

## SI Appendix

### 1. ATN Model

The dynamic allometric trophic network (ATN) models are built in a four step process: (1) the food-web structure is created by a structural model, (2) the trophic levels and body masses of the populations are calculated, (3) the parameters of an allometrically-scaled predator-prey model are calculated, and (4) a producer-nutrient model is randomly parameterized.

**(1) Food-web structure.** The network structure ("who eats whom") of the food webs follows the niche model (1). This stochastic model is based on algorithms that arrange trophic links,  $L$ , among species based on species richness,  $S$ , and connectance,  $C = L/S^2$ , as input parameters. The food-web structures predicted by the niche model have been successfully tested against empirical data (2).  $S$  and  $C$  were sampled from uniform distributions ( $10 \leq S \leq 30$ ,  $0.1 \leq C \leq 0.2$ ).

**(2) Trophic levels and body masses.** The feeding matrix created by the niche model was used to calculate the trophic levels of the species in the food webs as prey-averaged trophic levels (3). We calculated species' body masses starting with a basal species level of unity. The body masses of predators of successively higher trophic levels are calculated using average predator-prey body-mass ratios,  $Z$ , sampled from a log-normal distribution (mean = 10, SD = 100) for each food web independently. The body masses of predators,  $M_C$ , depend on trophic levels,  $T$ :

$$M_C = Z^{T-1}. \quad (1)$$

**(3) Allometric predator-prey model.** The population dynamics within these food webs follow an allometric predator-prey model (4, 5), where:

$$B_i' = r_i G_i(N) B_i - x_i B_i - \sum_{j=\text{consumers}} x_j y B_j F_{ji} / e_{ji} \quad (2a)$$

$$B_i' = -x_i B_i + \sum_{j=\text{resources}} x_i y B_i F_{ij} - \sum_{j=\text{consumers}} x_j y B_j F_{ji} / e_{ji} \quad (2b)$$

describe changes in relative, dimensionless biomass densities of primary producer (eq. 2a) and consumer species (eq. 2b). In these equations,  $B_i$  is the biomass density of population  $i$ ,  $r_i$  is  $i$ 's mass-specific maximum growth rate,  $G_i$  is  $i$ 's nutrient-dependent growth rate when  $i$  is a basal producer (see producer-nutrient model below),  $x_i$  is  $i$ 's mass-specific metabolic rate,  $y$  is the maximum consumption rate of the consumers relative to their metabolic rate,  $e_{ji}$  is  $j$ 's assimilation efficiency when consuming population  $i$ . The functional response,  $F_{ij}$ , describes the realized fraction of  $i$ 's maximum rate of consumption achieved when consuming species  $j$ :

$$F_{ij} = \frac{\omega_{ij} B_j^h}{B_0^h + c B_i B_0^h + \sum_{k=\text{resources}} \omega_{ik} B_k^h}, \quad (3)$$

where  $\omega_{ij}$  is the proportion (0-1) of  $i$ 's maximum consumption rate targeted to consuming  $j$  (hereafter: “preference”),  $B_0$  is the half saturation density of  $i$  ( $B_0=0.5$ ),  $h$  is the Hill-exponent which regulates the shape of the curve from Holling Type II to Holling Type III (6, 7), and  $c$  quantifies predator interference (6). The predator interference term in the denominator quantifies the degree to which individuals within population  $i$  interfere with one another's consumption activities, which reduces  $i$ 's per capita consumption if  $c > 0$  (8, 9). The  $F_{ij}$  can continuously vary between type II ( $h=1$ ,  $c=0$ ), type III ( $h=2$ ,  $c=0$ ), type II with predator interference ( $h=1$ ,  $c=1$ ) and type III with predator interference ( $h=2$ ,  $c=1$ ). We used uniform preferences of consumers with  $n$  resources ( $\omega_{ij} = 1/n$ ) – that is, consumers do not have an active prey preference, but rather feed according to the relative biomasses of their prey species. The per unit

biomass biological rates of production,  $R$ , metabolism,  $X$ , and maximum consumption,  $Y$ , follow negative-quarter power-law relationships with the species' body masses (10):

$$R_P = a_r M_P^{-0.25} . \quad (4a)$$

$$X_{C,P} = a_x M_C^{-0.25} , \quad (4b)$$

$$Y_C = a_y M_C^{-0.25} , \quad (4c)$$

where  $a_r$ ,  $a_x$  and  $a_y$  are allometric constants,  $M$  is the average body mass of individuals within the population, and the subscripts  $C$  and  $P$  indicate consumer and producer parameters, respectively (5). The time scale of the system is defined by normalizing the biological rates (4a-c) to the mass-specific growth rate of the basal population, and the maximum consumption rates are normalized by the metabolic rates:

$$r_i = 1, \quad (5a)$$

$$x_i = \frac{X_C}{R_P} = \frac{a_x}{a_r} \left( \frac{M_C}{M_P} \right)^{-0.25} , \quad (5b)$$

$$y_i = \frac{Y_C}{X_C} = \frac{a_y}{a_x} , \quad (5c)$$

Inserting equations 5a-c into equations 2a-b yields a population dynamic model with allometrically scaled parameters depending on the species' body masses (see above).

We used the following constant model parameters:  $y_i = 8$ ;  $e_{ij} = 0.85$  for carnivores and  $e_{ij} = 0.45$  for herbivores;  $a_x/a_r = 0.314$  for invertebrates, and  $a_x/a_r = 0.138$  for producers (4). We used random, normally distributed Hill exponents (mean = 1.5, SD = 0.25) and predator interference terms (mean = 0.5, SD = 0.25).

**(4) Producer-nutrient model.** Consistent with previous studies (11, 12), the growth of the producer species follows a well-established nutrient intake model:

$$G_i(N) = \text{MIN} \left( \frac{N_1}{K_{1i} + N_1}, \frac{N_2}{K_{2i} + N_2} \right) \quad (6)$$

that depends on the concentrations of two limiting nutrients  $N_l$ . This model has been widely used in theoretical plant ecology (13, 14) and successfully evaluated in experiments (15). In equation (6), MIN is the minimum operator, and  $K_{li}$  is species  $i$ 's half saturation density for nutrient  $l$ . Lower half saturation densities define higher nutrient-intake efficiencies. Therefore,  $G_i(N)$  follows a Monod Equation and is determined by the nutrient that is most limiting. The variation in the density of nutrient  $l$  is given by

$$N_l'(t) = D(S_l - N_l) - \sum_{i=1}^n (c_{li} r_i G_i(N) B_i), \quad (7)$$

where  $c_{li}$  is the content of nutrient  $l$  in the biomass of species  $i$ . Nutrients are exchanged at a turnover rate  $D$  – that is expressed relative to time scale of the growth rate of the producer species – with a supply concentration of  $S_l$ , and removal depends on the current nutrient concentration in the system,  $N_l$ . Throughout all simulations, the turnover rate ( $D = 0.25$ ) and the supply concentrations ( $S_l = 1$ ) were kept constant, and the first nutrient was the one most needed by all producer species as it has the highest content in their biomasses ( $c_{1i}=1$  and  $c_{2i}=0.5$ ). Since all producer species have similar  $r_i$  and  $x_i$ , the half saturation densities ( $K_{li}$ ) for the first nutrient define the competitive hierarchy amongst the producers: the nutrient-intake efficiency decreases with the half saturation density. We created a random competitive hierarchy amongst producer species by randomly varying the half-saturation densities determining their nutrient uptake uniformly between 0.1 and 0.2.

## 2. Simulations

Numerical integration of equations 2a,b used the Bulirsch-Stoer method with adaptive step sizes (16) and started with random initial biomass densities  $B_i$  selected independently, uniformly from the interval (0.05, 1). This algorithm was implemented in C++ and was chosen for its accuracy and computational efficiency; the results are not sensitive to the choice of integration algorithm. Each network was simulated once with all initial species present (Control), and once with each species removed independently, a total of (species present)+1 times. In each simulation, the biomass of each species was averaged from time step 50 to 200 (see Section 6 below) to yield a “removal matrix” for all species showing how each species’ time-averaged biomass differs with and without each of the other species. Many measures of this change have been used (17, 18). Here we use the difference in  $T$  biomass with vs. without  $R$  for two reasons: 1) When using a proportional change, such as a commonly used log ratio  $\log(B_T^+/B_T^-)$ , where  $B_T^+$  is the biomass of  $T$  with  $R$  present, and  $B_T^-$  is the biomass of  $T$  without  $R$  (17, 18), biomass densities  $\ll 1$  can cause interaction strength patterns to be dominated by very small changes in the biomass of extremely rare species. This disproportionate influence of small numbers can be alleviated by adding a constant to the numerator and denominator, but the choice of that constant influences the resulting patterns. If the constant is large relative to the mean observed biomasses, the log ratio approaches the simple difference between  $T$  biomass with vs. without  $R$ . Since our simulated biomass densities were on a relative scale with most being  $\ll 1$ , we had no *a priori* reason for choosing one constant over another. 2)  $I$  calculated as the difference of  $T$  biomass with vs. without  $R$  is transparent and can be used to derive other measures. See main text for results of randomization tests that generate an expected null pattern of  $I$  compared to our simulation results.

Species with biomass density less than  $10^{-30}$  after 200 time steps were considered extinct. Species that went extinct in the control runs were excluded from analysis. All 600 networks were persistent over 200 time steps: 16 networks lost one species, and 1 network lost 2 species in the control runs, yielding measurements of over 254,032 species interactions among 12,116 species (See “Non-equilibrium population dynamics and the time frame of the simulations” for results of longer time series).

### 3. Data Analysis

We tracked four broad categories of explanatory variables. **1) Global web structure** (19): initial and final species richness; initial and final connectance; number and proportion of top (*t*), intermediate (*i*), and basal (*b*) species; number and proportion of herbivores, carnivores, and omnivores; number of links between *t-i*, *t-b*, *i-b* and *i-i*, the total number of links and number of links per species; the mean, maximum, and standard deviation of the web’s resource-averaged trophic level (3), the mean, maximum, and standard deviation of shortest chain between each species and a basal species; and the clustering coefficient of the web (20). **2) Local network structure around *R* and *T***: resource-averaged trophic level and the number, total biomass, mean vulnerability (i.e., number of consumer species), and mean generality (i.e., number of resource species) of both consumer and resource species that are one, two, and three degrees from the *R* and *T* of each interaction. **3) *R* and *T* attributes**: mean biomass before *R* removal, mean body mass, consumer functional response shape (Hill coefficient), consumer interference, and the relative half saturation concentration of nutrient uptake for producers. **4) Attributes of the *R-T* pair**: degrees of separation, “simple” vs. “diffuse” connections (i.e., single vs. multiple paths from *R* to *T*), the net sign of all shortest paths (SP) and next-shortest paths (NSP) between *R* and *T*, the sum  $SP + NSP$ , and the weighted sum  $SP + NSP/2$ . The sign of each path of a given length

was determined by multiplying the signs of each link along the path, where all consumer-to-resource links are negative and resource-to-consumer links are positive. For diffuse connections, the signs of multiple paths of the same length were summed (+1 for positive and -1 for negative effects) (21). These variables were used to predict the sign and magnitude of the interaction between  $R$  and  $T$ . Both positive and negative  $I$  and *per capita I*, as well as the biomass, body mass, and density of  $R$  and  $T$  were all approximately log-normally distributed (e.g., Fig. S1).

To model  $I$ , we applied CART (22), a nonlinear modelling algorithm, to data from a random sample of half the webs (training data). Because the distribution of  $|I|$  is roughly lognormal (Fig. S3), the quadratic prediction error that CART seeks to minimize (by recursively “splitting” the data) is dominated by the tails of the distribution of  $I$ .

Therefore, we allowed CART to split the data until a terminal leaf captured points with a mean  $I$  within the 95th percentile of  $|I|$ . In the resulting tree ( $R^2 = 0.86$ )  $|I|$  and the sign of  $I$  were predicted by different sets of variables, and these results were used to guide subsequent analyses on log-transformed  $|I|$ . The models were then tested on the other 300 webs (test data). When modelling *per capita I*, we had an *a priori* interest in understanding how body size, which drives *per capita* maximum consumption rates for all species, explains variation in *per capita I* in a realistically complex community. We first used Reduced Major Axis regression to quantify the relationship between  $\log|per\ capita\ I|$  and  $\log(M_R)$ , where  $M_R$  is the body mass of  $R$ . To examine the influence of the network on this relationship, we first examined simply connected one degree effects of specialist consumers on resources that have only one consumer. We then included all diffuse one degree effects of consumers on their resources, and finally all possible interactions in the community. We then used CART to explore which of the remaining variables explain the residuals of the  $\log_{10}|per\ capita\ I|$  vs.  $\log(M_R)$  regression. These additional variables were combined with  $\log(M_R)$  in a multiple linear regression to

predict  $\log|per\ capita\ I|$ . As before, the models were built using half the webs, selected at random, and tested on the other half.

#### 4. Application to Field Data

The empirical data re-analyzed here measure the effects of predatory whelks, *Nucella emarginata*, on one of their primary resources, the mussel *Mytilus edulis*, both in the presence and absence of an alternative resource, the acorn barnacle *Balanus glandula* on the rocky, central coast of Oregon, USA, in mid-intertidal disturbance patches (23, 24). We quantified *per capita* and population  $I$  between whelks ( $R$ ) and mussels ( $T$ ) with barnacles present and absent, for each of two whelk biomass enclosure levels (high and low). To compare experimental results with our simulations, separate interaction strengths were estimated for each year of each of the three experiments by using the time-averaged mussel biomass in each replicate of each treatment. Replicates and experimental runs were spatially and/or temporally independent, and there was no consistent correlation in the time-averaged  $I$  across years ( $r$  ranged from 0.13 to -0.78). Thus, the experiment provided repeated estimates of interaction strength across natural variation in mussel and barnacle recruitment with space and time. Since all plots were initially scraped, data from the transient colonization phase were excluded from the analysis. Mussel biomass data used to predict  $I$  were from plots that enclosed whelks, which we considered analogous to the ‘ $R$  present’ condition of our simulation data. To compare the empirical results with our simulations, we converted the empirical data to relative biomass density and used the linear regression model that predicts *per capita*  $I$  in the ATN simulations:  $\log_{10}|per\ capita\ I| = -1.14 + (0.88 * \log_{10}(M_R)) + (0.71 * \log_{10}(B_T^+)) - (0.79 * \log_{10}(B_R))$  (Fig. 2c). One percent cover of mussels in the 400 cm<sup>2</sup> plots was the minimum biomass density unit. Mussels generally formed a mono-layer, so we assumed their biomass increased linear with cover. We used photos of the experimental plots to estimate that one percent mussel cover for this plot size was

approximately three young mussels and that the body size ratio of whelks to mussels was approximately two. In the treatments that included non-trophic interactions mediated through barnacles, the difference between the model predictions and the data were regressed against total barnacle cover in the paired ‘barnacle recruitment’ plots.

## 5. Macroecological patterns that emerged from the simulations.

Our food-web simulations draw  $\log_{10}$  body-mass ratios from a normal distribution (mean = 1, standard deviation (SD) = 2), which caused an increase in average body masses with the trophic levels (Fig S2a). This pattern is consistent with the most recent body mass data for natural food webs (25). Our food-web simulations produced a negative power law between mean species density and body mass, with an exponent of -1.38 (Fig. S2b, dotted line). The power law association was stronger for species with high biomass, where the exponent, -1.05 (Fig. S2b, red line), is generally consistent with a slope of -1 predicted by macroecological theory when multiple trophic levels are considered (10). The log of species’ mean density decreased with increasing trophic level ( $R^2 = 0.56$ ), while log of species’ mean biomass had no association with trophic level ( $R^2 = 0.07$ ). Positive and negative *per capita* as well as population *I* were all approximately log-normally distributed (Fig. S1).

## 6. Non-equilibrium population dynamics and the time frame of the simulations.

The populations in most of the simulated food webs exhibited complex population dynamics. To compare the biomasses among species or among runs with and without *R*,

we used a time average of species biomass rather than one single time point. This approach is common in field experiments where species also have non-equilibrium dynamics (23, 26). It allows for a stable estimate of the mean change in  $T$  biomass due to removing  $R$ . However, this approach is not sensitive to other kinds of potentially important changes to the temporal dynamics of a species, such as changes in the temporal variance of  $T$  (24) or the amplitude of  $T$ 's population cycling (27, 28). We considered two factors in choosing the averaging window: 1) How long an average is necessary for the estimated mean biomass to not be sensitive to the window size? 2) When does the influence of initial species biomasses disappear? Together, these two factors imply a minimum length of the time series.

Despite the complexity of the population dynamics and the strong temporal fluctuations of biomass densities, preliminary analyses showed that averaging over more than 100 time steps yielded constant mean biomasses (Fig. S3). This allowed calculating interaction strengths based on the mean biomass and density over 150 time steps even though the population dynamics did not reach equilibrium. Analyses of ten food webs showed that the population densities at time step 50 were not consistently correlated with the initial random population densities ( $p$  ranged from 0.227 to 0.992). This suggests initial random biomasses did not have a significant effect on the mean biomass calculated between time step 50 and 200. Hence, we used the mean biomass and density of each species from time steps 50 and 200 to calculate the *per capita* and population interaction strengths and other predictive variables that varied dynamically.

The most appropriate total length of the time series depends on the goals of the study.

Our goal was to investigate how much the biomass of a target species ( $T$ ) changes if

another species ( $R$ ) is removed? Thus, we focused on quantifying the strength of dynamic coupling between  $R$  and  $T$  rather than the number of  $T$  species that go secondarily extinct when  $R$  is removed (29-32). This dynamic coupling includes both positive and negative interactions and is critical for understanding how a given species' abundance changes when environmental or anthropogenic stressors change other species (33). Secondary extinction is a special case of positive interactions. Generally, interaction strengths depend on  $T$  biomasses with and without  $R$  unless the time series is too short or too long. For short time series, the influence of random initial species biomasses could mask the effect of  $R$  on  $T$ . For very long time series, many  $T$  may become secondarily extinct without  $R$ . If  $T$  biomass without  $R$  is zero, the interaction strength will be a trivial function of  $T$  biomass before removing  $R$ , and it may become impossible to distinguish differences in the dynamic coupling between  $R$  and these  $T$ . For example, if two species go secondarily extinct when  $R$  is removed and have the same biomass when  $R$  is present, by definition their interaction strengths are equal even if removing  $R$  drove one to extinction much sooner than the other. If the time series short enough that neither  $T$  goes extinct, the interaction strengths are sensitive to differences in their rates of decline after removing  $R$ .

Thus, to optimize our ability to quantify variation in interaction strengths among species, we chose a time series which was long enough that the signal of starting conditions disappears, but short enough that secondary extinctions are rare. To examine this issue, we simulated separate sets of 100 webs for 50, 100, 200, 500, 1000, 2000, and 5000 time steps, and for each time series length we quantified the proportion of interactions where  $T$  went secondarily extinct after removing  $R$ . For each time series

length, we excluded the first  $\frac{1}{4}$  of time steps from the window for calculating time-averaged biomasses.

For all webs, using the approach of Dunne et al. (29), we found that 10.9 % of all the species' removals led to structural extinction of  $T$  (i.e.,  $T$  loses all of its prey). The biomass change of a species,  $i$ , without predators or prey that is declining simply due to its own metabolic losses is described by:

$$dB_i/dt = -x_i B_i$$

where  $B_i$  is the biomass of species  $i$ , and  $x_i$  is its mass-specific metabolic rate. This has the solution

$$B_i = B_i(0)e^{-x_i t}$$

where  $B_i(0)$  is the initial population density. Taking this equal to the maximum initial density of 1 and using the extinction threshold of  $10^{-30}$ , the time to extinction,  $t$ , is

$$t = 69/x_i$$

The respiration rate is

$$x_i = 0.314W^{-0.25T_i - 1}$$

where  $W$  is the body mass ratio in the food web and  $T_i$  is the trophic level of species  $i$ . Assuming a maximum trophic level of  $T = 6$  and a body-mass ratio of  $W = 10$  the time to extinction is  $\sim 4000$ . Thus, 5000 time steps are enough to capture most structural secondary extinctions caused by removing  $R$ . At 5000 time steps, additional secondary extinctions of  $T$  caused by population dynamics resulted in 20% of all interactions that resulted in secondary extinction of  $T$  at the end of the time series after the removal of  $R$  (Fig. S4). These dynamic secondary extinctions increased sharply after 200 time steps, and levelled off at approximately 2000 time steps (Fig. S4). Using time averaged

biomasses for the extinction threshold distinguishes between  $T$  that go rapidly extinct vs. those that take almost the entire time series to go extinct.

As noted above, secondary extinction of  $T$  makes that class of positive interaction strengths of  $R$  on  $T$  a trivial function of  $T$  biomass with  $R$  present. It truncates the interaction strength, which reduces our ability to detect differences in the magnitude of short term dynamic coupling of  $R$  and  $T$ . Secondary extinctions have no effect on the negative interaction strengths measured. The general patterns that we observe at 200 time steps are very similar to those at longer time steps where dynamic secondary extinctions are more common (Table S1). For each simulation time length, the same few variables are selected by CART as the best predictors of the absolute magnitude and sign of the interactions. Simple linear regressions to explain variation in the absolute magnitude of  $|I|$  give similar results at each number of time steps (Table S1). For any given number of time steps, the results are similar whether secondary extinctions are included or excluded from the analyses (Table S1). The only clear quantitative difference in results with more time steps is that  $R_{\text{biomass}}$  predicts strong interactions better (Table S1, 5000 time steps). However this result is the same whether secondary extinctions are included or excluded (Table S1). We conclude that 1) secondary extinctions do not greatly alter the results of our study, and 2) the general patterns we observe are robust to different simulated time frames.

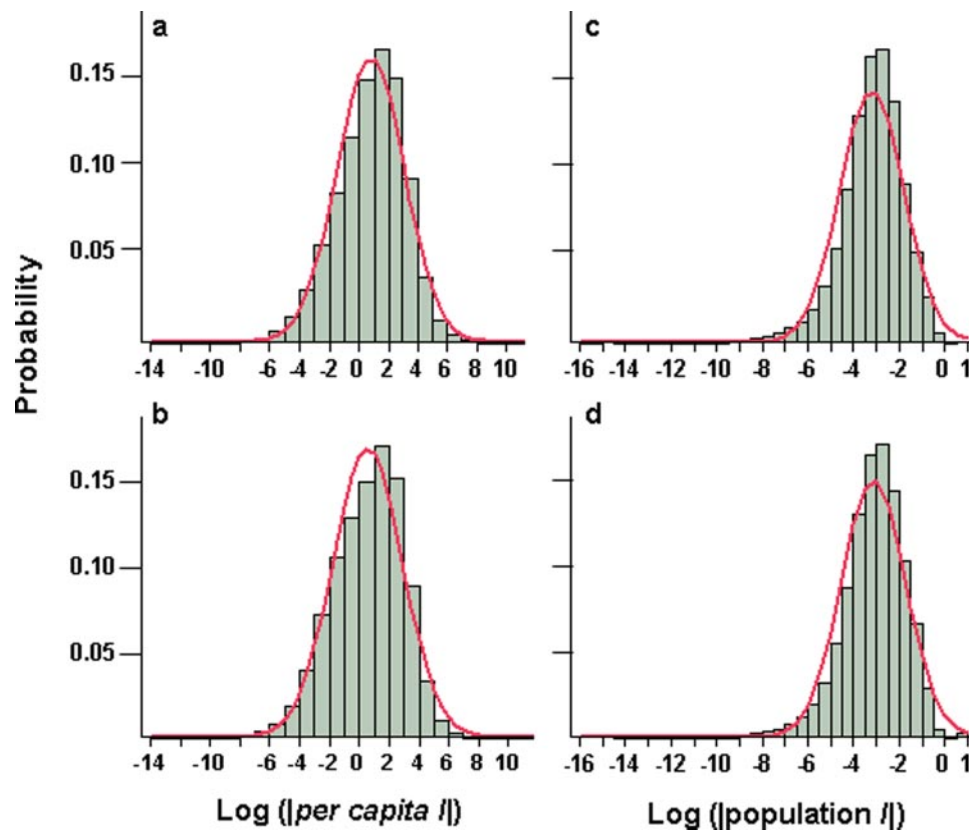
## References

1. Williams, R. J. & Martinez, N. D. (2000) Simple rules yield complex food webs *Nature* **404**, 180-183.
2. Williams, R. J. & Martinez, N. D. (2008) Success and its limits among structural models of complex food webs *Journal of Animal Ecology* **77**, 512-519.
3. Williams, R. J. & Martinez, N. D. (2004) Limits to trophic levels and omnivory in complex food webs: Theory and data *American Naturalist* **163**, E458-E468.
4. Brose, U., Williams, R. J., & Martinez, N. D. (2006) Allometric scaling enhances stability in complex food webs *Ecology Letters* **9**, 1228-1236.
5. Yodzis, P. & Innes, S. (1992) Body size and consumer-resource dynamics *American Naturalist* **139**, 1151-1175.
6. Rall, B. C., Guill, C., & Brose, U. (2008) Food-web connectance and predator interference dampen the paradox of enrichment *Oikos* **117**, 202-213.
7. Williams, R. J. & Martinez, N. D. (2004) Stabilization of chaotic and non-permanent food-web dynamics *European Physical Journal B* **38**, 297-303.
8. Beddington, J. R. (1975) Mutual Interference between Parasites or Predators and Its Effect on Searching Efficiency *Journal of Animal Ecology* **44**, 331-340.
9. DeAngelis, D. L., Goldstein, R. A., & O'Neill, R. V. (1975) A model for trophic interactions *Ecology* **56**, 881-892.
10. Brown, J. H., Gillooly, J. F., Allen, A. P., Savage, V. M., & West, G. B. (2004) Toward a metabolic theory of ecology *Ecology* **85**, 1771-1789.
11. Brose, U. (2008) Complex food webs prevent competitive exclusion among producer species *Proceedings of the Royal Society B-Biological Sciences* **275**, 2507-2514.
12. Brose, U., Berlow, E. L., & Martinez, N. D. (2005) Scaling up keystone effects from simple to complex ecological networks *Ecology Letters* **8**, 1317-1325.
13. Huisman, J. & Weissing, F. J. (1999) Biodiversity of plankton by species oscillations and chaos *Nature* **402**, 407-410.
14. Tilman, D. (1982) *Resource competition and community structure* (Princeton University Press, Princeton, New Jersey, USA).
15. Passarge, J., Hol, S., Escher, M., & Huisman, J. (2006) Competition for nutrients and light: Stable coexistence, alternative stable states, or competitive exclusion? *Ecological Monographs* **76**, 57-72.
16. Stoer, J. & Bulirsch, R. (1980) *Introduction to Numerical Analysis* (Springer-Verlag, New York).
17. Navarrete, S. A. & Menge, B. A. (1996) Keystone predation and interaction strength: interactive effects of predators on their main prey *Ecological Monographs* **66**, 409-429.
18. Berlow, E. L., Neutel, A.-M., Cohen, J. E., de Ruiter, P. C., Ebenman, B., Emmerson, M., Fox, J. W., Jansen, V. A. A., Iwan Jones, J., Kokkoris, G. D., *et al.* (2004) Interaction strengths in food webs: issues and opportunities *J Anim Ecology* **73**, 585-598.
19. Dunne, J. A., Williams, R. J., & Martinez, N. D. (2004) Network structure and robustness of marine food webs *Marine Ecology-Progress Series* **273**, 291-302.

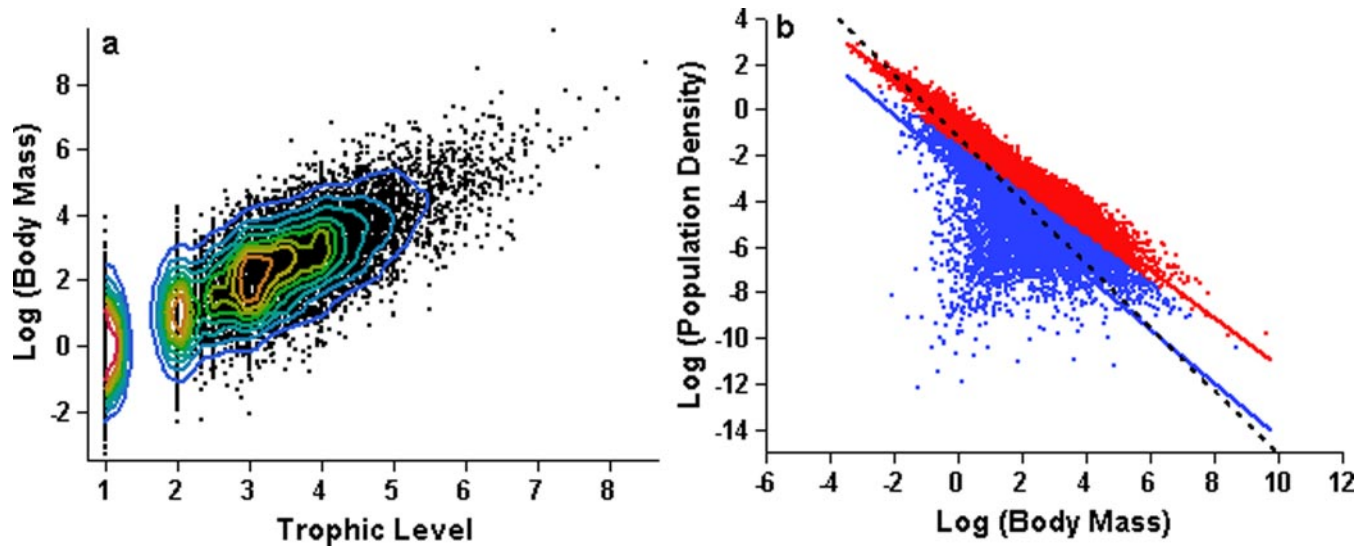
20. Williams, R. J., Berlow, E. L., Dunne, J. A., Barabási, A.-L., & Martinez, N. D. (2002) Two degrees of separation in complex food webs *Proceedings of the National Academy of Science* **99**, 12913-12916.
21. Jordan, F., Liu, W. C., & Davis, A. J. (2006) Topological keystone species: measures of positional importance in food webs *Oikos* **112**, 535-546.
22. De'ath, G. & Fabricius, K. E. (2000) Classification and regression trees: A powerful yet simple technique for ecological data analysis *Ecology* **81**, 3178-3192.
23. Berlow, E. L. (1997) From canalization to contingency: historic effects in a successional rocky intertidal community *Ecological Monographs* **67**, 435-460.
24. Berlow, E. L. (1999) Strong effects of weak interactions in ecological communities *Nature* **398**, 330-334.
25. Brose, U., Jonsson, T., Berlow, E. L., Warren, P., Banasek-Richter, C., Bersier, L. F., Blanchard, J. L., Brey, T., Carpenter, S. R., Blandenier, M. F. C., *et al.* (2006) Consumer-resource body-size relationships in natural food webs *Ecology* **87**, 2411-2417.
26. Wootton, J. T. (1993) Indirect effects and habitat use in an intertidal community: interaction chains and interaction modifications *American Naturalist* **141**, 71-89.
27. McCann, K., Hastings, A., & Huxel, G. R. (1998) Weak trophic interactions and the balance of nature *Nature* **395**, 794-798.
28. Otto, S. B., Rall, B. C., & Brose, U. (2008) Allometric degree distributions facilitate food-web stability *Nature* **450**, 1226-1229.
29. Dunne, J. A., Williams, R. J., & Martinez, N. D. (2002) Network structure and biodiversity loss in food webs: robustness increases with connectance *Ecology Letters* **5**, 558-567.
30. Ebenman, B., Law, R., & Borrvall, C. (2004) Community viability analysis: The response of ecological communities to species loss *Ecology* **85**, 2591-2600.
31. Sole, R. & Montoya, J. (2001) Complexity and fragility in ecological networks *Proceedings of the Royal Society of London Series B-Biological Sciences* **1480**, 2039-2045.
32. Srivastava, D. S. & Vellend, M. (2005) Biodiversity-ecosystem function research: Is it relevant to conservation? *Annual Review of Ecology Evolution and Systematics* **36**, 267-294.
33. Ives, A. R. & Cardinale, B. J. (2004) Food-web interactions govern the resistance of communities after non-random extinctions *Nature* **429**, 174-177.

# Supporting Information

Berlow *et al.* 10.1073/pnas.0806823106



**Fig. S1.** Emergent variation in interaction strengths. (a-b) The distribution of  $\log_{10}$  (per capita  $I$ ) for (a) positive and (b) negative interactions. (c-d) The distribution of  $\log_{10}$  (population  $I$ ) for (c) positive and (d) negative interactions.  $n = 114,114$  positive interactions for a and c,  $n = 139,918$  negative interactions for b and d. The red line is a fitted normal distribution.



**Fig. S2.** Macroecological assumptions and emergent patterns. (a)  $\text{Log}_{10}$  body mass depending on trophic level across all food webs simulated. Contour lines are 10% density quantiles from blue (low) to red (high). (b) Emergent relationship between species mean density and mean body mass for all species. Different colors indicate upper (red) and lower (blue) 50% quantiles of mean  $B_R$ . RMA Regressions of  $\text{log}_{10}$  density vs.  $\text{log}_{10}$  body mass for each biomass quantile (from low (blue line) to high (red line)): slope =  $-1.17$ ,  $-1.05$ ,  $R^2 = 0.36$ ,  $0.96$ ; error,  $n = 6,058$  species for each quantile. The overall relationship (dotted black line) has a slope of  $-1.38$  ( $R^2 = 0.59$ ).

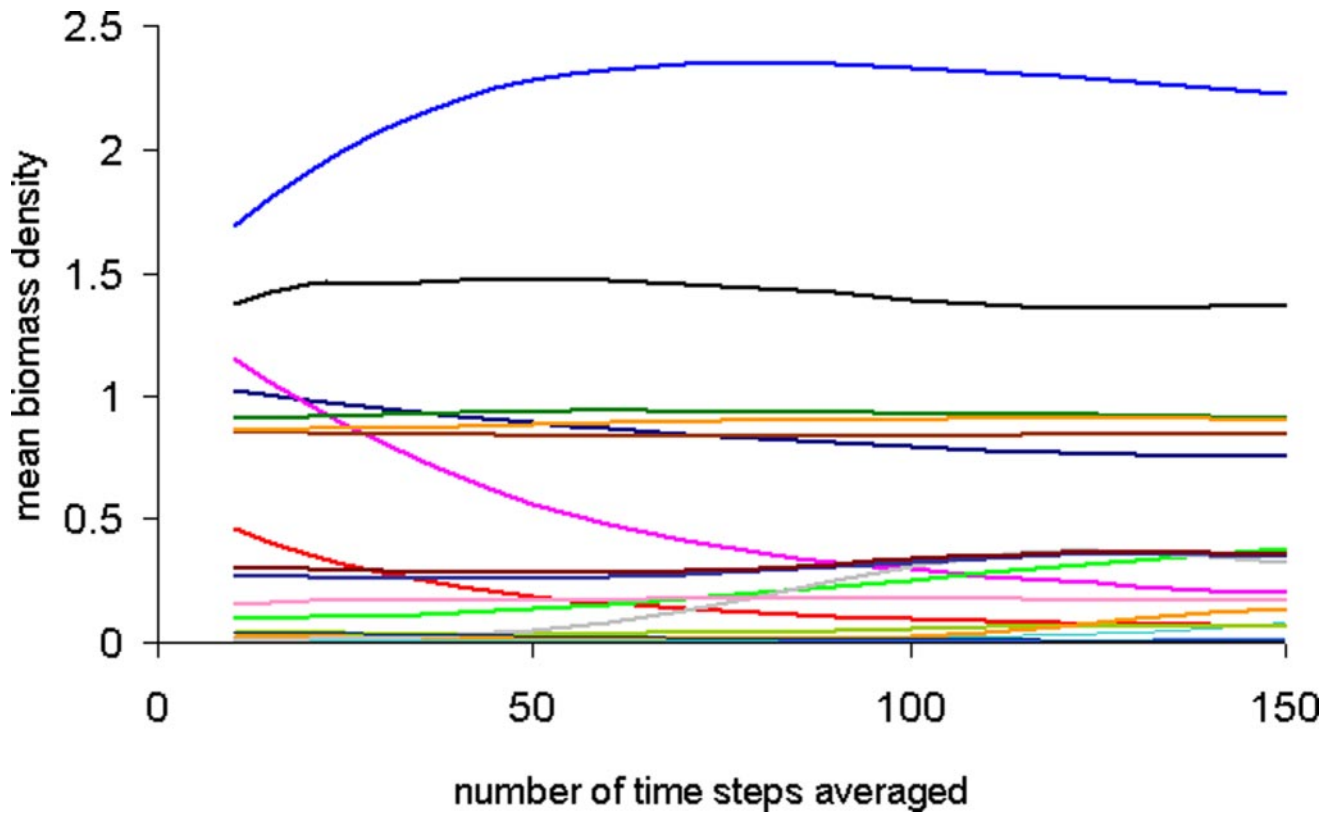
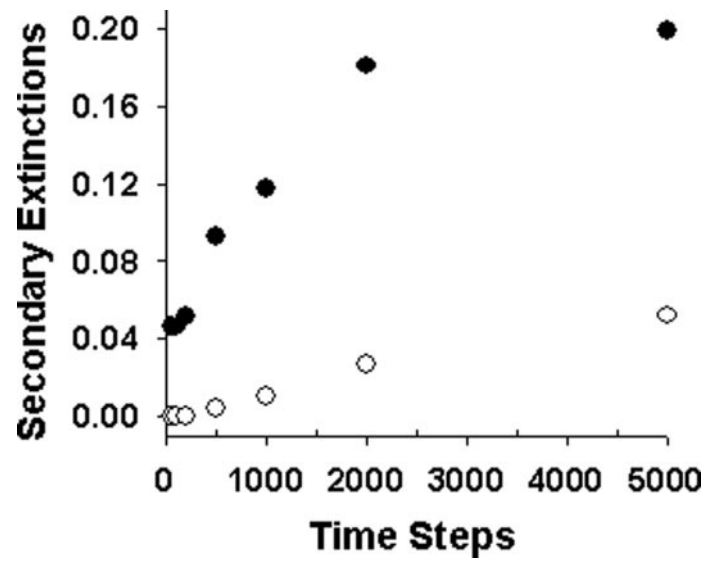


Fig. S3. Determining the time averaging window in an example 30 species niche-model food web. Mean biomass densities depending on the number of time steps averaged (starting after time step 50). Colors indicate different species



**Fig. S4.** Secondary extinctions over time. The proportion interactions where *R* removal caused secondary extinction of *T* for simulations of different length. Extinction thresholds were defined by either final *T* biomasses (solid symbols) or time-averaged *T* biomass (open symbols). One hundred different webs were used for each time series length. Cases where *R* or *T* did not persist in runs with *R* unmanipulated were excluded from the analysis.

**Table S1. Explaining variation in log ( $I/I_0$ ) for different simulation lengths**

time steps	2° extinctions	$n$	$R^2$	Intercept	$B_T^+$	$B_R$
50	included	40868	0.48	-1.20	0.68	0.20
50	excluded	38979	0.50	-1.19	0.69	0.21
100	included	42910	0.58	-1.30	0.71	0.21
100	excluded	40940	0.59	-1.27	0.72	0.21
200	included	43390	0.68	-1.33	0.73	0.20
200	excluded	41160	0.69	-1.29	0.73	0.20
500	included	37202	0.75	-1.42	0.72	0.23
500	excluded	33659	0.76	-1.38	0.73	0.22
1000	included	36462	0.74	-1.44	0.71	0.28
1000	excluded	31950	0.75	-1.42	0.71	0.27
2000	included	25181	0.75	-1.32	0.70	0.35
2000	excluded	20218	0.76	-1.33	0.70	0.34
5000	included	19242	0.68	-1.20	0.65	0.46
5000	excluded	14689	0.67	-1.22	0.63	0.46

Multiple linear regression results predicting log ( $I/I_0$ ) using log( $B_T^+$ ) and log( $B_R$ ) for different simulation lengths and with secondary extinctions included or excluded from the analysis.  $P < 0.0001$  in all cases.  $n$  = number of interactions.  $B_T$  and  $B_R$  are the biomass of  $T$  with  $R$  present and the biomass of  $R$ , respectively

## Other Supporting Information Files

[SI Appendix](#)