Body size and mobility explain species centralities in the Gulf of California food web

R. Olmo Gilabert[1], A. F. Navia[2], G. De La Cruz-Agüero[1], J. C. Molinero[3], U. Sommer[3] and M. Scotti[3,4]

[1]CICIMAR Centro Interdisciplinario de Ciencias Marinas del Instituto Politécnico Nacional, Apartado Postal 592, CP 23094, La Paz, Baja California Sur, México
[2]Fundación colombiana para la investigación y conservación de tiburones y rayas, SQUALUS. Calle 10° # 72-35, Apto. 301E, Cali, Valle, Colombia
[3]GEOMAR Helmholtz Centre for Ocean Research Kiel, Düsternbrooker Weg 20, 24105 Kiel, Germany
[4]Corresponding author. Email: marcoscot@gmail.com, phone: +49 (0) 431 600 4405

Abstract: Anthropic activities impact ecosystems worldwide thus contributing to the rapid erosion of biodiversity. The failure of traditional strategies targeting single species highlighted ecosystems as the most suitable scale to plan biodiversity management. Network analysis represents an ideal tool to model interactions in ecosystems and centrality indices have been extensively applied to quantify the structural and functional importance of species in food webs. However, many network studies fail in deciphering the ecological mechanisms that lead some species to occupy the most central positions in food webs. To address this question, we built a high-resolution food web of the Gulf of California and quantified species position using 15 centrality indices and the trophic level. We then modelled the values of each index as a function of traits and other attributes (e.g., habitat). We found that body size and mobility are the best predictors of indices that characterize species importance at local, meso- and global scale, especially in presence of data accounting for energy direction. This result extends previous findings that illustrated how a restricted set of trait-axes can predict whether two species interact in food webs. In particular, we show that traits can also help understanding the way species are affected by and mediate indirect effects. The traits allow focusing on the processes that shape the food web, rather than providing case-specific indications as the taxonomy-based approach. We suggest that future network studies should consider the traits to explicitly identify the causal relationships that link anthropic impacts to role changes of species in food webs.

Abbreviations: ESM–Electronic Supplementary Material; EBM–Ecosystem-Based management; SST–Sea Surface Temperature.

Introduction

Anthropic activities threaten global biodiversity in the oceans by direct causes like overfishing, pollution and habitat destruction (Lotze et al. 2006), or indirectly through eutrophication and climate change (Worm et al. 2000, Harley et al. 2006). Since biodiversity regulates processes (e.g., primary production and nutrient cycling; Worm et al. 2006), properties (e.g., resistance to invasion; Stachowicz et al. 2002) and their maintenance over time and space, its decline has profound consequences on the functioning of ecosystems. Although causal relationships are difficult to infer, the accelerating biodiversity loss is expected to be critical for the provision of goods (e.g., decrease of catch from fisheries; Worm et al. 2009) and services (e.g., increased exposure to flooding events and loss of nursery habitats; Díaz et al. 2006, Orth et al. 2006) to human populations. This scenario requires effective actions of ecosystem management to arrest the erosion and sustain the recovery of biodiversity. The failure of traditional strategies for fisheries management suggested shifting the focus from single species or sectors to actions targeting the whole ecosystem, a principle that lies behind the approach of "Ecosystem-Based Management" (EBM; Long et al. 2015).

Long et al. (2015) identified fifteen essential elements to successfully implement EBM. Among those elements the consideration of ecosystem connections was classified as of high importance. Such principle can be interpreted as: (1) the connectivity of marine environments, (2) the synergistic effect of multiple stressors and (3) the interactions among species (Guerry 2005). The first refers to the flow of nutrients and energy that can establish links between marine environments. For example, sub-tidal kelps supply carbon to inter-tidal communities and are vital for maintaining dense populations of patellid limpets in the southwestern rocky shores of South Africa (Bustamante et al. 1995). The second focuses on the fact that the interplay between various perturbations and their impact on ecosystems are often complicated to unveil. Indeed, in the Black Sea the overexploitation of small pelagics started in the 1970s but its consequences were masked initially by the bottom-up effect due to eutrophication (Bodini et al. 2018). The third lies on the assumption that the modelling of trophic interactions can give clues on the spread of indirect effects in food webs. Bondavalli and Ulanowicz (1999)

showed that intraguild predation can cause unexpected beneficial effects of the American alligator upon its prey (e.g., frogs are consumed by both snakes and alligators, but the latter exert stronger predatory impact on snakes than on the common prey). Therefore, to consider ecosystem connections implies emphasis on processes (i.e., interactions) linking objects (e.g., marine environment, sources of stress or species), a perspective shared with the studies of network analysis.

Network analysis represents an ideal tool to model ecosystem connections in multitrophic systems, and centrality indices can be calculated to quantify species importance in trophic interaction networks. Scotti and Jordán (2015) found that less abundant species in the Prince William Sound food web are unevenly distributed towards the top of the trophic chain and such network position further accentuates their risk of local extinction, which is already high because of the small population size. Rocchi et al. (2017) proposed network analysis to merge the gap between EBM theory and practice in data-poor contexts. They showed that, in general, the structure of the food web is robust to the removal (i.e., local extinction) of most vulnerable species (i.e., main target of fisheries) in the coastal ecosystem of Baja California Sur. Studies that rank species using centrality indices assume that being central in the trophic network corresponds to being functionally important. These studies provide indications based on the taxonomy (i.e., identity of the nodes), without informing about the ecological attributes (e.g., body size, habitat,

spatial and temporal distribution) of the most central species. The way species traits (and other attributes) are overlooked is in contrast with the goal of understanding how the position of species in the biotic interaction milieu (i.e., the realized niche) is governed by functional traits (McGill et al. 2006).

In this work, a highly disaggregated network of the Gulf of California food web was constructed. Sixteen indices (centralities and trophic level) were computed and regressed against species attributes (i.e., traits, together with features that classify spatial and temporal distribution). The attributes of the species (i.e., nodes) are independent of their network positions and allowed studying the ecological aspects that can explain node centralities and trophic level. Our goals were: (1) to investigate whether attributes can be used to predict centrality scores of species; and (2) to identify if there exists a restricted set of attributes able to characterize the position of species in the food web.

Methods

Study area

The study area is the Gulf of California, a marginal sea of the Pacific Ocean between the Baja California Peninsula and the Mexican mainland (Fig. 1). It spans from 24º N up to 32º N of latitude and is strongly influenced by the discharge of sediments from the Colorado River, with shallow waters
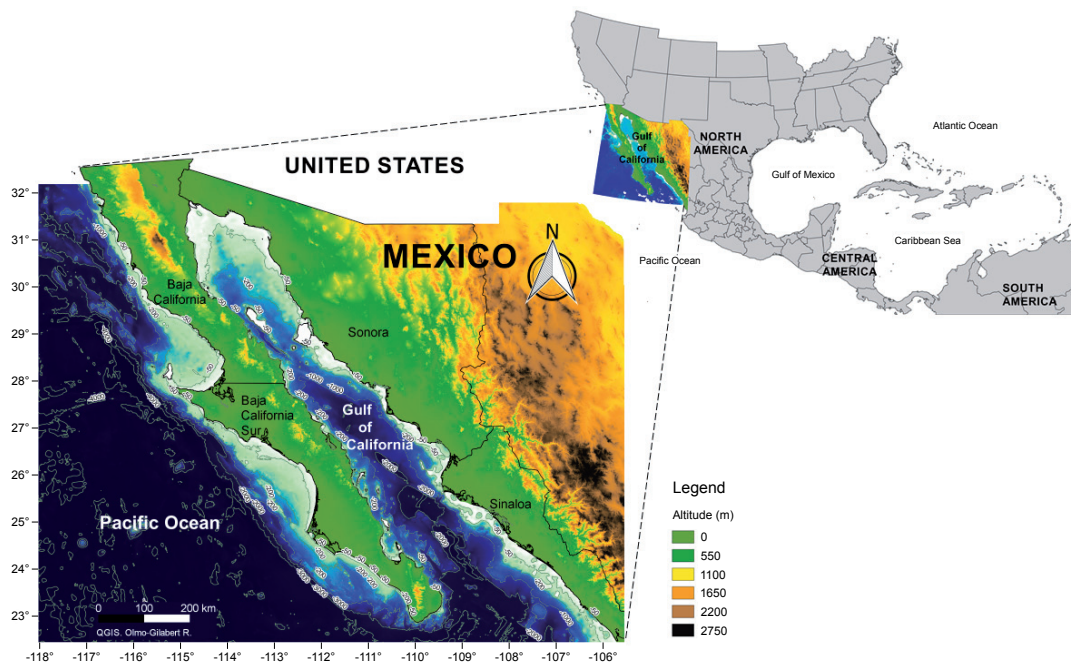

Figure 1. Map of the Gulf of California where isobaths are visualized and blue colors reflect sea depth. The map was generated with QGIS (version 2.14.20) http://www.qgis.org. The bathymetry was retrieved from the British Oceanographic Data Center (BODC) through the General Bathymetric Chart of the Oceans (GEBCO) https://www.bodc.ac.uk/data/online_delivery/gebco/. The relief layers for the continental area were obtained at a resolution of 60 m from the INEGI website and are based on the Continuo de Elevaciones Mexicano 3.0 (CEM 3.0) - discharge http://www.inegi.org.mx/geo/contenidos/datosrelieve/continental/default.aspx. The map of Mexico in the upper right corner is in vector format; it was obtained from the shape file "Mex_adm1.shp", which contains the states of the Mexican Republic. The US map layer (grey map on the top right) was obtained from the United States Geological Survey (USGS) https://www.sciencebase.gov/catalog/item/4fb555ebe4b04cb937751db9. All webpages were accessed on November 3, 2018.

in the north and deeper waters in the southern part (200-3000 m). The continental shelf is wider on the east coast due to the absence of high flow rivers on the west part. The Gulf of California has a characteristic continental climate, with sea surface temperature (SST) ranging between 13 °C and 31 °C and salinities of 35-35.8 PSU. In the northern part there are four endemic species. Three of them are economically important for fisheries (*Cynoscion othonopterus* (Jordan & Gilbert), *Micropogonias megalops* (Gilbert), and *Totoaba macdonaldi* (Gilbert)) while the fourth is *Phocoena sinus* (Norris & McFarland), the smallest and most endangered cetacean in the world. The rich biodiversity and the provision of relevant ecosystem services (e.g., resources for fisheries) call for the implementation of effective strategies of conservation (Lluch-Cota et al. 2007), an effort highlighted by the presence of 12 protected areas and 5 biosphere reserves (Sánchez-Ibarra et al. 2013).

Food web construction

The food web was constructed using literature data (e.g., gut content analysis; see Electronic Supplementary Material 1, ESM1) and consists of 317 taxa (nodes; *n* = 317), 74% of which were identified at species level. All nodes stand for living compartments (32 primary producers and 285 consumers) and the 3971 trophic interactions (links; *l* = 3971) are weighted with predator's feeding preferences, which consider the proportions of food that each predator receives through its prey (i.e., the total amount of food consumed by each predator is set to 1 and the relative contribution of each prey corresponds to the weight of the link). Therefore, the feeding preferences are always included in the interval (0, 1]. The only case with link weight = 1 is when the predator consumes a single prey type (i.e., diet specialization). Species distribution was validated through maps built with Mapmaker in the ModestR environment (García-Rosello et al. 2013), the Ocean Biogeographic Information System[1] (OBIS) and iSpecies[2]. The diet of 262 nodes (91.93% of the consumers) was defined using literature data specific to taxa that appear in the food web, while surrogate species (i.e., species that belong to the same genus, with comparable body size and maximum depth distribution) were considered for the remaining 23 nodes (8.07%; ESM1 at page 15). In particular, the diet of 186 consumers (65.26%) is based on information from the Gulf of California or areas nearby (i.e., Pacific Ocean in front of California and Central America). Main objectives were constructing a high resolution food web with most nodes representing single species and retrieving information on feeding preferences. All flows of energy from prey/resources (first column) to predators/consumers (second column), together with feeding preferences (third column), are stored in the ESM2 (text file with tab-separated values).

1 www.iobis.org
2 www.ispecies.org
[both websites last accessed on November 3, 2018]

Attributes of the taxa

To investigate the ecological meaning of various centrality indices we built a database including different attributes (e.g., traits and habitat; see ESM1). Metadata concerning all nodes were collected and the centrality scores (dependent variables) were regressed against the ecological attributes (independent variables). The following attributes were taken into account: (1) body size (classes of maximum body length measured in cm, with intervals representing 10-fold increases), (2) habitat (nine categories: reef, benthic, benthic-neritic, benthic-pelagic, pelagic-neritic, demersal-neritic, pelagic, pelagic-oceanic, or demersal), (3) maximum depth (to distinguish among: species close to surface (0,10] m, subtidal zone close to coast and reefs (10, 100] m, continental shelf (100, 200] m, continental slope in the upper part of mesopelagic zone (200, 500] m, lower part of mesopelagic zone (500, 1000] m, aphotic zone (1000, 2000] m, or deeper zones where large size migrants can live (2000, 3000] m), (4) mobility (i.e., taxa were classified as sessile or moving at various velocities: "low" for taxa displaying mainly passive movements that cannot be efficient for predator avoidance; "medium" that indicates the capability to move actively to escape from predators, referring to movements restricted to specific systems as seagrass beds and coral reefs; "high" that is used for species that move fast and cross different habitats), (5) direction of spatial movement (absence of move, move in latitude and/or longitude, water column move, movements both in latitude/longitude and water column), (6) seasonality (present in summer, winter, or both seasons), (7) daily activity (diurnal, nocturnal, or both), and (8) congregatory behavior (e.g., schooling for fish; it includes two classes: yes or no). Altogether, 2536 traits were considered and 89.04% of them was retrieved from the literature for taxa present in the food web. The remaining 10.96% was estimated with the opinion of experts (Prof. Gustavo De La Cruz-Agüero and Dr. Andrés Felipe Navia). The traits not found in the literature refer to: body size (20 nodes), daily activity (147 nodes) and congregatory behavior (111 nodes). The attributes of each node are in the ESM3. Frequency distributions of the eight attributes are reported in the ESM1, Fig. S1.

Centrality indices and trophic level

Centrality indices were selected to model the importance of nodes with respect to energy circulation and the spread of indirect effects in the food web. The position of each node *i* in was characterized by degree, betweenness, closeness, importance score, keystone indices and topological importance (Fig. 2). Such indices were chosen for their capability of describing node importance from local to global network scale, taking into account the diffusion of both vertical (i.e., bottom-up and top-down) and horizontal (e.g., to discriminate among trophic cascade and apparent competition) effects (Scotti and Jordán 2015). Feeding preferences were used to calculate the weighted forms of betweenness and closeness centralities. All other indices were presented in the unweighted version as

they were computed using the architecture of trophic interactions only. The centrality indices account for the role of the nodes at local (degree), meso- (keystone indices and topological importance), and global scale (closeness, betweenness, and importance score). In addition to the centrality indices also the trophic level was quantified. The trophic level of each consumer depends on the set of resources eaten; the contribution of the resources is then weighted by considering the feeding preferences of the consumer.

The degree of i is the sum of all directed interactions in which the node is involved (dc_i); it can be decomposed into in-degree (total number of prey/resources; $dc_{in,i}$) and out-degree (total number of predators/consumers; $dc_{out,i}$) (Jordán et al. 2006):

$$dc_i = dc_{in,i} + dc_{out,i} \qquad (1)$$

Betweenness centrality measures how frequently a node i lies on the shortest paths (i.e., geodesics) connecting any pair of nodes j and k in the network (Jordán et al. 2006). Betweenness can be calculated either by considering the di-

rectionality of flows (directed betweenness; $bc_{dir,i}$) or ignoring it (undirected betweenness; $bc_{und,i}$); the first accounts for the importance of nodes in spreading bottom-up effects, while the second combines vertical (bottom-up and top-down) and horizontal effects. In presence of undirected data, and in a network of n nodes, the normalized betweenness is:

$$bc_{und,i} = \frac{\sum_{j=1}^{n} \sum_{k=1}^{n} \frac{c_{jk}(i)}{c_{jk}}}{\frac{(n-1)(n-2)}{2}}; i \neq j \neq k \qquad (2)$$

where c_{jk} is the total number of shortest paths between j and k, and $c_{jk}(i)$ are the geodesics linking j with k and crossing i. The total score is computed by summing the relative number of times the geodesics between all pairs of nodes pass through i. The denominator is used to normalize the total score and stands for the number of all pairs of nodes excluding i. Directed betweenness is calculated in analogy with (2), and the only modification concerns the normalization that uses $[(n-1)(n-2)]$ as denominator. Nodes with high betweenness
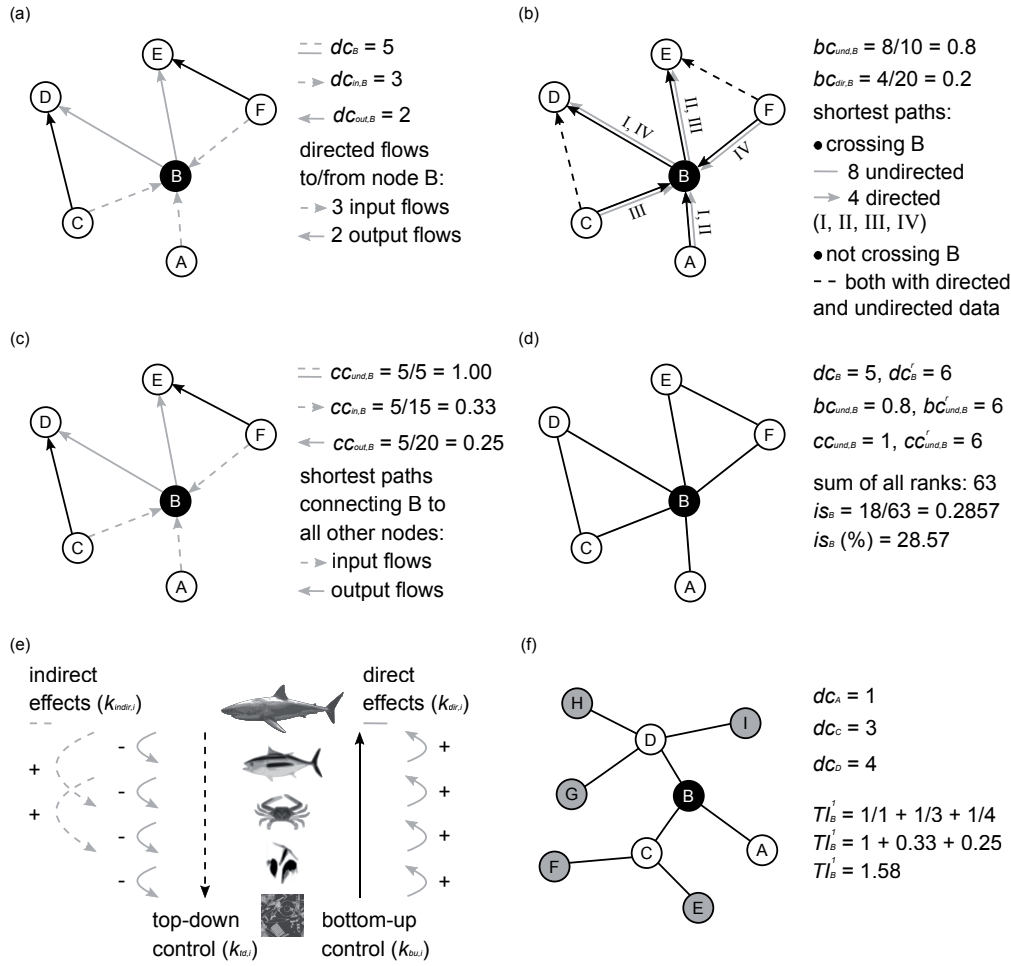

Figure 2. Schematic representation of centrality indices used in this study. These indices quantify the role of nodes with reference to local (degree, a), global (betweenness, b; closeness, c; and importance score, d), and meso- (all keystone indices but $k_{dir,i}$, which accounts for local effects only - see Fig. 3, e; and topological importance, f) scales. In the first three charts (a-c) the arrow-headed links leave the prey and point to predators. Conceptual diagram is illustrated for the keystone indices while all other centralities are calculated for node B.

are in the shortest paths linking many pairs of nodes and are expected to play relevant role in the spread of indirect effects.

Closeness centrality depends on the length of geodesics connecting a given node i to all other nodes in the network (Jordán et al. 2006). Undirected closeness is computed using the shortest paths linking node i to all other nodes, irrespective of energy flow direction ($cc_{und,i}$). Two directed versions can be calculated when shortest paths detection is constrained by energy flow orientation (i.e., strict bottom-up perspective): (a) in-closeness ($cc_{in,i}$) accounts for shortest paths from all network nodes to a given node i, while (b) out-closeness ($cc_{out,i}$) is based on the shortest paths from a given node i to all other network nodes. In directed networks there might be nodes that are not reachable when following the energy flow (e.g., two nodes that are apical predators); in such cases the shortest path is conventionally set as equal to n. In-closeness is obtained by summing all shortest paths to node i, and the size of the network (n) is used to normalize the score (the denominator is the number of all nodes except the one for which closeness is computed):

$$cc_{in,i} = \left[\frac{\sum_{k=1}^{n} d(p_k, p_i)}{n-1} \right]^{-1} \qquad (3)$$

where $d(p_k, p_i)$ is the number of steps in the geodesics linking the nodes k to i. Out-closeness is calculated according to the same principle, but considering all shortest paths from node i: $d(p_i, p_k)$. The equation for $cc_{und,i}$ has the same form of (3) but the search of geodesics from and to i does not depend on energy flow orientation. Thus, undirected closeness indicates how fast perturbations of specific nodes can propagate through the food web, irrespective of energy flow direction. Undirected closeness can also give clues on the exposure of each node to perturbations that target any other node in the food web, with the spread of effects not constrained by energy flow direction. In the directed versions of closeness, the shortest paths follow the direction of energy circulation and spread perturbations either from low trophic level nodes to the target taxa (when the in-closeness of this latter is calculated) or from the target taxa to higher trophic level nodes (when the out-closeness of the target node is computed). Therefore, in-closeness quantifies how fast a given node can be affected by perturbations on other nodes, while out-closeness defines how fast the perturbation targeting a specific node can diffuse to the rest of the network. Large values are found for: (a) nodes that can be quickly influenced by perturbations of various food web species ($cc_{und,i}$ and $cc_{in,i}$), and (b) nodes the disturbance of which can rapidly spread and affect many species in the food web ($cc_{und,i}$ and $cc_{out,i}$).

The focus of degree centrality is on local interactions, while betweenness and closeness characterize node position under a whole-network perspective (by taking into account the relative position compared to all pairs of nodes or all other nodes, respectively). In the attempt of simultaneously considering these alternative aspects, an importance score (is_i) was calculated for each node i by combining the local and global information portrayed by degree, undirected betweenness and

undirected closeness. These indices were selected because they are the most widely used centralities in food web analysis (e.g., see Scotti and Jordán 2010, 2015). First, for each centrality the nodes were ranked in descending order (i.e., the value of the most central node is 317). Second, for each node i the rank values (denoted by the r superscripts) were summed:

$$dc_i^r + bc_{und,i}^r + cc_{und,i}^r \qquad (4)$$

Finally, each score in (4) was normalized by dividing it with the sum of all cumulative ranks:

$$is_i = \frac{dc_i^r + bc_{und,i}^r + cc_{und,i}^r}{\sum_{j=1}^{n} dc_j^r + bc_{und,j}^r + cc_{und,j}^r} \qquad (5)$$

The meso-scale importance of any given node i was characterized by the keystone index (k_i) (Jordán et al. 2006). Such index requires as input the binary matrix of trophic interactions without cycles. It is calculated by summing bottom-up ($k_{bu,i}$) and top-down ($k_{td,i}$) keystone indices

$$k_{bu,i} = \sum_{j=1}^{m} \frac{1}{dc_{in,j}} (1 + k_{bu,j}) \qquad (6)$$

$$k_{td,i} = \sum_{z=1}^{q} \frac{1}{dc_{out,z}} (1 + k_{td,z}) \qquad (7)$$

$$k_i = k_{bu,i} + k_{td,i} \qquad (8)$$

where m is the number of predators feeding on species i and q is the number of prey eaten by species i; $dc_{in,j}$ is the number of prey of j (in-degree) and $dc_{out,z}$ is the number of predators of z (out-degree). The term $k_{bu,j}$ stands for the bottom-up keystone index of predator j, while $k_{td,z}$ indicates the top-down keystone index of prey z. Both the bottom-up and the top-down variants include parts that refer to direct and indirect effects. Therefore, the keystone index can be rewritten by summing two other terms: the keystone indices of direct ($k_{dir,i}$) and indirect ($k_{indir,i}$) effects

$$k_{dir,i} = \sum_{j=1}^{m} \frac{1}{dc_{in,j}} + \sum_{z=1}^{q} \frac{1}{dc_{out,z}} \qquad (9)$$

$$k_{indir,i} = \sum_{j=1}^{m} \frac{k_{bu,j}}{dc_{in,j}} + \sum_{z=1}^{q} \frac{k_{td,z}}{dc_{out,z}} \qquad (10)$$

$$k_i = k_{dir,i} + k_{indir,i} \qquad (11)$$

Both $k_{bu,i} + k_{td,i}$ and $k_{dir,i} + k_{indir,i}$ equal k_i. The keystone index k_i is particularly suitable to quantify the relevance of vertical interactions over horizontal ones, being thus appropriate for identifying the impacts related to trophic cascade rather than those due to apparent competition.

Topological importance was used to model meso-scale effects by also taking into account exploitative and apparent competitions (Jordán et al. 2006). This index considers

all interactions as undirected. The effect of j on i when these two nodes are m steps from each other is $a_{m,ij}$. The simplest way to calculate $a_{m,ij}$ is when $m = 1$; in such case the effect corresponds to the reciprocal of the degree of node i: $a_{1,ij} = 1/dc_i$. Indirect chain effects are assumed to be additive and multiplicative. If two pathways of length $m = 2$ connect i to j by passing through the nodes q and h, then the net effect of j on i ($a_{2,ij}$) is calculated by summing the two individual 2-steps effects:

$$a_{2,ij(q)} = a_{1,qj} \times a_{1,iq} \qquad (12)$$

$$a_{2,ij(h)} = a_{1,hj} \times a_{1,ih} \qquad (13)$$

$$a_{2,ij} = a_{2,ij(q)} + a_{2,ij(h)} \qquad (14)$$

where $a_{2,ij(q)}$ and $a_{2,ij(h)}$ are the individual effects of j on i transmitted via 2-steps pathways that cross nodes q and h, respectively. In a network with n nodes the effect of all species on node i via pathways of exactly m steps is

$$\alpha_{m,i} = \sum_{j=1}^{n} a_{m,ij} \qquad (15)$$

while the total effect of species i on all other species via m-steps pathways is:

$$\beta_{m,i} = \sum_{j=1}^{n} a_{m,ji} \qquad (16)$$

The topological importance of any node i (TI_i^m) is defined as the average total effect transmitted through all chains of length up to m steps (i.e., if $m = 3$, then the chains of length 1, 2, and 3 steps are considered):

$$TI_i^m = \frac{\sum_{z=1}^{m} \beta_{z,i}}{m} = \frac{\sum_{z=1}^{m} \sum_{j=1}^{n} a_{z,ji}}{m} \qquad (17)$$

Here the index was calculated up to four steps ($m = 4$) because previous works demonstrated that meso-scale, horizontal effects in food webs tend to converge when such threshold is attained (Scotti and Jordán 2015). The topological importance complements the meso-scale description provided by keystone indices.

Besides centrality indices, the position of nodes in food webs is often characterized by the trophic level. The trophic level quantifies the average distance of nodes from the external source of energy (i.e., the average number of steps through which energy is transferred from primary producers to consumers; Scotti et al. 2006). In a food web with n nodes, the trophic level of any species i (TL_i) depends on the trophic level and the relative amount of energy received from its prey:

$$TL_i = 1 + \sum_{j=1}^{n} TL_j \times g_{ji} \qquad (18)$$

where TL_j is the trophic level of prey j and g_{ji} indicates the feeding preference of i for j (this value is in the interval $0 \leq$ $g_{ji} \leq 1$). The trophic level is calculated from local information and takes into account energy flow direction. It represents a relevant indicator as the distribution of the most central species, calculated using various indices, shows regular patterns along the trophic hierarchy (Scotti and Jordán 2010).

Network analysis was performed in the R statistical environment (R Core Team 2017) except for the keystone indices, which were obtained with the FLKS 1.1 software program (Jordán et al. 2006). All versions of degree, betweenness and closeness were calculated using the 'igraph' package (Csárdi and Nepusz 2006); the importance score, the topological importance and the trophic level were computed with ad hoc R scripts.

Statistical analysis

Centrality indices and trophic level (i.e., dependent variables) were regressed against species attributes (i.e., independent variables). Eklöf et al. (2013) showed that a limited set of trait axes is required to model interactions in ecological networks (they found that the minimum number of dimensions has mean 2.665 and the "best" number of dimensions has mean 1.395). Therefore, we limited the assessment to models including either a single variable or two variables (i.e., all possible combinations of two variables were explored). The predictive power of both single attributes and combinations of two attributes was tested using generalized linear models (GLMs). GLMs were fit with the *glm* function from the 'stats' R package (R Core Team 2017). Gaussian, gamma or Poisson distributions were used to model the centrality indices and the choice of the distribution depends on the properties of the index (e.g., Poisson distribution to model integer values with zeroes as in the case of in-degree centrality; see Table 1). For all models, either identity or logarithmic link function was used. Model selection was based on the Akaike Information Criterion (AIC). The amount of deviance explained by each GLM (i.e., D^2) was calculated with the *Dsquared* function from the 'modEvA' R package (Barbosa et al. 2016). When an overall statistically significant difference in group means was detected with ANOVA, post-hoc comparisons were performed with the Tukey's Honest Significant Difference (HSD) test (*glht* function from the 'multcomp' R package; Hothorn et al. 2008). All p-values of multiple comparisons performed with Tukey's test were adjusted with the single-step method. Pairwise similarities between all indices (centralities and trophic level) were quantified with Spearman's rank correlation coefficients. These coefficients were used to construct the dendrogram of similarities, based on Ward's minimum variance clustering method.

Results

Body size and mobility are the best predictors to model centrality indices and trophic level in the Gulf of California food web (Table 1). They appear in 12 (body size) and 8 (mobility) GLMs (Fig. 3). However, a two-way ANOVA run to examine the main effect of the two traits showed that only mobility is positively associated with GLMs that account

Table 1. ANOVA results for centrality indices and trophic level. The column "statistics" reports χ^2 (dc_{in}, dc_{out}, dc) or F values (for all other indices). The last two columns indicate families and link functions used in the GLMs.

test	term	statistics	p	df	family	link
dc_{in}	body size	587.870	< 0.001	4	Poisson	log
(D^2 = 0.386)	mobility	531.150	< 0.001	3		
dc_{out}	body size	1291.030	< 0.001	4	Poisson	log
(D^2 = 0.378)	habitat	992.580	< 0.001	8		
dc	mobility	436.820	< 0.001	3	Poisson	log
(D^2 = 0.204)	direction of spatial movement	536.620	< 0.001	3		
bc_{und}	direction of spatial movement	2.663	0.048	3	Gaussian	identity
(D^2 = 0.025)						
bc_{dir}	mobility	8.335	< 0.001	3	Gaussian	identity
(D^2 = 0.103)	seasonality	5.433	0.005	2		
cc_{in}	body size	66.379	< 0.001	4	Gamma	log
(D^2 = 0.536)	mobility	52.111	< 0.001	3		
cc_{out}	body size	71.520	< 0.001	4	Gamma	log
(D^2 = 0.536)	mobility	25.069	< 0.001	3		
cc_{und}	body size	8.270	< 0.001	4	Gamma	log
(D^2 = 0.070)						
is	body size	11.946	< 0.001	4	Gamma	log
(D^2 = 0.182)	mobility	10.630	< 0.001	3		
k_{bu}	body size	17.507	< 0.001	4	Gaussian	identity
(D^2 = 0.212)	mobility	4.382	0.005	3		
k_{td}	body size	6.843	< 0.001	4	Gaussian	identity
(D^2 = 0.137)	direction of spatial movement	7.264	< 0.001	3		
k_{dir}	direction of spatial movement	14.116	< 0.001	3	Gaussian	identity
(D^2 = 0.143)	activity	4.670	0.010	2		
K_{indir}	body size	4.891	< 0.001	4	Gaussian	identity
(D^2 = 0.091)	direction of spatial movement	3.831	0.010	3		
k	body size	4.890	< 0.001	4	Gaussian	identity
(D^2 = 0.094)	direction of spatial movement	5.746	< 0.001	3		
TI^4	body size	6.582	< 0.001	4	Gamma	log
(D^2 = 0.187)	direction of spatial movement	12.204	< 0.001	3		
TL	body size	81.012	< 0.001	4	Gamma	log
(D^2 = 0.666)	mobility	102.751	< 0.001	3		

for larger proportions of deviance (i.e., with D^2 > 0.2) [χ^2 (1, 15) = 7.411, p = 0.006], while no significant relationship was found for the body size [χ^2 (1, 15) = 0.796, p = 0.372]. Larger D^2 values were obtained for local indices calculated using directed data [two-way ANOVA with main effects only; directed data: χ^2 (1, 15) = 8.392, p = 0.004; local importance: χ^2 (1, 15) = 4.035, p = 0.045]. The fact that centralities computed using directed data and quantifying node importance at local scale are better predicted by traits can be only partially explained by similarities among the indices. Indeed, the ranking of taxa based on centralities and trophic level allows obtaining three clusters (Fig. 4) and two of them include the indices modelled by GLMs with D^2 larger than 0.2. The indices with D^2 > 0.2 are in clusters that mainly account for the energy that either enters (dc_{in}, dc, cc_{in} and TL) or leaves (dc_{out}, cc_{out} and k_{bu}) the nodes for which centralities are computed. The third cluster includes many indices based on undirected data (e.g., bc_{und}, TI^4) and poorly explained by traits (i.e., D^2 < 0.2). Also the importance score (is) belongs to the third cluster related to undirected data, due to similarities in node rankings determined by the use of bc_{und} and cc_{und}. Body size has pervasive consequences on the number of prey consumed

by each taxon (dc_{in}). The number of prey per taxon significantly increases from the class of smallest size organisms (0, 1] to the class of large size animals (100, 1000], while the trophic habits of largest megafauna (1000, 10000] do not differ from those of species in the small (1, 10] and medium (10, 100] size classes (see Tukey's test in Table S2, ESM1). Also, animals with medium and high mobility consume significantly more prey than sessile organisms and animals with low mobility. The relevance of body size in trophic interactions is corroborated by the decreasing number of consumers per taxon with increasing body size of the prey (dc_{out}). However, such trend is non-monotonic and taxa belonging to the class of smallest size animals (0, 1] have significantly less consumers than organisms in the small (1, 10] and medium (10, 100] size classes (see Tukey's test in Table S5, ESM1). The number of predators per taxon is also affected by the habitat, with prey inhabiting coral reefs and demersal zones showing significantly less predators than prey from pelagic and benthic habitats. In-closeness (cc_{in}) and trophic level (TL) display similarities with in-degree centrality (dc_{in}), with body size and mobility being strong predictors that explain large proportions of deviance (D^2 > 0.2; see Tukey's

Figure 3. Summary of best predictors used to model centrality indices and trophic level. The indices quantify either local or non-local (i.e., at meso- and global-scale) importance and are computed with either directed or undirected data. Traits and other attributes (e.g., habitat) are the predictors taken into account to describe the scores of the indices. Body size and mobility are the most informative traits for modelling centrality indices and trophic level. The last column summarizes the GLMs capable of explaining the largest amount of deviance.

index	local	direct.	body size	habitat	depth	mobility	spatial	season	activity	congr.	D² > 0.2
dc_{in}	■	■	■			■					■
dc_{out}			■	■							
dc		■	■			■	■				
bc_{und}						■	■				
bc_{dir}		■				■	■	■			
cc_{in}			■			■					■
cc_{out}		■	■								
cc_{und}		■				■					■
is						■					
k_{bu}			■						■		■
k_{td}			■				■				
k_{dir}	■								■		
k_{indir}			■			■					
k						■					
TI^4						■	■				
TL			■			■					■
counts	5	8	12	1	0	8	7	1	1	0	7

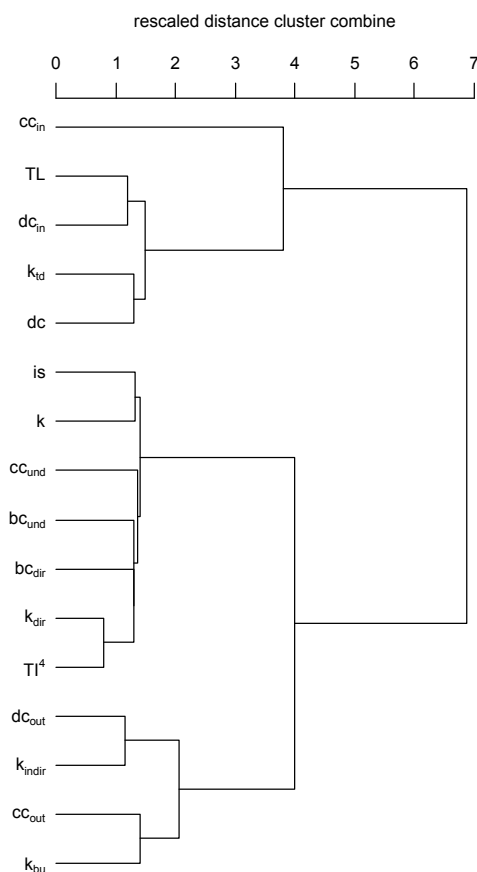rescaled distance cluster combine

Figure 4. Dendrogram of similarities among indices, which was built using Spearman's rank correlation coefficients. The dendrogram was assembled with the Ward's minimum variance clustering method. Key of the codes: cc_{in} = in-closeness, TL = trophic level, dc_{in} = in-degree, k_{td} = top-down keystone index, dc = degree, is = importance score, k = keystone index, cc_{und} = undirected closeness, bc_{und} = undirected betweenness, bc_{dir} = directed betweenness, k_{dir} = keystone index, direct effects, TI^4 = topological importance up to 4 steps, dc_{out} = out-degree, k_{indir} = keystone index, indirect effects, cc_{out} = out-closeness, k_{bu} = bottom-up keystone index.

test in Tables S16-S17 and S45-S46, ESM1). Animals with medium and high mobility have significantly higher cc_{in} and TL than sessile organisms and taxa with low mobility. Differences among body size groups are less straightforward than in the case of dc_{in}: (1) medium (10, 100] and large (100, 1000] size animals have significantly higher cc_{in} than organisms in smaller size classes and large megafauna (this latter holds for large animals only); (2) organisms of smallest size (0, 1] show significantly lower TL than animals belonging to all other size classes; (3) large size animals (100, 1000] feed at significantly higher TL than all other taxa, with the exception of large megafauna (that include both filter-feeding baleen whales and sperm whale). Out-closeness (cc_{out}) and bottom-up keystone index (k_{bu}) vary with body size and mobility with analogous patterns, thus deviating from out-degree (dc_{out}) that was better modeled with body size and habitat (see Tukey's test in Tables S5-S6, ESM1). Both cc_{out} and k_{bu} decrease with body size and mobility, being significantly higher for smaller size animals. Sessile and less motile organisms have significantly higher chances of affecting taxa that receive energy (either directly or indirectly) from them (i.e., they have higher cc_{out} than taxa that lie downstream in the trophic hierarchy). Degree (dc) is the only index based on undirected data for which two traits can explain a sufficiently large amount of deviance ($D^2 > 0.2$). It attains largest values in presence of animals that move along the water column with medium mobility. The scores calculated for each node using all 15 centrality indices and the trophic level are stored in the ESM3. In the ESM1 we illustrated frequency distributions of centrality indices and trophic level (Fig. S2), and nodes' similarities defined by the 15 centrality indices (non-metric multidimensional scaling in Fig. S19).

Discussion

Network analysis is often applied to quantify the structural and functional importance of species in food webs. For example, centrality indices can be used to identify which spe-

cies should be protected to increase the resilience of marine food webs (Bornatowski et al. 2014, Scotti and Jordán 2015, Navia et al. 2016, Rocchi et al. 2017). However, poor details are known about the ecological meaning of centrality indices. Here we calculated centralities and trophic level of species and trophospecies (i.e., groups of species that share the same prey and are eaten by the same predators) in the Gulf of California food web. To investigate the possible mechanisms behind the rankings provided by different centralities we modelled the values calculated for each index as a function of independent ecological attributes, such as traits and habitat. We found that body size and mobility are the best predictors, particularly in the case of indices that describe species importance at local scale by considering the direction of energy flow in the food web (Table 1, Fig. 3).

Theoretical models have shown that a single trait allows generating networks with structural properties coherent with those of empirical food webs. Body size was evoked as an important trait in theoretical models that order the nodes along one niche axis to define the probabilities of having prey (e.g., cascade and niche models; Cohen and Newman 1985, Williams and Martinez 2000). Also, Petchey et al. (2008) started from the body size of species to construct an optimal foraging model and used allometries of foraging variables for predicting the structure of real food webs. Further research elucidated that body size is a key trait to predict whether two species interact in food webs (Eklöf et al. 2013). Our work confirms the pivotal role that body size has in shaping the architecture of trophic interactions (Table 1). Larger animals consume more types of prey than species of smaller size, even though large megafauna (e.g., blue whale and sperm whale) do not comply with such trend. The predators per species decrease with the body size but the smallest organisms present in the food web (e.g., dinoflagellates and cladocerans) are not the most vulnerable. The non-univocal trend found for smallest organisms could be due to the low resolution of the microbial food web, with all organisms of size ≤ 1 cm grouped in the same class (Sommer et al. 2018). These relationships linking body size to numbers of prey (dc_{in}) and predators (dc_{out}) can explain why species feeding higher in the trophic hierarchy have more prey than those with lower trophic level (Cohen and Newman 1985, Williams and Martinez 2000; see Fig. S20 in ESM1 for the comparison between the trends of centrality indices along the trophic hierarchy in the Gulf of California food web and in theoretical models). Also in the Gulf of California food web, species at higher trophic levels (TL) have larger body size than species towards the bottom of the trophic chain. The skewed distribution of trophic interactions along the food chain holds for diverse aquatic food webs (Scotti et al. 2009a,b). The uneven distribution is due to the fact that more prey types are consumed by predators at higher trophic levels compared to what happens for species feeding lower in the trophic chain. However, such pattern is reversed when trophic interactions are weighted by the actual amount of energy/matter flowing in the food web (Scotti et al. 2009b). Species at lower trophic levels have less prey types but feed more evenly on them than predators that lie towards the top of the trophic hierarchy. The latter exhibit

wider potential dietary breadths than lower trophic level consumers, but they satisfy the energy demand by preferentially feeding on small subsets of few prey (i.e., with feeding specialization). Our data confirm the strong positive correlation between trophic level and number of prey per predator (Spearman's ρ = 0.666, p < 0.001), with body size distribution along the trophic hierarchy responsible for such relationship. The study of feeding preferences reveals that also in the Gulf of California food web the strongest interactions are unevenly distributed towards the top of the trophic chain, with few interactions responsible for most of energy/matter flow (Spearman's ρ = 0.640, p < 0.001). An exception is represented by Eden's whale (*Balaenoptera edeni* Anderson) that, despite its body size, feeds at intermediate trophic level (TL = 3.279) and shows no trophic specialization for its prey (i.e., fish and planktonic crustaceans). The prevalence of weak links towards the top of the trophic hierarchy might however have been affected by the principles adopted to construct the food webs analyzed by Scotti et al. (2009a,b). Indeed, in most weighted food webs the nodes at lower trophic levels lump together various species, while better resolution is attained at higher trophic levels (i.e., with a 1:1 correspondence between species and nodes). Our analysis was based on a high resolution food web composed of 317 nodes, 74% of which represent single species. It confirmed that species at higher trophic levels have larger size, display more potential prey and higher specialization than species feeding lower in the trophic chain, thus suggesting that the findings of Scotti et al. (2009a,b) were not necessarily biased by inhomogeneous node resolution along the trophic chain. Nevertheless, microbial trophic interactions are poorly represented in the current version of the Gulf of California food web and this might have led to underestimating both dietary breadth and trophic level of various taxa (Sommer et al. 2018). For example, the heterotrophic flagellates are crucial in the microbial food web of marine systems but do not form a homogeneous trophic guild (Moustaka-Gouni et al. 2016). Thus, the inclusion of heterotrophic flagellates and the adoption of a fine taxonomic description might modify the centrality of tunicates and other members of the "jelly food chain" (Sommer et al. 2018).

The relevance of body size goes beyond its predictive power in quantifying species importance at local scale (i.e., when it is used to model dc_{in}, dc_{out} and TL). In particular, the relationships linking body size to directed closeness centralities are consistent with trends observed using in- and out-degree (see Table 1 and Tukey's tests in Tables S2, S5, S16 and S19, ESM1). First, cc_{in} increases with body size as it happens to the number of prey per predator (dc_{in}). Therefore, when the exposure to perturbations is investigated by considering energy intake (cc_{in}), large species that feed at high trophic levels are more quickly influenced compared to small organisms towards the bottom of the trophic hierarchy. The high vulnerability exhibited by species with more prey contradicts the expectation that wide dietary breadths act as buffer to prevent negative consequences due to biomass fluctuations of some resources (MacArthur 1955). The striking exposure of generalist species to perturbations can however be explained with the presence of multiple pathways through which the distur-

bance reaches the species at high trophic levels. Second, cc_{out} shows an effect of body size that is coherent with what found for dc_{out} (i.e., small body size animals are the most central). Hence, if constraints imposed by energy flow direction are taken into account, the spread of perturbations (cc_{out}) occurs faster when anthropic pressures target small organisms at low trophic levels rather than large animals at the top of the trophic chain. Body size, a relatively simple trait to determine, can thus predict indirect effects in addition to influencing the way species interact (Eklöf et al. 2013). These results can be explained with the similarities among trophic level, directed degree and closeness centralities (Fig. 4).

Mobility is the second most frequently used predictor to model centrality indices in the Gulf of California food web and is significantly associated to GLMs that explain largest amounts of deviance ($D^2 > 0.2$; see Table 1 and Fig. 3). Mobility has been already shown to yield higher overlap between predicted and empirical interactions than body size (Eklöf et al. 2013). The importance that mobility has in determining the role played by species in food webs is further confirmed by our study. This is because we found that mobility can also predict the scores of centralities quantifying indirect effects (i.e., bc_{dir}, cc_{in}, cc_{out}, is and k_{bu}). The relevance of mobility contrasts with some theoretical models that aim at assembling networks structurally equivalent to empirical food webs starting from body size alone (Williams and Martinez 2000, Petchey et al. 2008). Allesina et al. (2008) proposed an alternative framework for the inclusion of multidimensional niches. In their food web model the predators can choose the prey depending on more traits. Since body size and mobility are the two independent variables used to model dc_{in} (Table 1), these traits define the most promising bi-dimensional niche space to account for most feeding relations in the Gulf of California food web. Our results do not exclude that different combinations of traits can represent the most suitable dimensions to model feeding relationships in other food webs. For example, habitat of the resources is one of the traits with the largest explanatory power for predicting trophic interactions in the Kongsfjorden food web (Eklöf et al. 2013). This agrees with the relevance that body size and habitat have as predictors of the number of predators per prey (dc_{out}) in the Gulf of California food web (Table 1).

Our study is based on a highly resolved food web, with most of the nodes representing single species. The detailed description of nodes and trophic interactions is an essential requisite to investigate whether traits and attributes can predict the scores of centrality indices. This is because the unambiguous classification of traits and attributes gets more complicated when various taxa are grouped in the same node. Also, the prevalence of nodes that include single species allows discriminating differences that strictly depend on feeding preferences (i.e., when two species consume the same set of prey but feed on them in different proportions). However, some limits persist even in presence of high resolution food webs that exhibit 1:1 correspondence between nodes and species. First, intraspecific trait variation cannot be taken into account despite it has large ecological effects (Bolnick et al. 2011). Second, some issues can emerge in the treatment of traits that change along the ontogenetic development of species (Reiss et al. 2009). Moreover, the limits imposed by the resolution can also affect the food web structure. This is the case of our model that represents most fish species as single nodes but does not consider ontogenetic diet shifts, which are common in many fish when becoming adult (Valenzuela-Quiñonez et al. 2017).

In the Gulf of California food web, body size and mobility are good predictors of centrality indices that encompass diverse definitions of species importance (i.e., constrained by energy flow direction or not, and with reference to local-, meso- or global-scale; see Fig. 3). This result integrates previous findings that showed how a few functional trait-axes are required to model feeding interactions in food webs (Eklöf et al. 2013). The feeding interactions (i.e., directed degree centralities, which define the interaction milieu; McGill et al. 2006) in the Gulf of California food web are constrained by three traits: body size, mobility and habitat. The novelty of our work stands in the evidence that a restricted subset of traits characterizes the role of species also at meso- and global-scale. We suggest that the use of a trait-based approach to model feeding interactions and species importance in trophic networks can help finding general patterns (McGill et al. 2006). This is because it allows focusing on the processes that shape the ecological communities, rather than providing highly contingent rules as nomenclatural studies. Functional traits can explain the ecological meaning of network analysis results (e.g., scores of centrality indices), thus making possible to establish indirect but explicit links between the changes of environmental conditions (e.g., global warming that modifies the frequency of some traits) and their consequences on food web structure. Therefore, merging trait ecology with network analysis is essential to understand how rapidly changing conditions due to human activities may impact interactions and persistence of species in food webs.

References

Allesina, S., D. Alonso and M. Pascual. 2008. A general model for food web structure. *Science* 320:658–661.

Barbosa A.M., J.A. Brown, A. Jimenez-Valverde and R. Real. 2016. modEvA: Model Evaluation and Analysis. R package version 1.3.2. URL https://CRAN.R-project.org/package=modEvA

Bodini, A., M. Rocchi and M. Scotti. 2018. Insights into the ecology of the Black Sea through the qualitative loop analysis of the community structure. *Limnol. Oceanogr.* 63:968–984.

Bolnick, D.I., P. Amarasekare, M.S. Araújo, R. Bürger, J.M. Levine, M. Novak, V.H.W. Rudolf, S.J. Schreiber, M.C. Urban and D.A. Vasseur. 2011. Why intraspecific trait variation matters in community ecology. *Trends Ecol. Evol.* 26:183–192.

Bondavalli, C. and R.E. Ulanowicz. 1999. Unexpected effects of predators upon their prey: the case of the American alligator. *Ecosystems* 2:49–63.

Bornatowski, H., A.F. Navia, R.R. Braga, V. Abilhoa and M.F. Maia-Corrêa. 2014. Ecological importance of sharks and rays in a structural food web analysis in southern Brazil. *ICES J. Mar. Sci.* 71:1586–1592.

Bustamante, R.H., G.M Branch and S. Eekhout. 1995. Maintenance of an exceptional intertidal grazer biomass in South Africa: subsidy by subtidal kelps. *Ecology* 76:2314–2329.

Cohen, J.E. and C.M. Newman. 1985. A stochastic theory of community food webs I. Models and aggregated data. *Proc. R. Soc. Lond. B* 224:421–448.

Csárdi, G. and T. Nepusz. 2006. The igraph software package for complex network research. *InterJournal, Complex Systems* 1695:1–9.

Díaz, S., J. Fargione, F.S. Chapin III and D. Tilman. 2006. Biodiversity loss threatens human well-being. *PLOS Biol.* 4:e277.

Eklöf, A., U. Jacob, J. Kopp, J. Bosch, R. Castro-Urgal, N.P. Chacoff, B. Dalsgaard, C. Sassi, M. Galetti, P.R. Guimaraes, S.B. Lomascolo, et al. 2013. The dimensionality of ecological networks. *Ecol. Lett.* 16:577–583.

García-Rosello, E., C. Guisande, J. González-Dacosta, J. Heine, P. Pelayo-Villamil, A. Manjarrés-Hernández, A. Vaamonde and C. Granado-Lorencio. 2013. ModestR: a software tool for managing and analyzing species distribution map databases. *Ecography* 36:1202–1207.

Guerry, A.D. 2005. Icarus and Daedalus: conceptual and tactical lessons for marine ecosystem-based management. *Front. Ecol. Environ.* 3:202–211.

Harley, C.D., A. Randall Hughes, K.M. Hultgren, B.G. Miner, C.J. Sorte, C.S. Thornber, L.F. Rodriguez, L. Tomanek and S.L. Williams. 2006. The impacts of climate change in coastal marine systems. *Ecol. Lett.* 9:228–241.

Hothorn, T., F. Bretz and P. Westfall. 2008. Simultaneous inference in general parametric models. *Biom. J.* 50:346–363.

Jordán, F., W.C. Liu and A.J. Davis. 2006. Topological keystone species: measures of positional importance in food webs. *Oikos* 112: 535–546.

Lluch-Cota, S.E., E.A. Aragón-Noriega, F. Arreguín-Sánchez, D. Aurioles-Gamboa, J.J. Bautista-Romero, R.C. Brusca, R. Cervantes-Duarte, et al. 2007. The Gulf of California: review of ecosystem status and sustainability challenges. *Progr. Oceanogr.* 73:1–26.

Long, R.D., A. Charles and R.L. Stephenson. 2015. Key principles of marine ecosystem-based management. *Mar. Policy* 57:53–60.

Lotze, H.K., H.S. Lenihan, B.J. Bourque, R.H. Bradbury, R.G. Cooke, M.C. Kay, S.M. Kidwell, M.X. Kirby, C.H. Peterson and J.B. Jackson. 2006. Depletion, degradation, and recovery potential of estuaries and coastal seas. *Science* 312:1806–1809.

MacArthur, R. 1955. Fluctuations of animal populations and a measure of community stability. *Ecology* 36:533–536.

McGill, B.J., B.J. Enquist, E. Weiher and M. Westoby. 2006. Rebuilding community ecology from functional traits. *Trends Ecol. Evol.* 21:178–185.

Moustaka-Gouni, M., K.A. Kormas, M. Scotti, E. Vardaka and U. Sommer. 2016. Warming and acidification effects on planktonic heterotrophic pico-and nanoflagellates in a mesocosm experiment. *Protist* 167:389–410.

Navia, A.F., V.H. Cruz, A. Giraldo and A. Barausse. 2016. The structure of a marine tropical food web, and its implications for ecosystem-based fisheries management. *Ecol. Model.* 328:23–33.

Orth, R.J., T.J. Carruthers, W.C. Dennison, C.M. Duarte, J.W. Fourqurean, K.L. Heck, A.R. Hughes, G.A. Kendrick, W.J. Kenworthy, F. Olyarnik, F.T. Short, et al. 2006. A global crisis for seagrass ecosystems. *BioScience* 56:987–996.

Petchey, O.L., A.P. Beckerman, J.O. Riede and P.H. Warren. 2008. Size, foraging, and food web structure. *Proc. Natl. Acad. Sci. U.S.A.* 105:4191–4196.

R Core Team. 2017. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/

Reiss, J., J.R. Bridle, J.M. Montoya and G. Woodward. 2009. Emerging horizons in biodiversity and ecosystem functioning research. *Trends Ecol. Evol.* 24:505–514.

Rocchi, M., M. Scotti, F. Micheli and A. Bodini. 2017. Key species and impact of fishery through food web analysis: a case study from Baja California Sur, Mexico. *J. Mar. Syst.* 165:92–102.

Sánchez-Ibarra, C., D.M. Bermúdez-García, J.E. Bezaury-Creel, C. Lasch-Thaler, N. Rodríguez-Dowdell, N. Cárdenas-Torres, S. Rojas-González de Castilla, A. Gondor (eds.). 2013. Plan de Acción para la Conservación y Aprovechamiento Sustentable de la Biodiversidad Terrestre y Marina de la Región Golfo de California y Pacífico Sudcaliforniano. Comisión Nacional de Áreas Naturales Protegidas (CONANP), The Nature Conservancy (TNC), Fondo Mexicano para la Conservación de la Naturaleza, A.C., 294 pp. México.

Scotti, M. and F. Jordán. 2010. Relationships between centrality indices and trophic levels in food webs. *Community Ecol.* 11:59–67.

Scotti, M. and F. Jordán. 2015. The structural importance of less abundant species in Prince William Sound food web. *Isr. J. Ecol. Evol.* 61:77–89.

Scotti, M., C. Bondavalli and A. Bodini. 2009a. Linking trophic positions and flow structure constraints in ecological networks: energy transfer efficiency or topology effect? *Ecol. Model.* 220: 3070–3080.

Scotti, M., C. Bondavalli, A. Bodini and S. Allesina. 2009b. Using trophic hierarchy to understand food web structure. *Oikos* 118: 1695–1702.

Scotti, M., S. Allesina, C. Bondavalli, A. Bodini and L.G. Abarca-Arenas. 2006. Effective trophic positions in ecological acyclic networks. *Ecol. Model.* 198:495–505.

Sommer, U., E. Charalampous, M. Scotti and M. Moustaka-Gouni. 2018. Big fish eat small fish: implications for food chain length? *Community Ecol.* 19:107–115.

Stachowicz, J.J., H. Fried, R.W. Osman and R.B. Whitlatch. 2002. Biodiversity, invasion resistance, and marine ecosystem function: reconciling pattern and process. *Ecology* 83:2575–2590.

Valenzuela-Quiñonez, F., F. Galván-Magaña, D.A. Ebert and E.A. Aragón-Noriega. 2017. Feeding habits and trophic level of the shovelnose guitarfish (*Pseudobatos productus*) in the upper Gulf of California. *J. Mar. Biol. Assoc. U.K.* 1–10.

Williams, R.J. and N.D. Martinez. 2000. Simple rules yield complex food webs. *Nature* 404:180–183.

Worm, B., E.B. Barbier, N. Beaumont, J.E. Duffy, C. Folke, B.S. Halpern, J.B. Jackson, H.K. Lotze, F. Micheli, S.R. Palumbi, E. Sala, et al. 2006. Impacts of biodiversity loss on ocean ecosystem services. *Science* 314:787–790.

Worm, B., H.K. Lotze and U. Sommer. 2000. Coastal food web structure, carbon storage, and nitrogen retention regulated by consumer pressure and nutrient loading. *Limnol. Oceanogr.* 45: 339–349.

Worm, B., R. Hilborn, J.K. Baum, T.A. Branch, J.S. Collie, C. Costello, M.J. Fogarty, E.A. Fulton, J.A. Hutchings, S. Jennings, O.P. Jensen et al. 2009. Rebuilding global fisheries. *Science* 325: 578–585.

Electronic Supplementary Material 1 (ESM1)

1. List of references consulted to retrieve node properties: diet, traits and attributes (pages 2-14)

2. Surrogate species of those in the Gulf of California food web (pages 15-16)

Figure S1. Bar plots that visualize the frequency distribution of traits and attributes for the 317 nodes of the Gulf of California food web

Figure S2. Bar plots that illustrate the frequency distributions of centrality indices and trophic level for the 317 nodes of the Gulf of California food web

3. Generalized linear models and Tukey's tests (pages 22-55)

Figure S19. Non-metric multidimensional scaling (nMDS) that represents the distance between the nodes depending on the centrality indices

Figure S20. Relations between centralities and trophic level in theoretical models

References

References consulted for the diet of the species in the Gulf of California food web (pages 59-77)

References consulted for traits and attributes of the species in the Gulf of California food web (pages 78-85)

Electronic Supplementary Material 2 (ESM2)

Text file (with tab-separated values) composed of three columns that summarize all trophic interactions in the Gulf of California food web: (i) prey/resources (first column), (ii) predators/consumers (second column); feeding preferences (third column); each row indicates a trophic interaction

Electronic Supplementary Material 3 (ESM3)

Excel file with centrality indices, trophic level and attributes (traits and habitat) for all 317 nodes