

Linking structure and function in food webs: maximization of different ecological functions generates distinct food web structures

Jian D. L. Yen^{1*}, Reniel B. Cabral^{2†§}, Mauricio Cantor^{3§}, Ian Hatton^{4§},
Susanne Kortsch^{5§}, Joana Patrício^{6§} and Masato Yamamichi^{7‡§}

¹School of Physics and Astronomy, Monash University, Melbourne, VIC 3800, Australia; ²National Institute of Physics, University of the Philippines Diliman, Quezon City 1101, Philippines; ³Department of Biology, Dalhousie University, B3H 4J1, Halifax, Canada; ⁴Department of Biology, McGill University, H3A 1B1, Montreal, Canada; ⁵Norwegian College of Fishery Science, UIT the Arctic University of Norway, 9037 Tromsø, Norway; ⁶European Commission, Joint Research Centre, Institute for Environment and Sustainability, 21027 Ispra, Italy; ⁷Department of Ecology and Evolutionary Biology, Cornell University, Ithaca, NY 14853-2701, USA

Summary

1. Trophic interactions are central to ecosystem functioning, but the link between food web structure and ecosystem functioning remains obscure. Regularities (i.e. consistent patterns) in food web structure suggest the possibility of regularities in ecosystem functioning, which might be used to relate structure to function.
2. We introduce a novel, genetic algorithm approach to simulate food webs with maximized throughput (a proxy for ecosystem functioning) and compare the structure of these simulated food webs to real empirical food webs using common metrics of food web structure. We repeat this analysis using robustness to secondary extinctions (a proxy for ecosystem resilience) instead of throughput to determine the relative contributions of ecosystem functioning and ecosystem resilience to food web structure.
3. Simulated food webs that maximized robustness were similar to real food webs when connectance (i.e. levels of interaction across the food web) was high, but this result did not extend to food webs with low connectance. Simulated food webs that maximized throughput or a combination of throughput and robustness were not similar to any real food webs.
4. Simulated maximum-throughput food webs differed markedly from maximum-robustness food webs, which suggests that maximizing different ecological functions can generate distinct food web structures. Based on our results, food web structure would appear to have a stronger relationship with ecosystem resilience than with ecosystem throughput.
5. Our genetic algorithm approach is general and is well suited to large, realistically complex food webs. Genetic algorithms can incorporate constraints on structure and can generate outputs that can be compared directly to empirical data. Our method can be used to explore a range of maximization or minimization hypotheses, providing new perspectives on the links between structure and function in ecological systems.

Key-words: biodiversity, ecological networks, ecosystem functioning, ecosystem resilience, genetic algorithm, secondary extinctions, total system throughput, trophic interactions

Introduction

The processes that define ecosystem functioning (e.g. productivity, regulatory processes, nutrient cycling and energy flow) result from complex interactions between biotic and abiotic components of ecosystems, typically linked through trophic interactions (Loreau 2010). Despite increased understanding of the mechanisms

*Correspondence author. E-mail: Jian.Yen@monash.edu

†Present address: Sustainable Fisheries Group, Bren School of Environmental Science and Management and Marine Science Institute, University of California, Santa Barbara, CA, USA

‡Present address: Hakubi Center for Advanced Research, Kyoto University, Sakyo, Kyoto, Japan

§arranged in alphabetical order.

underlying trophic interactions among species (e.g. Dunne, Williams & Martinez 2002a; Bascompte *et al.* 2003), inferring ecosystem functioning from food web structure remains a central challenge in ecology (Thompson *et al.* 2012). This challenge is complicated by the remarkable complexity (e.g. many species, indirect trophic interactions, temporal changes in interaction strength and topology) and diversity (e.g. aquatic vs. terrestrial, below-ground vs. above-ground) of real food webs.

Despite ecological community composition being highly dynamic, food web structure displays regularities, such as similarities in the number of realized trophic links (connectance) and similar distributions of feeding links into and out of a given species (degree distribution; Dunne, Williams & Martinez 2002a). Food webs are structurally complex and have many properties that are non-random and are persistent across ecosystem types and through time (e.g. modules of highly connected species) (De Angelis 1975; Stouffer & Bascompte 2010). These consistent patterns suggest some level of order underlying food web structure, which might lead to regularities in ecosystem functioning that, if identified, could be used to link food web structure to ecosystem functioning.

Throughput is one possible proxy for ecosystem functioning, being positively correlated with several functions, including biomass production, nutrient cycling, ecosystem metabolism and organic matter transformation (Costanza & Mageau 1999). High throughput is associated with high rates of biomass production and with fast cycling of matter and nutrients through an ecosystem (Costanza & Mageau 1999; Gascuel *et al.* 2008). Thermodynamics provides some justification for the maximization of throughput, with several studies exploring links between maximization principles and ecological patterns (Kleidon, Malhi & Cox 2010; Yen *et al.* 2014). Thermodynamic approaches have been used to study the outcome of interspecific competition experiments (DeLong 2008), the broad-scale distribution of vegetation (del Jesus *et al.* 2012), and the distribution of organisms among size classes in a community (Yen *et al.* 2015). These studies showed that maximized productivity is consistent with many observed properties of ecological populations and communities. Given the close links between productivity and throughput, evidence for maximized productivity suggests that throughput also might be maximized in some ecological systems. A key question is whether maximized throughput is associated with particular food web structures.

Maximization of throughput might be constrained by food web stability and robustness. Food webs with very high throughput might be unstable, which would make them susceptible to collapse (Gascuel *et al.* 2008). Therefore, it is important to consider ecosystem functioning in the context of food web persistence. Food web persistence can be captured by the notion of resilience or robustness, which emphasizes a food web's capacity to absorb and recover from perturbations (Holling 1973; May 1973). Topological robustness is one possible measure of resili-

ence and measures the capacity of a food web to absorb the loss of a single species (Dunne, Williams & Martinez 2002b). Linking maximized throughput to the robustness of a food web provides a more complete measure of ecosystem functioning that considers the traditional notion of ecosystem functioning alongside a requirement that the food web can persist.

An additional consideration is the known association between food web structure and species richness (Vermaat, Dunne & Gilbert 2009). For example, species-rich food webs often have a lower proportion of realized species–species links than species-poor food webs and typically contain longer chains and have more links per species than species-poor food webs (Vermaat, Dunne & Gilbert 2009). A measure that is closely related to species richness is connectance, which is the proportion of realized trophic links out of all possible links (May 1972; Martinez 1992). Connectance often is negatively correlated with species richness and, as a result, has a strong influence on food web structure, with high connectance being associated with many intermediate species, high levels of omnivory and high levels of clustering (Vermaat, Dunne & Gilbert 2009; Riede *et al.* 2010).

We develop a genetic algorithm approach to simulate food webs subject to general maximization (or minimization) criteria. A genetic algorithm is a heuristic method capable of converging on solutions to complex problems if explicit maximization or minimization criteria can be specified. We use this method to simulate food webs with maximized throughput, maximized robustness and a combination of these measures. We have two main goals: (i) to determine whether the maximization of different ecological functions (e.g. throughput and robustness) generates realistic food web structures and (ii) to determine whether particular conditions (e.g. high or low connectance) or ecosystem types (e.g. aquatic or terrestrial) are associated with the relationship between food web structure and function. Our genetic algorithm method is general and could be used to simulate ecological systems under a range of maximization and minimization hypotheses. This method accounts for the known relationships between food web structure and species richness and connectance (Vermaat, Dunne & Gilbert 2009). We apply our approach to a comprehensive collection of 85 binary (i.e. feeding relations are present or absent) and 21 weighted (i.e. feeding relations are weighted by consumption) empirical food webs from freshwater, marine, estuarine and terrestrial ecosystems (Tables S1, S2, Supporting information; Fig. 1).

Materials and methods

FOOD WEB DEFINITIONS

Food webs were depicted as networks whose nodes (species) are connected by directed links representing a trophic interaction (e.g. predation). A food web network was represented by an

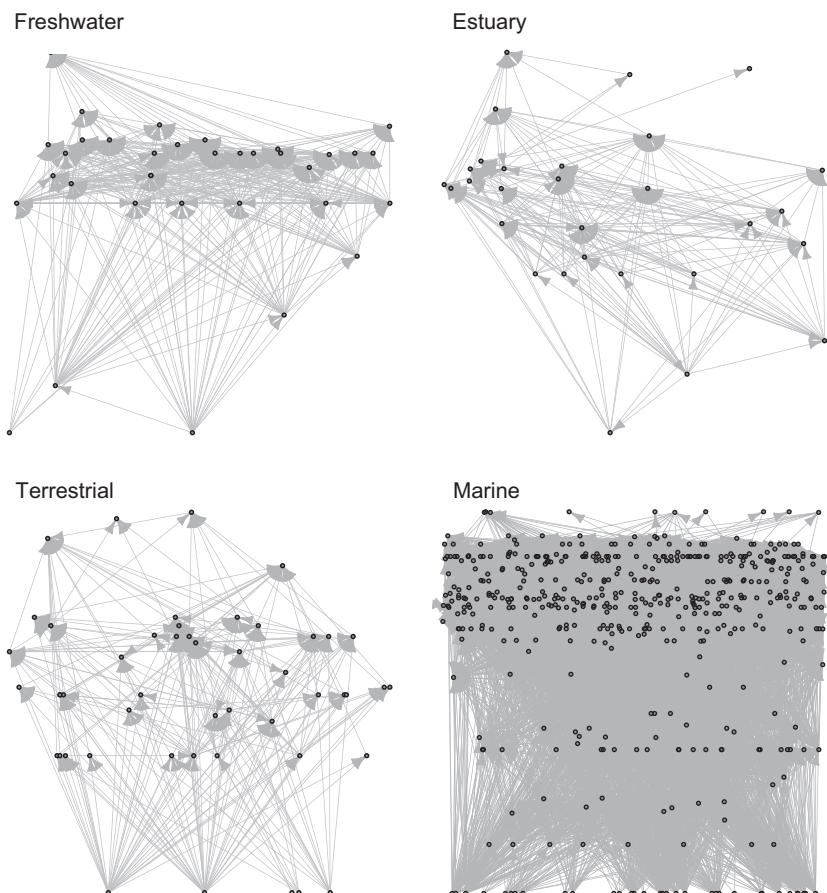


Fig. 1. Examples of real food webs used in our study. Four different ecosystem types are shown: freshwater [Skipwith pond; species richness (S) = 35; connectance (C) = 0.31], estuary (Benguela reef; S = 29; C = 0.24), terrestrial (mangrove island; S = 49; C = 0.10) and marine (Weddell sea; S = 492; C = 0.07). Complete information on the 106 empirical food webs is in Tables S1 and S2 (Supporting information).

$n \times n$ adjacency matrix \mathbf{A} , where n is the number of taxa in the food web. For binary food webs, each element a_{ij} is equal to 1 if the species in row i eats the species in the column j and is equal to 0 otherwise. For weighted food webs, each element a_{ij} is equal to the relative rate of energy transfer (e.g. grams of carbon per day) from node j to node i .

FOOD WEB DATA

We collated data for 106 food webs, encompassing terrestrial, freshwater, estuarine and marine ecosystems (Tables S1, S2, Supporting information; Opitz 1996; Link 2002; Thompson & Townsend 2004; Riede *et al.* 2010; Raymond *et al.* 2011; Planque *et al.* 2014; Csardi 2015). In these data, 85 food webs were binary, while 21 food webs contained information on relative link weights (Tables S1, S2, Supporting information). Species richness for the 85 binary food webs ranged from 27 to 586 species and connectance ranged from 0.01 to 0.31. Species richness for the 21 weighted food webs ranged from 20 to 128 species and connectance ranged from 0.05 to 0.22. Species richness and connectance were weakly negatively correlated for these data (Pearson's r = -0.22 for binary food webs, r = -0.37 for weighted food webs). Weighted food webs contain detailed information on food web functioning (e.g. relative rates of energy transfer), but estimates of link weights are scarce and can be difficult to estimate reliably. The inclusion of binary food webs alongside weighted food webs allowed us to test our method on a small number of weighted food webs, while also applying our method to a large number of high-resolution food webs (>100 species) that accurately represent the topology of general food web networks. Binary versions of all weighted food webs

were included in the binary data set, so that our results are based on 106 binary food webs and 21 weighted food webs.

DESCRIBING A FOOD WEB: METRICS OF FOOD WEB STRUCTURE

We compared simulated and real binary food webs using 10 metrics of food web structure. We classified these metrics as 'local' and 'global' metrics based on whether they emphasize species-species links or properties of the entire food web, respectively. At the local level, we calculated number of basal species, number of intermediate species, number of top species, generality of consumers and vulnerability of consumers. At the global level, we calculated mean path length, mean trophic level, mean omnivory index, modularity and transitivity (a measure of clustering; Table 1).

We compared simulated and real weighted food webs using modified versions of the five local metrics (Table 1). These modifications accounted for the additional information provided by link weights on the topological structure of food webs and follow Bersier, Banašek-Richter & Cattin (2002).

DEFINING FOOD WEB THROUGHPUT AND ROBUSTNESS

A standard definition of food web throughput is the sum of all link weights. However, maximizing this measure of throughput provides little information on food web structure because maximum throughput occurs when connectance is as high as possible

Table 1. Metrics used to compare simulated and real food webs. Five local and five global metrics were used. Definitions: L is the number of nonzero links in the food web; $I(condition)$ is an indicator function equal to 1 if *condition* is true, 0 otherwise; p is a path within the food web; n_p is the number of paths in the food web; S is the number of species; TL_i is the trophic level of species i ; x_{ij} is the link weight from species j to species i ; $x_{\bullet i}$ is the sum of all links into species i ; $x_{i \bullet}$ is the sum of all links out of species i ; $x_{\bullet \bullet}$ is the sum of all link weights; m is the number of edges; k_i is the degree of species i ; c_i is the type (cluster) of species i ; s_i is the sum of edge weights for edges adjacent to species i . All weighted formulae are adapted from Bersier, Banašek-Richter & Cattin (2002)

Scale	Metric	Formula	Weighted formula ¹	Description
Local	No. of basal species	$N_B = \sum_i I(S_i = \text{basal})$	$\sum_{j=1}^S I(d_j = 0); d_j = \frac{x_{\bullet j} n_{N,j}}{x_{\bullet j} n_{N,j} + x_{\bullet j} n_{P,j}}$	The number of species not feeding on any other species (i.e. primary producers)
	No. of intermediate species	$N_I = \sum_i I(S_i = \text{intermediate})$	$\sum_{j=1}^S I(0 < d_j < 0.99); d_j = \frac{x_{\bullet j} n_{N,j}}{x_{\bullet j} n_{N,j} + x_{\bullet j} n_{P,j}}$	The number of species either feeding on other species or with at least one predator
	No. of top species	$N_T = \sum_i I(S_i = \text{top})$	$\sum_{j=1}^S I(0.99 < d_j < 1); d_j = \frac{x_{\bullet j} n_{N,j}}{x_{\bullet j} n_{N,j} + x_{\bullet j} n_{P,j}}$	The number of species with no predators
	Generality	$\frac{L}{N_T + N_I}$	$\sum_{k=1}^S \frac{x_{k \bullet}}{x_{\bullet \bullet}} n_{N,k}$	The number of resource taxa a given consumer taxon has
	Vulnerability	$\frac{L}{N_B + N_I}$	$\sum_{k=1}^S \frac{x_{\bullet k}}{x_{\bullet \bullet}} n_{P,k}$	The number of predators a given taxon has
Global	Mean path length	$\frac{1}{n_p} \sum_p \text{length}(p)$	N/A	The average length of all complete paths (p) through the food web
	Mean trophic level	$\frac{1}{S} \sum_i TL_i$	N/A	The average trophic level of all nodes in the food web
	Mean omnivory index	$\frac{1}{S} \sum_i \sum_j (TL_j - (TL_i - 1))^2 \frac{x_{ij}}{x_{\bullet i}}$	N/A	Index for the average number of omnivorous links per species; details in Kones <i>et al.</i> (2009)
	Modularity	$\frac{1}{2m} \sum_{i,j} (x_{ij} - \frac{k_i k_j}{2m}) I(c_i = c_j)$	N/A	Index for the number of clusters of highly connected species; details in Kones <i>et al.</i> (2009)
	Transitivity	$\frac{1}{s_i(k_i-1)} \sum_{j,h} \frac{x_{ij} + x_{ih}}{2} x_{ij} x_{ih} x_{jh}$	N/A	Index for the transitivity (a measure of clustering) of the food web; details in Kones <i>et al.</i> (2009)

¹In all formulae, $H_{N,k} = -\sum_{i=1}^S \frac{x_{ki}}{x_{\bullet k}} \log_2 \left(\frac{x_{ki}}{x_{\bullet k}} \right)$; $H_{P,k} = -\sum_{j=1}^S \frac{x_{jk}}{x_{\bullet k}} \log_2 \left(\frac{x_{jk}}{x_{\bullet k}} \right)$; $n_{N,k} = \begin{cases} 2^{H_{N,k}} & ; x_{k \bullet} = 0 \\ 0 & ; x_{k \bullet} \neq 0 \end{cases}$; $n_{P,k} = \begin{cases} 2^{H_{P,k}} & ; x_{\bullet k} = 0 \\ 0 & ; x_{\bullet k} \neq 0 \end{cases}$

and all link weights are as large as possible. If connectance is known, maximizing a standard definition of throughput provides no additional information on food web structure. We used a modified definition of food web throughput, trophic-weighted throughput, to link throughput more closely to food web structure (Table 2). Trophic-weighted throughput multiplies each link weight by a factor of ten for each increase in trophic level, which incorporates the c. 10% transfer efficiency commonly used for aquatic organisms (Pauly & Christensen 1995).

There are several possible measures of food web robustness; we used the negative of the average number of secondary extinctions (Table 2). The number of secondary extinctions is the number of species that go extinct following removal of a given node (species). The average number of secondary extinctions is the average of this value over all nodes (i.e. remove each node one at a time and average the recorded numbers of secondary extinctions) (Table 2). We used the negative of this number because we wanted to maximize robustness, which

Table 2. Fitness functions used in the genetic algorithm. The variable x_{ij} is the link weight from species j to species i ; TL is trophic level; S is the number of species; and e_i is the number of species going extinct following removal of species i

Fitness function	Formula	Description
Weighted throughput	$\sum_{i,j} 10^{TL_i} x_{ij}$	Total system throughput weighted by trophic level (TL); this definition implies that higher trophic levels use more energy than lower trophic levels
Robustness (inverse secondary extinctions [SE])	$\frac{-1}{S} \sum_i e_i$	The inverse of the average number of species that go extinct following the removal of species i (e_i). The negative sign is included because maximizing robustness is equivalent to minimizing secondary extinctions (i.e. we wanted to maximize the inverse of SE)
Weighted throughput and Robustness	N/A	Simultaneous maximization of weighted throughput and robustness

involves minimizing the number of secondary extinctions (Table 2).

SIMULATING FOOD WEBS USING A GENETIC ALGORITHM

We used a genetic algorithm (Michalewicz 1996) to simulate food webs that maximized trophic-weighted throughput and robustness (Table 2). Both throughput and robustness were used as independent measures of fitness (single-objective optimization), but we also ran the genetic algorithm with both fitness functions combined (multi-objective optimization; Miettinen 1999).

For each of the 106 empirical food webs, we simulated food webs with the same number of species (S) and the same connectance (C) as the empirical food web, but with throughput, robustness, or both throughput and robustness maximized. We constrained S and C because these quantities have a large influence on food web structure (Riede *et al.* 2010). For a binary empirical food web with S species and connectance equal to C , our method was to (see also Fig. 2):

- 1 Initialize 100 random binary food webs with S species and connectance equal to C .
- 2 Draw random weights for all nonzero links from a uniform distribution on [0, 1].
- 3 Calculate fitness (e.g. throughput, robustness, or both) and retain the 20 food webs with highest fitness.
- 4 Generate 80 new food webs by drawing one of the 20 retained food webs at random and subjecting it to a mutation process:
 - 4.1. Draw the number of links to mutate, N_{change} , from a discrete uniform distribution on $[N_{zero}, N_{positive}]$, where N_{zero}

and $N_{positive}$ are the number of zero and nonzero links, respectively.

- 4.2. Select N_{change} positive links and set them equal to zero and select N_{change} zero links and assign them a weight drawn randomly from a uniform distribution on [0-01, 1].
- 4.3. Assign a new weight to one-third of the positive links (selected randomly from all positive links). The weight was drawn randomly from a uniform distribution on [0-01, 1].
- 5 Repeat steps 3 and 4 for 2000 iterations. Convergence was assessed using plots of fitness against iteration.
- 6 After 2000 iterations, select the 20 food webs with the highest fitness and calculate local and global metrics of food web structure for each. These food webs are the final, simulated maximum-throughput and/or maximum-robustness food webs.
- 7 Calculate the median value of each metric of food web structure for the 20 selected food webs and compare this value to the metric for the real food web.

We followed these same steps for the 21 weighted food webs, but modified steps 1 and 2 to generate 100 random weighted food webs using a null model (see Null model comparison below), and based steps 6 and 7 on the five weighted local metrics rather than all ten local and global metrics (Fig. 2).

Convergence to maximum throughput took an average of 160 out of 2000 iterations (minimum = 20, maximum = 250). Convergence to maximum robustness took an average of 15 out of 2000 iterations (minimum = 1, maximum = 155). Convergence when maximizing throughput and robustness simultaneously took an average of 150 out of 2000 iterations (minimum = 27, maximum = 244). An example of the changes in food web structure through successive iterations of the genetic algorithm is in Fig. S1.

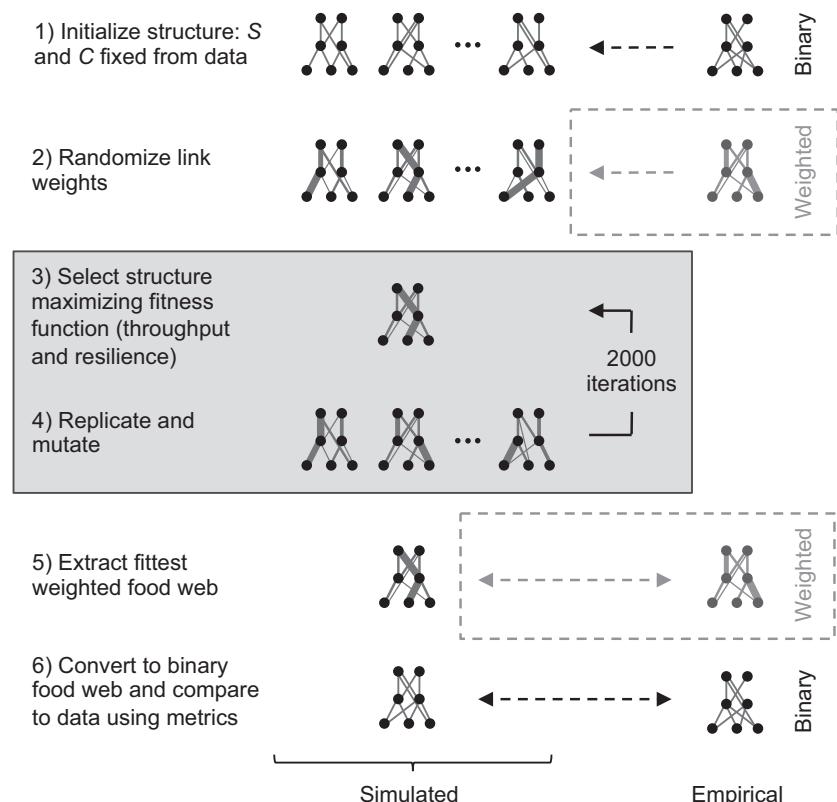


Fig. 2. Schematic of the genetic algorithm method used to simulate binary and weighted food webs maximizing a particular fitness function. Steps 1–6 outline the method for binary food webs, while dashed boxes are alternative start and endpoints when weighted food web data are available. The grey shaded box (steps 3 and 4) is the genetic algorithm (described in detail in *Simulating food webs using a genetic algorithm*, see main text). This general approach could be adapted to study the links between structure and function for more general systems (i.e. systems other than food webs), as long as the function to be optimized is well defined and can be calculated from the relevant measure of structure.

NULL MODEL COMPARISON

For each empirical food web, the genetic algorithm simulated one set of food webs with maximum throughput, one set of food webs with maximum robustness, and one set of food webs with both maximum throughput and robustness. We used a null model to test whether simulated food webs were significantly different to real food webs. This null model was used to generate a null distribution for each of the metrics of food web structure, which estimates the plausible range of variation in each metric for a particular empirical food web. If the metrics for the simulated food webs were outside the 2.5% and 97.5% quantiles of the null distribution, then simulated food webs were considered to differ significantly from the real food webs (Gotelli & Graves 1996).

We used a null model that constrained the number of species S and the connectance C , which matched the constraints on our genetic algorithm. For weighted food webs, this algorithm also constrained the distribution of link weights. Given S and C , we used a swap algorithm to shuffle the trophic interactions (i.e. the nonzero entries in the adjacency matrix) among the species, while constraining the row totals (i.e. the number of species consuming a particular species) and the column totals (the number of species consumed by a particular species) (see Gotelli & Graves 1996). This null model could be applied to binary and weighted food webs and was used to generate 1000 food webs.

SOFTWARE

The genetic algorithm, null models and all processing of outputs were implemented in R 3.1.2 (R Core Team 2014). All code has been compiled into an R package and is included as Supporting Information (Appendix S1, Supporting information; also available at <https://github.com/jdyen/fwga>).

Results

Our results serve to validate our method, which links theoretical ecosystem functions to empirical data. As applied to food web structure, our results largely align with previous work, but there are several notable findings.

Local metrics of food web structure suggested that highly connected, simulated food webs were similar to real food webs, regardless of whether throughput or robustness was being maximized (Fig. 3, see also Figs S2, S3, Supporting information). Where simulated binary food webs differed from real binary food webs (i.e. when connectance was low), simulated food webs tended to have fewer basal and top species, more intermediate species, and lower generality and vulnerability than real food webs (Fig. 3). Simulated weighted food webs tended to be similar to real weighted food webs, with several exceptions (Fig. S4, Supporting information). Five weighted data sets (Rhode river, St Marks, Narragansett, Lake

Michigan and Florida bay [wet season]) had more basal species and fewer intermediate species in simulated than in real food webs, while four weighted data sets (Florida bay [dry season], Graminoid [wet and dry season] and Cypress [dry season]) had higher generality and vulnerability in simulated than in real food webs (Fig. S4, Supporting information). The null model provided narrow bounds around the real food web, suggesting that there is little variation in local metrics if connectance is constrained (Figs 3, S2–S4, Supporting information).

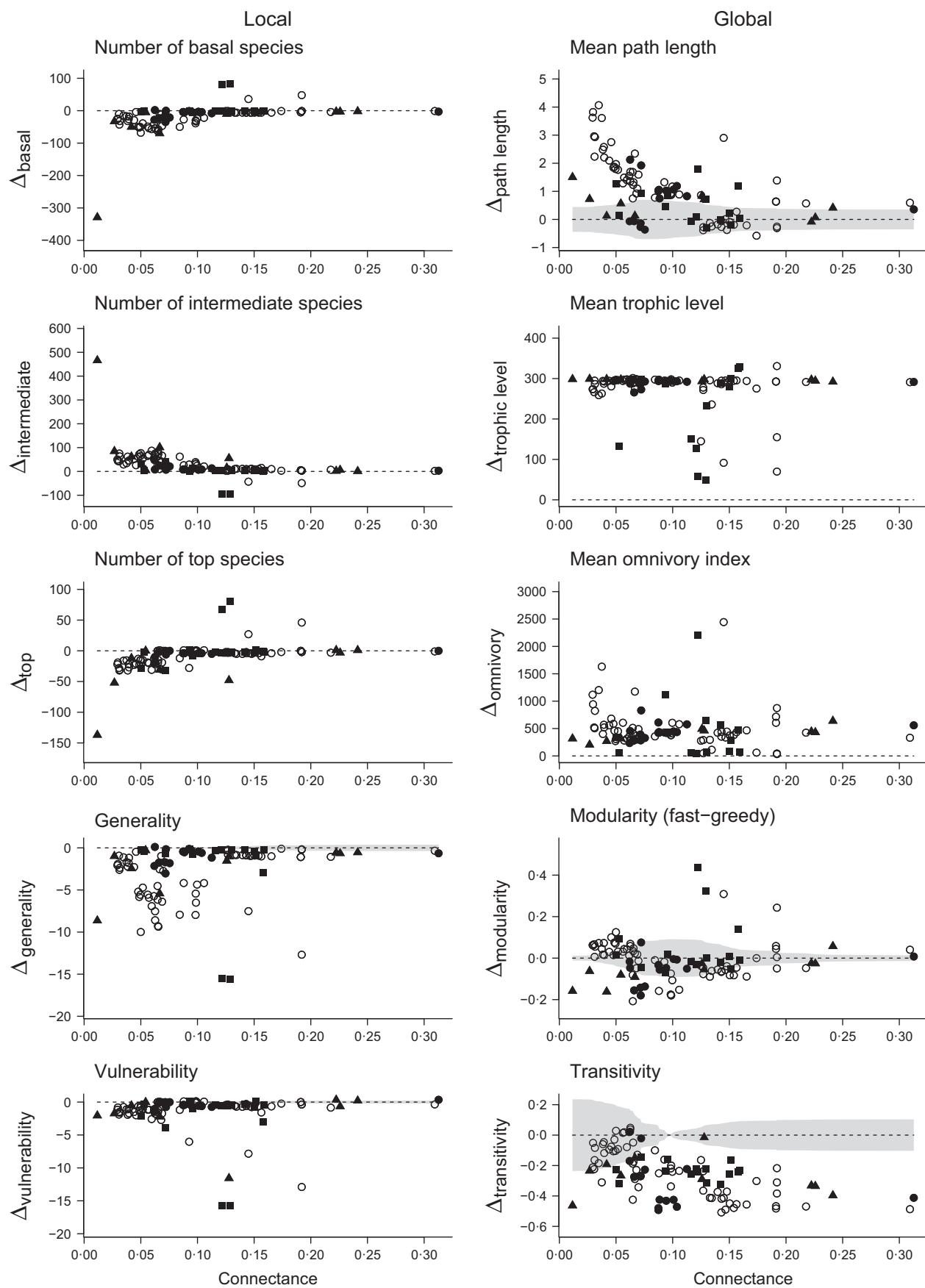
Global metrics of food web structure highlighted clear differences between simulated and real binary food webs (Fig. 3, see also Figs S2, S3, Supporting information). When connectance was low, simulated food webs had longer paths, higher average trophic levels and higher levels of omnivory than real food webs (Figs 3, S2, S3, Supporting information). Simulated food webs tended to have lower levels of transitivity than real food webs, but differences in transitivity were not associated with connectance (Figs 3, S2, S3, Supporting information).

Global metrics of food web structure indicated that maximizing throughput or maximizing robustness simultaneously led to unrealistically high mean trophic levels (mean = 295, standard deviation = 10), high levels of omnivory and lower transitivity than in real food webs (Figs 3, 4). Estimated trophic levels are not a true ‘trophic level’ because of the effect of loops on trophic-level calculations, but the high average trophic levels are representative of the substantial differences between real food webs and simulated food webs with maximum throughput or maximum robustness.

Maximizing throughput or maximizing robustness simultaneously generated similar food web structures (Figs 3, S2, Supporting information). In contrast, maximizing robustness alone generated food web structures that were similar to real food webs when connectance was high (Figs 4, S3, Supporting information). In particular, maximum-robustness food webs had more realistic average trophic levels than maximum-throughput food webs (mean trophic level = 9, standard deviation = 26; c.f. mean = 5, standard deviation = 9 for real food webs) (Figs 4, S2, Supporting information).

The relationship between simulated and real food webs did not depend on ecosystem type (Fig. 3). Although many of the simulated food webs that differed from real food webs were based on marine and freshwater ecosystems, this result is confounded by connectance (e.g. all marine food webs had low connectance; Fig. 3). When connectance was high, simulated maximum-robustness food webs appeared similar to real food webs regardless of ecosystem type (Fig. S3, Supporting information).

Fig. 3. Local (left column) and global (right column) metrics comparing real food webs and food webs simulated with maximum throughput and robustness. Δ_{metric} (y-axis) is the difference between simulated and real food webs for a given metric (real food web is the zero line, Δ values are based on an average of 20 food webs). Grey shading is the null model range around the fitted food web (mean distance of null food web models from the real food webs – projected symmetrically). Freshwater (open circles), estuarine (closed squares), marine (closed triangle) and terrestrial (closed circles) food webs are shown.



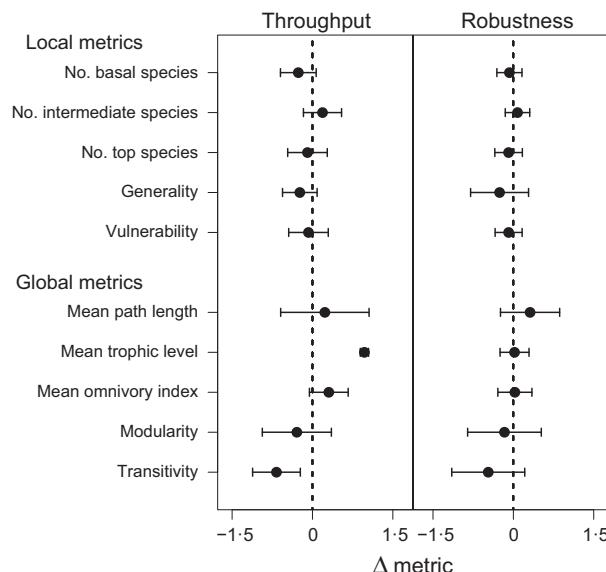


Fig. 4. Differences between simulated and real food webs for local and global metrics, based on food webs with connectance >0.1 . Results for maximum-throughput (left column) and maximum-robustness (right column) food webs are shown. Δ metric is the difference between simulated and real food webs (real food web is the zero line), standardized by the maximum Δ value for a given metric. Points are mean Δ values, and bars denote 95% confidence intervals around the mean value.

Discussion

Our study suggests that food web structure can be linked to ecosystem functioning and resilience. In particular, we found that highly connected food webs maximize resilience but do not maximize throughput. These results indicate that the maximization of different ecological functions can generate distinct food web structures, which could be used to test proposed links between food web structure and function. Two key features of food webs were revealed: (i) topological food web robustness is linked strongly to connectance; and (ii) ecosystem throughput does not appear to be maximized in real food webs. Our study also highlights that global metrics of food web structure (i.e. those based on properties of the entire food web) were better at distinguishing different food web structures than were local metrics of food web structure (i.e. those based on species–species links).

We found that maximizing different ecological functions generated distinct food web structures. Although the maximization of robustness generated realistic food webs when connectance was high, the simultaneous maximization of robustness and throughput generated food webs that differed markedly from real food webs. This result suggests that maximization of trophic-weighted throughput is unlikely to occur in real food webs, even when constrained by robustness. A broader exploration of constraints (e.g. direct constraints on the number of trophic levels) or alternative definitions of throughput might determine whether maximization of ecosystem func-

tioning can predict observed food web structures. The use of constraints or more-nuanced definitions of ecosystem functioning would allow tests of complex hypotheses about food web function. For example, the genetic algorithm could quantify ecosystem functioning with dynamic, size-structured food web models (e.g. Binzer *et al.* 2016), which would link food web structure to a range of ecological functions. Irrespective of the exact definition of ecosystem functioning, our approach could be used more broadly to link structure and function in food webs.

While several studies have used genetic algorithms to study food webs (e.g. Ruiz-Moreno, Pascual & Riolo 2006; Allesina & Pascual 2009; Eklöf *et al.* 2013), to our knowledge, our study is the first to use a genetic algorithm to explore the relationship between food web structure and ecosystem functioning. The genetic algorithm mimics the process of natural selection, but it cannot be used to investigate how natural selection might maximize ecosystem functioning or robustness. The process of evolution followed by the genetic algorithm is not a realistic evolutionary assembly process for real food webs, since entire food webs cannot be thought to compete with one another or evolve through large recombination events. Nonetheless, the genetic algorithm is a convenient and adaptable method that estimates near-optimal food web structures. Our genetic algorithm approach generated distinct food web structures when different ecological functions were maximized, so we believe this to be a useful method for exploring how different ecological functions are related to food web structure.

The observation of many weak links in real food webs (McCann, Hastings & Huxel 1998; Neutel, Heesterbeek & de Ruiter 2002; Rooney *et al.* 2006) might explain why maximum robustness was more apparent than maximum throughput in the food webs studied here. A mixture of strong and weak links has been shown to drive stability and robustness in food webs, with a combination of fast and slow energy channels conferring robustness on food webs (McCann, Hastings & Huxel 1998; Rooney *et al.* 2006). In contrast, these weak links directly reduce throughput from its maximum possible value. We used trophic-weighted throughput in our study, which is a more nuanced definition of throughput (higher trophic levels require more energy than lower trophic levels due to inefficient energy conversion among species). However, maximizing this definition of throughput led to an unrealistically high average trophic level, which is not observed in real food webs. It is possible that alternative definitions of throughput could yield more realistic food webs; the applicability of our method to weighted food webs supports further tests of maximum-throughput hypotheses. Alternatively, additional constraints on food web structure, such as restricting the number of closed loops, might lead to more realistic food web structures.

Our results suggest that more highly connected food webs are likely to be more robust, which is consistent with previous studies (Dunne, Williams & Martinez 2002b;

Eklöf & Ebenman 2006; but see Vieira & Almeida-Neto 2015). While several studies have observed relationships between connectance and robustness (e.g. Dunne, Williams & Martinez 2002b), few have explored reasons for this link (but see Petchey, Brose & Rall 2010; Thébaud & Fontaine 2010). Maximized robustness in highly connected food webs might be related to the overall level of interaction within a food web. High connectance typically is associated with short chains and with high numbers of generalist species, which would be expected to lead to fewer secondary extinctions than if food webs had long chains and many specialist species (Dunne, Williams & Martinez 2002a; Kortsch *et al.* 2015).

Local metrics of structure were strikingly similar across all simulated food webs, despite markedly different structures in maximum-throughput and maximum-robustness food webs. This result suggests that local metrics have little power to distinguish different food web structures, at least when connectance is known (i.e. these metrics are correlated with connectance). The link between local food web metrics and connectance has been observed previously (Vermaat, Dunne & Gilbert 2009), and our results reinforce the notion that, if connectance is known, other local metrics may not provide much information about food web structure. Modifying local metrics to account for link weights distinguished different food web structures better than binary metrics, which suggests that link-weight distributions might be a useful measure (Bersier, Banašek-Richter & Cattin 2002). Global metrics of food web structure were better at distinguishing different structures than local metrics. In particular, the related measures of mean trophic level and the level of omnivory were able to distinguish maximum-throughput and maximum-robustness structures.

While we were able to distinguish food webs with maximized throughput and robustness, we did not find clear differences among food webs from different ecosystem types. Marine and freshwater food webs appeared less similar to optimized food webs in some cases, but this pattern was confounded by low connectance for many of these food webs (Fig. 3). Based on our global food web metrics, different ecosystem types appeared to have similar levels of robustness and throughput (for a given connectance), suggesting that connectance might be a more important driver of food web structure and function than ecosystem type (but see Digel *et al.* 2014). This result supports the notion that there may be general rules underlying food web structure and indicates that these rules might be linked more closely to connectance than to particular ecosystems.

Our genetic algorithm approach is general and could be used to explore links between structure and function in systems other than food webs. For example, hypotheses about ecosystem function (e.g. maximum productivity) could be used to predict community structure (e.g. size distributions; Yen *et al.* 2014, 2015). More broadly, our approach could be used to relate structure and function

in a range of biological networks (e.g. gene-regulation networks, biochemical networks, mutualistic networks, social networks) (Bascompte *et al.* 2003; Gunawardena 2012; Smith *et al.* 2015). The only requirements are that the function to be maximized can be defined clearly and that the predicted structure can be compared with empirical structures. Our software (Appendix S1, Supporting information) could be applied directly to any system where structure can be represented as a matrix (e.g. site-by-species community matrices) and where ecosystem function can be calculated directly from this matrix (e.g. production as a weighted sum of species abundances). Broad application of our approach would provide substantial insight into the relationship between structure and function in a range of biological systems and might strengthen links among different fields of biology (e.g. organismal biology, food web ecology and community ecology).

We have introduced a general method for assessing whether structure is related to function in ecological systems. We found that maximizing different ecological functions (e.g. throughput and robustness) led to distinct food web structures, which suggests that there are predictable links between structure and function. Maximizing trophic-weighted throughput did not generate realistic food web structures, which suggests that simple measures of throughput are not maximized in real food webs. In contrast, maximizing robustness generated food web structures similar to those observed in real food webs, which suggests that ecosystem resilience might be maximized in real food webs. Broadly, our genetic algorithm approach provides a framework for predicting structure from different ecosystem functions and could be used to test a range of hypotheses about the links between structure and function in ecological systems.

Acknowledgements

We thank Jennifer Dunne and Swati Patel for insightful comments, Luís Seone for assistance with the implementation of the genetic algorithm and the Santa Fe Institute Complex Systems Summer School 2013 for facilitating this work. Ulrich Brose and Charles Novae de Santana provided empirical food web data. We acknowledge the following organizations for their support: JDLY – Monash University, the Sir James McNeill Foundation and the Victorian Life Sciences Computation Initiative; RC – National Institute of Physics, College of Science, UP Diliman and the Santa Fe Institute; MC – CNPq Brazil, Killam Trusts Canada and the Cetacean Society International; SK – the Troms Science Foundation and the Arctic University of Norway, Tromsø; JP – MARE – Marine and Environmental Sciences Centre and the University of Coimbra, Portugal; MY – the Japan Society for the Promotion of Science (JSPS) Postdoctoral Fellowship for Research Abroad (24-869).

Data accessibility

Food web data are listed in Tables S1 and S2 and were extracted from the GlobalWeb data base (<http://globalwebdb.com>; Thompson *et al.* 2012), Appendix 1 in Riede *et al.* (2010), Raymond *et al.* (2011), Planque *et al.* (2014) and the igraphdata R package (Csardi 2015).

Genetic algorithm software has been collated into an R package, and package binaries have been uploaded as online supporting information.

References

- Allesina, S. & Pascual, M. (2009) Googling food webs: can an eigenvector measure species' importance for coextinctions? *PLoS Computational Biology*, **5**, e1000494.
- Bascompte, J., Jordano, P., Melián, C.J. & Olesen, J.M. (2003) The nested assembly of plant-animal mutualistic networks. *Proceedings of the National Academy of Sciences USA*, **100**, 9383–9387.
- Bersier, L.-F., Banašek-Richter, C. & Cattin, M.-F. (2002) Quantitative descriptors of food-web matrices. *Ecology*, **83**, 2394–2407.
- Binzer, A., Guill, C., Rall, B.C. & Brose, U. (2016) Interactive effects of warming, eutrophication and size-structure: impacts of biodiversity and food-web structure. *Global Change Biology*, **22**, 220–227.
- Costanza, R. & Mageau, M. (1999) What is a healthy ecosystem? *Aquatic Ecology*, **33**, 105–115.
- Csardi, G. (2015) igraphdata: A collection of network data sets for the ‘igraph’ package. *R package version 1.0.1*. URL: <http://CRAN.R-project.org/package=igraphdata>
- De Angelis, D.L. (1975) Stability and connectance in food web models. *Ecology*, **56**, 238–243.
- DeLong, J.P. (2008) The maximum power principle predicts the outcomes of two-species competition experiments. *Oikos*, **117**, 1329–1336.
- Digel, C., Curtsdotter, A., Riede, J., Klärner, B. & Brose, U. (2014) Unravelling the complex structure of forest soil food webs: higher omnivory and more trophic levels. *Oikos*, **123**, 1157–1172.
- Dunne, J.A., Williams, R.J. & Martinez, N.D. (2002a) Food-web structure and network theory: the role of connectance and size. *Proceedings of the National Academy of Sciences USA*, **99**, 12917–12922.
- Dunne, J.A., Williams, R.J. & Martinez, N.D. (2002b) Network structure and biodiversity loss in food webs: robustness increases with connectance. *Ecology Letters*, **5**, 558–567.
- Eklöf, A. & Ebenman, B. (2006) Species loss and secondary extinctions in simple and complex model communities. *Journal of Animal Ecology*, **75**, 239–246.
- Eklöf, A., Jacob, U., Kopp, J., Bosch, J., Castro-Urgal, R., Chacoff, N.P. et al. (2013) The dimensionality of ecological networks. *Ecology Letters*, **16**, 577–583.
- Gascuel, D., Morissette, L., Palomares, M.L.D. & Christensen, V. (2008) Trophic flow kinetics in marine ecosystems: toward a theoretical approach to ecosystem functioning. *Ecological Modelling*, **217**, 33–47.
- Gotelli, N.J. & Graves, G.R. (1996) *Null Models in Ecology*. Smithsonian Institution Press, Washington, DC.
- Gunawardena, J. (2012) A linear framework for time-scale separation in nonlinear biochemical systems. *PLoS ONE*, **7**, e36321.
- Holling, C.S. (1973) Resilience and stability of ecological systems. *Annual Review of Ecology and Systematics*, **4**, 1–23.
- del Jesus, M., Foti, R., Rinaldo, A. & Rodriguez-Iturbe, I. (2012) Maximum entropy production, carbon assimilation, and the spatial organization of vegetation in river basins. *Proceedings of the National Academy of Sciences USA*, **109**, 20837–20841.
- Kleidon, A., Malhi, Y. & Cox, P.M. (2010) Maximum entropy production in environmental and ecological systems. *Philosophical Transactions of the Royal Society B. Biological Sciences*, **365**, 1297–1302.
- Kones, J.K., Soetaert, K., van Oevelen, D. & Owino, J.O. (2009) Are network indices robust indicators of food web functioning? A Monte Carlo approach. *Ecological Modelling*, **220**, 370–382.
- Kortsch, S., Primicerio, R., Fossheim, M., Dolgov, A.V. & Aschan, M. (2015) Climate change alters the structure of arctic marine food webs due to poleward shifts of boreal generalists. *Proceedings of the Royal Society B. Biological Sciences*, **282**, 20151546.
- Link, J. (2002) Does food web theory work for marine ecosystems? *Marine Ecology Progress Series*, **230**, 1–9.
- Loreau, M. (2010) *From Populations to Ecosystems: Theoretical Foundations for a New Ecological Synthesis*. Princeton University Press, Princeton, NH, USA.
- Martinez, N.D. (1992) Constant connectance in community food webs. *American Naturalist*, **139**, 1208–1218.
- May, R.M. (1972) Will a large complex system be stable? *Nature*, **238**, 413–414.
- May, R.M. (1973) Qualitative stability in model ecosystems. *Ecology*, **54**, 638–641.
- McCann, K., Hastings, A. & Huxel, G.R. (1998) Weak trophic interactions and the balance of nature. *Nature*, **395**, 794–798.
- Michalewicz, Z. (1996) *Genetic Algorithms+ Data Structures = Evolution Programs*. Springer-Verlag, New York, NY, USA.
- Miettinen, K. (1999) *Nonlinear Multiobjective Optimization*. Kluwer Academic Publishers, Boston, MA, USA.
- Neutel, A.M., Heesterbeek, J.A.P. & de Ruiter, P.C. (2002) Stability in real food webs: weak links in long loops. *Science*, **296**, 1120–1123.
- Opitz, S. (1996) *Trophic interactions in Caribbean coral reefs*. ICLARM Technical Report 43. International Center for Living Aquatic Resources Management, Makati City, Philippines.
- Pauly, D. & Christensen, V. (1995) Primary production required to sustain global fisheries. *Nature*, **374**, 255–257.
- Petchey, O.L., Brose, U. & Rall, B.C. (2010) Predicting the effects of temperature on food web connectance. *Philosophical Transactions of the Royal Society B. Biological Sciences*, **365**, 2081–2091.
- Planque, B., Primicerio, R., Michalsen, K., Aschan, M., Certain, G., Dalpadado, P. et al. (2014) Who eats whom in the Barents Sea: a food web topology from plankton to whales. *Ecology*, **95**, 1430.
- R Core Team (2014) *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. URL: <http://www.R-project.org/>
- Raymond, B., Marshall, M., Nevitt, G., Gillies, C.L., van den Hoff, J., Stark, J.S. et al. (2011) A Southern Ocean dietary database. *Ecology*, **92**, 1188.
- Riede, J.O., Rall, B.C., Banasek-Richter, C., Navarrete, S.A., Wieters, E.A., Emmerson, M.C. et al. (2010) Scaling of food-web properties with diversity and complexity across ecosystems. *Advances in Ecological Research*, Vol. 42 (ed G. Woodward), pp. 139–170. Academic Press, Burlington, MA, USA.
- Rooney, N., McCann, K., Gellner, G. & Moore, J.C. (2006) Structural asymmetry and the stability of diverse food webs. *Nature*, **442**, 265–269.
- Ruiz-Moreno, D., Pascual, M. & Riolo, R. (2006) Exploring network space with genetic algorithms: modularity, resilience, and reactivity. *Ecological Networks – Linking Structure to Dynamics in Food Webs* (eds M. Pascual & J.A. Dunne), pp. 187–208. Oxford University Press, New York, USA.
- Smith, C., Pechuan, X., Puzio, R.S., Biro, D. & Bergman, A. (2015) Potential unsatisfiability of cyclic constraints on stochastic biological networks biases selection towards hierarchical architectures. *Journal of the Royal Society Interface*, **12**, 20150179.
- Stouffer, D.B. & Bascompte, J. (2010) Understanding food-web persistence from local to global scales. *Ecology Letters*, **13**, 154–161.
- Thébault, E. & Fontaine, C. (2010) Stability of ecological communities and the architecture of mutualistic and trophic networks. *Science*, **329**, 853–856.
- Thompson, R.M. & Townsend, C.R. (2004) Land-use influences on New Zealand stream communities: effects on species composition, functional organisation, and food-web structure. *New Zealand Journal of Marine and Freshwater Research*, **38**, 595–608.
- Thompson, R.M., Brose, U., Dunne, J.A., Hall, R.O. Jr, Hladz, S., Kitching, R.L. et al. (2012) Food webs: reconciling the structure and function of biodiversity. *Trends in Ecology & Evolution*, **27**, 689–697.
- Vermaat, J.E., Dunne, J.A. & Gilbert, A.J. (2009) Major dimensions in food-web structure properties. *Ecology*, **90**, 278–282.
- Vieira, M.C. & Almeida-Neto, M. (2015) A simple stochastic model for complex coextinctions in mutualistic networks: robustness decreases with connectance. *Ecology Letters*, **18**, 144–152.
- Yen, J.D.L., Paganin, D.M., Thomson, J.R. & Mac Nally, R. (2014) Thermodynamic extremization principles and their relevance to ecology. *Austral Ecology*, **39**, 619–632.
- Yen, J.D.L., Paganin, D.M., Thomson, J.R. & Mac Nally, R. (2015) Thermodynamics predicts density-dependent energy use in organisms and ecological communities. *Physical Review E*, **91**, 042708.

Received 10 May 2015; accepted 26 November 2015

Handling Editor: Eoin O'Gorman

Supporting Information

Additional Supporting Information may be found in the online version of this article.

Table S1. Complete list of binary data sets used in analyses.

Table S2. Complete list of weighted data sets used in analyses.

Fig. S1. Example of food-web development through successive iterations of the genetic algorithm.

Fig. S2. Plots of local and global metrics comparing real food webs and food webs simulated with maximum throughput.

Fig. S3. Plots of local and global metrics comparing real food webs and food webs simulated with maximum robustness.

Fig. S4. Plots of local metrics comparing real weighted food webs and food webs simulated with maximum throughput or maximum robustness.

Appendix S1. R package containing code for genetic algorithm software (built and tested using R version 3.1.2).