drossel, B. The effect of dispersal between patches on the stability of large trophic food webs. *Theor. Ecol.* **8**, 233–244 (2015).
31. Pellissier, L. et al. Comparing species interaction networks along environmental gradients. *Biol. Rev.* **93**, 785–800 (2018).
32. Quimbayo, J. P. et al. The global structure of marine cleaning mutualistic networks. *Glob. Ecol. Biogeogr.* **27**, 1238–1250 (2018).
33. Montoya, J. M. & Raffaelli, D. Climate Change, biotic interactions and ecosystem services. *Phil. Trans. R. Soc. Lond. B* **365**, 2013–2018 (2010).
34. 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).
35. Boyce, D. G., Frank, K. T., Worm, B. & Leggett, W. C. Spatial patterns and predictors of trophic control in marine ecosystems. *Ecol. Lett.* **18**, 1001–1011 (2015).
36. Barabás, G. & Allesina, S. Predicting global community properties from uncertain estimates of interaction strengths. *J. R. Soc. Interface* **12**, 20,150,218 (2015).
37. Olivier, P. & Planque, B. Complexity and structural properties of food webs in the Barents Sea. *Oikos* **126**, 1339–1346 (2017).
38. Colléter, M. et al. Global overview of the applications of the Ecopath with Ecosim modeling approach using the EcoBase models repository. *Ecol. Model.* **302**, 42–53 (2015).
39. Hattab, T. et al. Forecasting fine-scale changes in the food-web structure of coastal marine communities under climate change. *Ecography* **39**, 1227–1237 (2016).
40. Roslin, T. & Majaneva, S. The use of DNA barcodes in food web construction—terrestrial and aquatic ecologists unite! *Genome* **59**, 603–628 (2016).
41. Gascuel, D. et al. A future for marine fisheries in Europe (Manifesto of the Association Francaise d'Halieumetrie). *Fish. Res.* **109**, 1–6 (2011).
42. Parravicini, V. et al. Global patterns and predictors of tropical reef fish species richness. *Ecography* **36**, 1254–1262 (2013).
43. Bartomeus, I. et al. A common framework for identifying linkage rules across different types of interactions. *Funct. Ecol.* **30**, 1894–1903 (2016).
44. Williams, R. J. & Martinez, N. D. Simple rules yield complex food webs. *Nature* **404**, 180–183 (2000).
45. Gravel, D., Poisot, T., Albouy, C., Velez, L. & Mouillot, D. Inferring food web structure from predator-prey body size relationships. *Methods Ecol. Evol.* **4**, 1083–1090 (2013).
46. Reygondeau, G. et al. Dynamic biogeochemical provinces in the global ocean. *Glob. Biogeochem. Cycles* **27**, 1046–1058 (2013).
47. Longhurst, A. *Ecological Geography of the Sea* (Academic Press, 1998).
48. Pons, P. & Latapy, M. Computing communities in large networks using random walks. *J. Graph Algorithms Appl.* **10**, 191–218 (2006).
49. Bersier, L.-F., Banašek-Richter, C. & Cattin, M.-F. Quantitative descriptors of food-web matrices. *Ecology* **83**, 2394–2407 (2002).
50. Memmott, J., Waser, N. M. & Price, M. V. Tolerance of pollination networks to species extinctions. *Proc. Biol. Sci.* **271**, 2605–2611 (2004).
51. Wessel, P. & Smith, W. H. F. A global, self-consistent, hierarchical, high-resolution shoreline database. *J. Geophys. Res.* **101**, 8741–8743 (1996).
52. Breiman, L., Friedman, J. H., Olshen, R. A. & Stone, C. J. *Classification and Regression Trees* (Chapman & Hall/CRC, 1984).
53. Quinn, G. P. & Keough, M. J. *Experimental Design and Data Analysis for Biologists* (Cambridge Univ. Press, 2002).

Author contributions

The design of the study is the result of a working group on the global variation of ecological networks. W.A. provided the data for species occurrences. C.A. computed the data with the help of S.A.W. and conducted the analyses. D.G. contributed the analytical tools to build the metaweb. C.A., T.P., L.P. and D.G. wrote the manuscript with input from all authors.

Competing interests

The authors declare no competing interests.

Additional information

nature research

Corresponding author(s):   camille albouy

Last updated by author(s):   2019/05/30

Reporting Summary

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

Statistics

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

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

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

Software and code

Policy information about availability of computer code

Data collection	R software version 3.4.
Data analysis	R software version 3.4.

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

Data

Policy information about availability of data

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

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

A full data availability statement is included in the manuscript

Field-specific reporting

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

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

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

Ecological, evolutionary & environmental sciences study design

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

Study description	Here we have achieved the unprecedented feat of high-resolution mapping of the global diversity of all known marine fish species and their ecological interactions. Using this map, we compute for the first time the complex network of feeding interactions among all fish species around the world in order to better understand the consequences of extinctions and the propagation of perturbations. Among the key focal points of our manuscript, we find the particularly surprising result that this global food web is highly connected geographically, indicating that marine food webs have an elevated robustness to extinctions, yet their spatial connectivity may also propagate perturbations rapidly across oceans. Maps of local food webs also reveal that coastal ecosystems are more robust to extinction than those from the open water, an observation that generalizes previous findings based on punctual studies.
Research sample	Species data were obtained from the Ocean Biogeographic Information System (OBIS, http://www.iobis.org) on 08/27/2014 that represents the best available data set for a global study of marine fish food webs. We inventoried 16,238,200 occurrence records from 34,883 entries. We cleaned the data by identifying the synonyms, misspellings and rare species (only one occurrence) and by restricting it to species present in the marine environment according to FishBase (13). Synonyms were converted to accepted names. This resulted in a set of 11,503,257 occurrences for 11,345 fish species around the world. We considered the entire occurrences available in the OBIS database spanning from 1826 to 2013 as the occurrences inferior to 1950 represent solely 0.28% of the entire data set. To counteract some known biases in OBIS data (e.g., not all species/regions are equally represented), we reconstructed distribution maps for each species, defined as the convex polygon surrounding the area where each species was observed (see for details Fig. S3). The resulting polygon was divided into four parts across the world to integrate possible discontinuity between the two hemispheres and the Atlantic and Pacific Oceans. Then we refined each species distribution map by removing areas where maximal depths fell outside the minimum or maximum known depth range of the species (13). Final distribution maps of well-known species were checked visually and reviewed by the authors according to their expertise. We also compared the distribution of ~3500 tropical species present in this database to the Gaspar database frequently use in reef fish analyses (32). Finally, we aggregated fish distributions on a 1° resolution grid covering all oceans. All the data are freely available at this address https://figshare.com/s/c9ca229cc1f3548f8b5c. This scale allows computational feasibility and reproducibility of all analysis conducted here. We compiled the following environmental descriptors: distance to land, sea surface temperature (SST), sea surface salinity (SSS) and chlorophyll-a. Environmental descriptors were compiled from different sources and resolutions, and were fit to our grid system (Table S2).
Sampling strategy	NA
Data collection	NA
Timing and spatial scale	NA
Data exclusions	NA
Reproducibility	NA
Randomization	NA
Blinding	NA

Did the study involve field work? ☐ Yes ☒ No

Reporting for specific materials, systems and methods

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

Materials & experimental systems

n/a	Involved in the study
☒	Antibodies
☒	Eukaryotic cell lines
☒	Palaeontology
☒	Animals and other organisms
☒	Human research participants
☒	Clinical data

Methods

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