PROCEEDINGS OF THE ROYAL SOCIETY B | BIOLOGICAL SCIENCES

Universal power-law diet partitioning by marine fish and squid with surprising stability −diversity implications

Axel G. Rossberg, Keith D. Farnsworth, Keisuke Satoh and John K. Pinnegar

Supplementary data	"Data Supplement" http://rspb.royalsocietypublishing.org/content/suppl/2010/11/09/rspb.2010.1483.DC1.html
References	**This article cites 49 articles, 3 of which can be accessed free** http://rspb.royalsocietypublishing.org/content/early/2010/11/09/rspb.2010.1483.full.html#ref-list-1
P<P	Published online 10 November 2010 in advance of the print journal.
Subject collections	Articles on similar topics can be found in the following collections ecology (1898 articles)
Email alerting service	Receive free email alerts when new articles cite this article - sign up in the box at the top right-hand corner of the article or click **here**

Advance online articles have been peer reviewed and accepted for publication but have not yet appeared in the paper journal (edited, typeset versions may be posted when available prior to final publication). Advance online articles are citable and establish publication priority; they are indexed by PubMed from initial publication. Citations to Advance online articles must include the digital object identifier (DOIs) and date of initial publication.

To subscribe to *Proc. R. Soc. B* go to: **http://rspb.royalsocietypublishing.org/subscriptions**

Universal power-law diet partitioning by marine fish and squid with surprising stability–diversity implications

Axel G. Rossberg[1],*, **Keith D. Farnsworth**[1], **Keisuke Satoh**[2] **and John K. Pinnegar**[3]

[1]*School of Biological Sciences, Medical Biology Centre, Queen's University Belfast, 97 Lisburn Road, Belfast BT9 7BL, UK*
[2]*Tropical Tuna Resources Division, National Research Institute of Far Seas Fisheries, 5-7-1 Orido, Shimizu, Shizuoka 424-8633, Japan*
[3]*Lowestoft Laboratory, Centre for Environment, Fisheries and Aquaculture Science (CEFAS), Pakefield Road, Lowestoft, Suffolk NR33 0HT, UK*

A central question in community ecology is how the number of trophic links relates to community species richness. For simple dynamical food-web models, link density (the ratio of links to species) is bounded from above as the number of species increases; but empirical data suggest that it increases without bounds. We found a new empirical upper bound on link density in large marine communities with emphasis on fish and squid, using novel methods that avoid known sources of bias in traditional approaches. Bounds are expressed in terms of the diet-partitioning function (DPF): the average number of resources contributing more than a fraction f to a consumer's diet, as a function of f. All observed DPF follow a functional form closely related to a power law, with power-law exponents independent of species richness at the measurement accuracy. Results imply universal upper bounds on link density across the oceans. However, the inherently scale-free nature of power-law diet partitioning suggests that the DPF itself is a better defined characterization of network structure than link density.

Keywords: community structure; stability; complexity–diversity; interaction strength; species richness; food webs

1. INTRODUCTION

Relationships between biodiversity and the stability and complexity of ecological communities are central to understanding their assembly, structure, function and persistence, and hence important for conservation. Odum [1], MacArthur [2] and others initiated the so-called complexity–diversity–stability debate nearly 60 years ago. By now, we learned that the answers to the questions asked then critically depend on their precise formulation [3]. Here we address one specific set of questions—originally posed by MacArthur [2,4] and May [5] and still being discussed today [6–8]—concerning the relation between the number of species S and the number of trophic links L (i.e. consumer–resource pairs) in large natural food webs. Because link counts L inevitably increase with community size, they are 'normalized' (descaled) by dividing by the potential number of links, giving the connectance or 'complexity' $C = L/S^2$, or by species richness, giving the *link density* $Z = L/S$ of the community (see table 1 for a list of symbols).

Mathematical arguments and simulations by Gardner & Ashby [4], May [5] and others (using simple models of large random communities) showed that there are upper limits to link density, $Z = CS$, beyond which community steady states become unstable, ultimately leading to species extinctions [9]. Communities with numerically large Z should therefore not be observable in the field. Early empirical studies appeared to confirm this prediction (e.g. [10,11]), but later work using larger food-web datasets contradicted the theoretical expectations (e.g. [8,12–15]). In these studies, scaling laws such as $Z \propto S^\alpha$ were found, with values of α ranging from 0.3 [15] to 1 [14], conforming with the long-standing empirical intuition that 'complexity begets stability'. These observations motivated an intense search for mechanisms that could stabilize communities with large S and Z (e.g. [16–25]).

However, there are a number of issues that, although known and acknowledged in principle [7,26–31], have not been fully and jointly accounted for in the analyses of food-web data.

Three facts in particular are important: most trophic links are weak [32,33] empirical food-web datasets vary in their resolution of taxonomic or functional groups [27,28], and criteria for recording or excluding particular trophic links vary between datasets [30]. Without giving these three issues sufficient consideration, conclusions drawn from analyses of food-web topologies regarding complexity–diversity and complexity–stability relations might be premature.

To overcome these problems, we use here the *diet-partitioning function* (DPF), proposed by Rossberg *et al.* together with a protocol for calculating it from incompletely resolved diet data [34]. We express link strengths in

* Author for correspondence (axel@rossberg.net).

Table 1. List of important symbols.

symbol	meaning
α, γ	exponents for scaling with S
ν	exponent for scaling of Z_c with r
ν_i	ν measured in system i
$\bar{\nu}$	sample mean of ν
$\Delta\nu_i$	residuals of regression of ν against log-effort
C	food-web connectance
E	total number of non-empty stomachs sampled
e_i	standard error of ν_i
f_{ij}	diet fraction: relative contribution of i to j's diet
f	diet-fraction threshold
L	number of trophic links in a food web
r	diet-ratio threshold $= f/(1-f)$
S	species richness
S_c	consumer species richness
S_p	producer species richness
S_{fish}	fish species richness
s_ν	sample standard deviation of ν
Z	links per species
Z_c	links per consumer
$Z_c(f)$	diet-partitioning function (DPF)

Figure 1. Approximate locations of sampling sites.

terms of gravimetric or volumetric *diet fractions*, i.e. proportional contributions to a consumer's diet—a common theoretical [35] and empirical [30,31] convention—such that f_{ij} is the fraction of the total diet of species j that is made up of species i. The DPF $Z_c(f)$ is defined as the community average of the number of prey species that contribute more than a fraction f to the diet of a consumer species, where f takes values between 0 and 1. The value of $Z_c(f)$ can be interpreted as the mean number of trophic links per consumer, where only those links that are stronger than f are counted.

If S_c is the number of consumer species and S_p is the number of producer (autotrophic) species in a community (so that $S_c + S_p = S$) then, irrespective of the criterion for counting links, the *consumer link density* $Z_c = L/S_c$ must always be larger than the link density $Z = L/(S_c + S_p)$, because $S_p > 0$. Since Z_c equals the mean number of prey per consumer, it can be estimated from a sub-sample of consumer species. It is therefore measured more easily than Z for large communities.

In this paper, we use Z_c to test whether there is an upper bound to Z; specifically, we test if Z_c increases with species richness S because, if it does not, then $Z_c > Z$ implies that Z must be bounded from above for increasing S. Conversely, an unbounded increase of Z with S implies an unbounded increase of Z_c with S. Counting only links stronger than f, but keeping f variable, we report estimates of $Z_c = Z_c(f)$ in relation to S, taken from seven different marine stomach-contents datasets. Based on our analysis, we will argue that the DPF $Z_c(f)$ itself might be a more suitable characterization of communities than any particular value Z_c extracted from it.

2. MATERIAL AND METHODS
(a) *Calculation of DPF*
We use the *diet ratio* $r_{ij} = f_{ij}/(1 - f_{ij})$, comparing the contribution of the ith diet item to the sum of all other
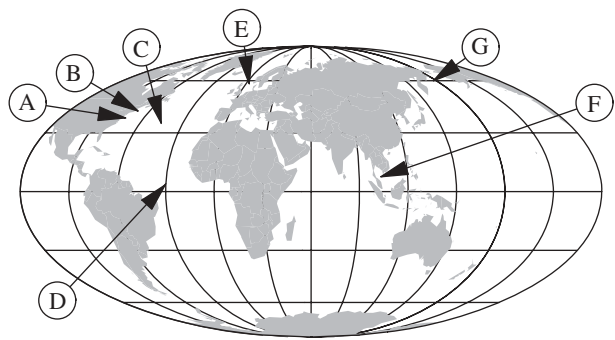
contributions to the diet of j, as an alternative measure to the diet fraction f_{ij}. Diet ratios can attain values in the range $[0;\infty]$ and, like 'odds' in statistics, are the natural choice on logarithmic scales.

The basic idea underlying the correction of the DPF for unresolved diet items is to first count only those diet items that are resolved to species level, and then to compensate for the proportion of diets left out. A detailed illustration of the method and hands-on instructions for its implementation are given in electronic supplementary material, appendix S1. For a systematic derivation, see Rossberg *et al.* [34].

(b) *Diet data*
The DPF $Z_c(f)$ was computed for seven stomach-contents datasets (labelled A–G) of marine fish and squid [36–40]. As illustrated in figure 1, sample sites span a broad latitudinal range. Key properties of the datasets are listed in table 2. Data sources differ considerably by sampling methods, sampling efforts and the kinds of species included. Details are discussed in electronic supplementary material, appendix S2.

(c) *Fitting curves to observed DPF*
Characterizations of the DPF in terms of theoretical or heuristic models should be based only on a range of diet fractions f for which the data are reliable. For obvious practical reasons, actual or conceivable diet items making very small (rare) contributions to a consumer's diet are unlikely to be observed and recorded, leading to an underestimation of the DPF for low threshold values. As f increases, the probability of observation gradually increases, up to a point above which sampling is essentially complete and empirical DPF, while still exhibiting measurement errors, are unbiased. Based on the number of non-empty stomachs sampled per consumer, we assume this point to be somewhere below $f = 0.02$, i.e. that sampling of contributions larger than $f = 0.02$ is essentially complete (see also electronic supplementary material, supplementary discussion S6). In terms of diet ratios, this corresponds to the range $0.02 \lesssim r \lesssim 50$, which we shall call the *reliable range*.

Models for the DPF, such as equation (3.2) below, were fitted to the data over this reliable range by a likelihood maximization technique. Details of the procedure are described in electronic supplementary material, appendix S3.

(d) *Species richness*
Traditionally, diversity–complexity studies measure diversity by the numbers of nodes distinguished in empirical food-web datasets. Since here only sub-samples of food webs are used, information on species richness has to be obtained independently. Fortunately, only relative richness estimates

Table 2. Datasets included in the analysis.

dataset	S_{fish} AquaMaps	S_{fish} FishBase + OBIS	consumer species included	resource species resolved	non-empty stomachs
(A) N.-W. Atlantic Shelf I	247	645	146	767	29 032
(B) N.-W. Atlantic Shelf II	247	645	117	216	50 027
(C) Open North Atlantic	335	679	17	19	188
(D) Open Tropical Atlantic	340	179	18	24	357
(E) North Sea	103	192	15	91	5599
(F) South China Sea	512	3827	18	14	~1007
(G) Eastern Bering Sea	92	250	25	137	17 688

are required for our argument. Further, as shown in a detailed, quantitative discussion of these questions in electronic supplementary material, supplementary discussion S6.2, rather coarse richness estimates are sufficient to support our main conclusion.

We use estimates of species richness based on the FishBase [41], OBIS [42] and AquaMaps [43]. These databases have global coverage, and therefore allow us to obtain, with a few exceptions, richness estimates specific to the study sites considered here. All three databases are substantially more detailed for fish than for most other taxa. To reduce biases owing to data gaps, relative richness is therefore measured in terms of richness of fish, S_{fish}, including Actinopterygii, Chondrichthyes and Agnatha for FishBase and OBIS, and only Actinopterygii (the dominating taxon) for AquaMaps. Overall, marine species richness is known to follow the same global trends as that of fish, with some variation in details [44]. This fact and the low demands on accuracy mentioned above justify our choice of S_{fish} as an estimate of relative richness. This measure has the additional advantage that its empirical uncertainty has been quantified [45], allowing a quantification of the uncertainty this implies for our main results. As shown in electronic supplementary material, supplementary discussion S6.2, this uncertainty is small when compared with that stemming from the dietary analysis.

Despite this robustness to uncertainties in richness estimates, we performed our analysis using two sets of values for S_{fish}, one derived from AquaMaps, the other from FishBase and OBIS. The reason is that both approaches have specific strengths: FishBase + OBIS data are based on direct observations and therefore methodologically more transparent; AquaMaps is more robust to data gaps, and richness estimates specific to the delineations of the study areas A–G can be obtained (see electronic supplementary material, appendix S4 for details). Estimated species richnesses S_{fish} are listed in table 2. The much lower FishBase + OBIS richness in the Tropical Atlantic (set D) when compared with the North Atlantic is inconsistent with other estimates [44] and most likely attributable to data gaps [42].

Because the FishBase richness estimate for fish in the 'South China Sea' (3827) is considerably larger than our estimate for study area F using AquaMaps (512, the most species-rich study area), the nominal accuracy of the estimated slope of link density versus diversity turns out to be substantially higher when using the FishBase + OBIS data. To caution against inflated accuracy, we therefore conservatively report detailed results using the AquaMaps richness measures below, followed by short summaries of the corresponding results using FishBase + OBIS.
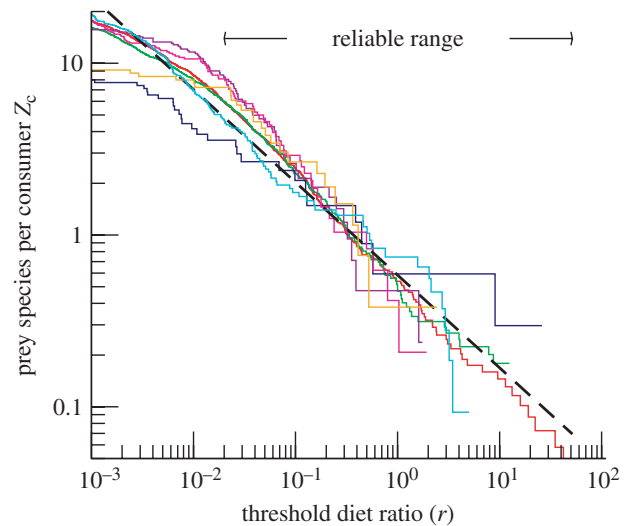
Figure 2. Comparison of seven empirical DPF from six marine areas. The diet ratio r is the contribution of one prey species to a consumer's diet relative to all other contributions. Solid lines: datasets A–G. Dashed line: power-law equation (3.2) with $\nu = 0.54$. Red, A; green, B; blue, C; violet, D; pink, E; orange, F; light blue, G.

3. RESULTS

(a) *Computation and characterization of empirical DPF*

The DPF estimated from sets A–G are shown in figure 2, with the threshold f expressed as a diet ratio, $r = f/(1 - f)$. Surprisingly, all DPF except for set C (blue) match a single curve, suggesting the possibility of a universal law for diet partitioning within these communities. Over the reliable range in r (see §2c), $0.02 \lesssim r \lesssim 50$, the curves appear to follow power laws (straight lines in log–log graphs), mirroring a recent similar observation of power law [46] rather than exponentially distributed [32,33] trophic fluxes.

This allows us to describe the DPF in the power-law form: $Z_c = K r^{-\nu}$. The constant of proportionality K can now be computed from the condition

$$\int_0^1 Z_c(f)\mathrm{d}f = 1, \tag{3.1}$$

which follows from the fact that all diet fractions of a consumer add up to one [34]. One obtains $K = \pi^{-1}\nu^{-1} \sin(\pi\nu)$ for $0 < \nu < 1$, giving

$$Z_c = \frac{\sin(\pi\nu)}{\pi\nu} r^{-\nu} = \frac{\sin(\pi\nu)}{\pi\nu}\left(\frac{f}{1-f}\right)^{-\nu}. \tag{3.2}$$

Table 3. Results of maximum-likehood fits of three functional forms to the observed DPF.

assumed functional form[a]:	$Z_c \propto r^{-\nu}$				$Z_c \propto \max[\ln(c_2/r),0]$			$Z_c \propto \exp(-c_1 r)$		
dataset	$\nu \pm$ s.e.m.	d.o.f.	χ^2	p	χ^2	p	ΔAIC[b]	χ^2	p	ΔAIC[b]
(A) N.-W. Atlantic Shelf I	0.53 ± 0.03	56	35.3	0.99	506.9	$<10^{-9}$	998.2	415.3	$<10^{-9}$	338.1
(B) N.-W. Atlantic Shelf II	0.56 ± 0.04	44	33.2	0.88	224.5	$<10^{-9}$	393.4	223.0	$<10^{-9}$	165.4
(C) Open North Atlantic	0.30 ± 0.09	10	13.5	0.20	79.3	$<10^{-9}$	98.5	26.3	0.003	9.4
(D) Open Tropical Atlantic	0.55 ± 0.11	19	10.7	0.93	38.3	0.005	34.5	30.8	0.04	8.1
(E) North Sea	0.50 ± 0.10	20	10.2	0.96	47.2	0.001	47.7	26.2	0.16	6.3
(F) South Chinese Sea	0.43 ± 0.12	13	11.3	0.58	52.6	10^{-6}	51.4	25.2	0.02	7.1
(G) Eastern Bering Sea	0.58 ± 0.07	22	25.3	0.28	111.2	$<10^{-9}$	145.3	63.9	10^{-5}	25.5

[a] ν, c_1, c_2 are fitting parameters; proportionality constants are given by equation (3.1).
[b] Excess AIC relative to power-law fits; positive values indicate that power law is the preferred model.

Beyond the reliable range, when $r \lesssim 0.02$, the empirical DPF in figure 2 decline in slope to near zero, clearly deviating from power laws. This may result from incomplete sampling of small diet fractions, but may conceivably show a true underlying deviation from power laws at very low diet-ratio thresholds. We consider resolving these possibilities in electronic supplementary material, supplementary discussion S6, and here restrict our analysis to the reliable range ($0.02 \lesssim r \lesssim 50$), addressing the following three questions: (i) are the empirical DPF consistent with power laws over the reliable range in r? (ii) What values of the exponent ν best fit the empirical DPF over the reliable range? (iii) How do the fitted values vary among study sites?

Estimates of ν from maximum-likelihood fits of the power law (3.2) to the empirical DPF over the reliable range and related statistics are listed in table 3 (details provided in electronic supplementary material, appendix S3). Goodness-of-fit by χ^2-statistics shows that all empirical DPF are compatible with power laws (table 3). For comparison, models with exponential and logarithmic dependencies of Z_c on r score considerably worse by χ^2-statistic and by the Akaike Information Criterion (AIC, see table 3 and also electronic supplementary material, appendix S3).

Thus, over the reliable range, all DPF $Z_c(f)$ statistically match power laws that are wholly characterized by the exponent ν. For power-law DPF, constancy of ν over a range of species richness S_{fish} implies constancy of consumer link density $Z_c(f)$ at varying species richness for any value of the diet-fraction threshold f. Comparisons of link densities across study areas are therefore most efficiently carried out as comparisons of the fitted exponents ν across study areas.

(b) Comparison of the power-law exponent ν across study sites

The maximum-likelihood fitting procedure for the power-law equation (3.2) yields Cramér-Rao lower bounds as estimates for the standard errors e_i of the fitted parameter ν (table 3) for each dataset i. This information (not entering conventional regression analysis) was used to estimate the accuracy of fitted regression models. Specifically, 10^6 Monte Carlo simulations of the null model

$$\nu_i = \nu^* + e_i \xi_i \quad (i = 1, \ldots, 7) \qquad (3.3)$$

with standard-normal ξ_i were generated. Species richnesses and sampling efforts were fixed as in table 2,

and, without loss of generality, we set $\nu^* = 0$. By evaluating the simulated data just as the measured data, standard errors of regression coefficients and p-values for null hypotheses were obtained. Confidence intervals were computed by offsetting the corresponding distributions of the simulated data by the estimated values obtained from the measured data (which is admitted by the linearity of the regression models and the symmetry of standard normal distributions).

(i) Cross-site comparison of exponents without correction for effort

Assuming that the power-law exponents for all study systems share a common value, the maximum-likelihood estimate $\bar{\nu}$ for this value can be computed as the weighted average of the values listed in table 2, with the weights chosen as the inverse variances of the estimation errors $1/e_i^2$. In what follows, all means and regressions are computed with this weighting of datasets. We obtain $\bar{\nu} = 0.54(2)$ (digits in parenthesis represent standard errors estimated using model (3.3)).

To test the null hypothesis that all exponents are equal, we computed the statistic $SS_0 = \sum_i (\nu_i - \bar{\nu})^2 e_i^{-2} = 8.4$. Comparison with model (3.3) shows that data are consistent with this hypothesis at $p = 0.21$, that is, SS_0 computed using simulated data from model (3.3) is larger than the empirical value in 21 per cent of all cases.

The weighted sample standard deviation of ν is $s_\nu = 0.032$, giving a coefficient of variation $CV = s_\nu/\bar{\nu} = 0.06$. By an argument detailed in electronic supplementary material, supplementary discussion S6.2, this small CV itself strongly constrains the scope for any systematic variation of ν with species richness, independent of the particular measure of species richness used.

However, specific tests for the dependence of ν on species richness have stronger statistical power than, e.g. the simple test for equal means above. For example, a regression of ν against AquaMaps estimates of species richness S_{fish} yields $\nu = 0.62(7) - 0.00040(27) \, S_{\text{fish}}$. The 95 per cent confidence region for this regression, together with measured ν values, is shown in figure 3a. Despite the weak negative trend, the regression coefficient is consistent with the null hypothesis that ν is independent of S_{fish} at $p = 0.14$.

Using the FishBase + OBIS richness estimates, the slope of the regression of ν on S_{fish} becomes $-0.000034(36)$, consistent with zero at $p = 0.35$. Inclusion or exclusion of dataset D in this analysis (for
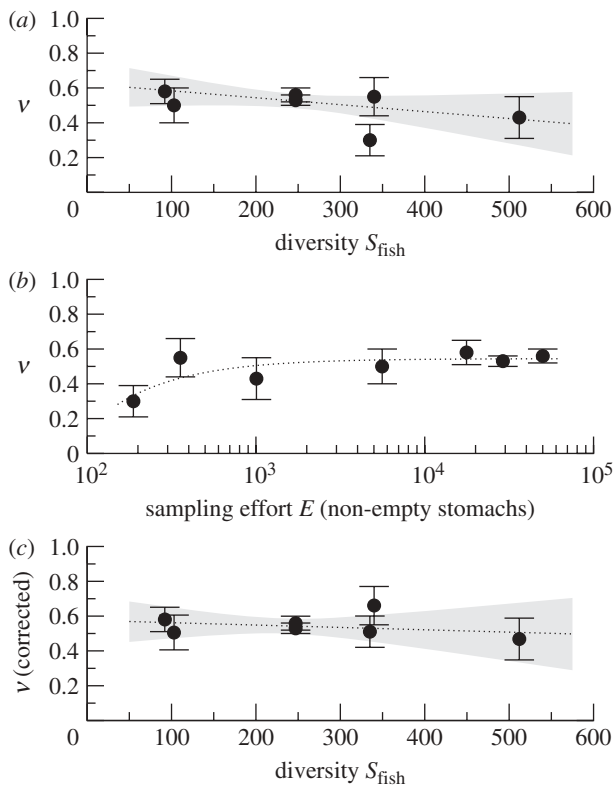
Figure 3. Dependencies of diet-partitioning exponent ν on species richness and sampling effort. Each point represents one dataset. In (c) ν is corrected for effort. Error bars indicate s.e. Dotted lines are regressions of ν against (a) species richness, (b) inverse effort and (c) both. Shaded areas in panels (a) and (c) indicate 95% confidence regions of regression lines.

its questionable richness estimate) gives numerically identical results.

(ii) *Cross-site comparison of exponents with correction for effort*
Figure 3b shows empirical values of ν drawn against sampling effort E, measured in terms of the number of non-empty stomachs analysed. The graph indicates a weak dependence of ν on effort. To check this hypothesis we compared the AIC for model fits of the forms $\nu = a$ (ΔAIC $= 0$), $\nu = a + bE$ (ΔAIC $= -0.98$), $\nu = a + bE^{-1}$ (ΔAIC $= -3.95$) and $\nu = a + b \log E$ (ΔAIC $= -3.01$). While not unequivocal, the negative ΔAIC for the last three models suggest that values of ν are somewhat biased by sampling effort. A bias proportional to E^{-1} (third model) is favoured by the AIC and appears plausible because asymptotic low-sample-size biases of this form are often encountered in statistics. The dotted line in figure 3b is the corresponding fitted curve. It is given by $\nu = 0.54(2) - 40(16)E^{-1}$, with the second term describing a bias at low sampling efforts. Exponents ν corrected for this bias are shown in figure 3c.

Indeed, while numerically leaving the overall estimate $\bar{\nu} = 0.54(2)$ unchanged, the correction for effort improves the consistency of measured diet-partitioning exponents ν across study sites. Based on the statistic $SS = \sum_i (\Delta \nu_i)^2 \, e_i^{-2} = 2.4$, where $\Delta \nu_i$ are the residuals of the regression against E^{-1}, the null hypothesis of equal ν across study sites after correction is consistent with the data with $p = 0.78$ (by comparison of SS with simulations of model (3.3), including simulated corrections for effort).

Correction for effort also removes the weak negative dependence of diet breadth on species richness. A combined regression of ν against sampling effort and species richness (AquaMaps) yields $\nu = 0.57(7) - 36(18)E^{-1} - 0.00014(30)S_{\mathrm{fish}}$. The 95 per cent confidence region of this regression for infinite effort is indicated in figure 3c. Since we simultaneously regressed against effort and species richness, this confidence region accounts for uncertainties by both the correction for effort and the regression on S_{fish}. (Strong correlations between E^{-1} and S_{fish} would broaden this region rather than narrowing it.) Comparing the regression slope with simulations of model (3.3), the null hypothesis that corrected ν do not depend on S_{fish} is fully supported ($p = 0.65$). A statistic characterizing the strength of the dependence of one empirical quantity on another (rather than the strength of evidence for a dependence) is the elasticity at the sample means ([47], see also electronic supplementary material, appendix S6.2), obtained by normalizing the regression slope to the sample means of the two variables. With weighted mean species richness $\bar{S}_{\mathrm{fish}} = 243.1$, the elasticity of the dependence of ν on S_{fish} evaluates to $-0.06(14)$. When assuming a power-law relation between ν and S_{fish}, we obtain $\nu = 0.67(S_{\mathrm{fish}})^{-0.04}$ with a 95% confidence interval $(-0.12, 0.25)$ for the exponent, again after correcting for insufficient sampling (and using an adaptation of model (3.3) to handle this case).

When using FishBase + OBIS richness estimates, the p-value for independence of ν from S_{fish} becomes 0.42 when including set D and 0.55 when excluding it. The regression slope for the dependence of corrected ν on S_{fish} becomes $-0.000029(36)$ (elasticity $-0.036(45)$) when including and $-0.000022(37)$ (elasticity $-0.028(47)$) when excluding set D.

All results reported above are consistent with constant ν, and therefore constant consumer link density across study systems, and the data impose tight bounds on any conceivable dependence of link-density on species richness.

4. DISCUSSION
(a) *Diversity–complexity implications*
We found that, to the accuracy of our analysis, all DPF followed a power law over the reliable range, which covers more than three orders of magnitude in diet ratios, and that the power-law exponent ν is independent of species richness S_{fish}. We cannot exclude a weak positive or negative dependence of ν on species richness, but a steady decline of ν to zero as S_{fish} goes to zero seems unlikely in view of the low upper bound for the exponent of a conceived power-law relation $\nu \propto S_{\mathrm{fish}}^{\gamma}$ ($-0.25 \leq \gamma \leq 0.12$).

If the DPF exponent ν is independent of local species richness, then so is consumer link density $Z_c(f)$ at any fixed threshold f within the reliable range. This, in turn, implies an upper bound on the conventional link density Z (with links thresholded at f) that is independent of community size. This means that Z could, for example, approach some constant, or may steadily decrease as species richness increases. Even if the DPF slightly deviated from a perfect power law, this would lead only to small additional uncertainties and not affect the role of $Z_c(f)$ as an upper bound on Z. Observations and

theory suggest that consumers have broader diets at higher trophic levels [23,48], so the inevitable over-sampling of consumers from higher trophic levels will rather over- than underestimate Z_c, reinforcing the interpretation of Z_c as an upper bound for Z.

Despite these clear findings, it is not a simple matter to draw definite conclusions in terms of link density Z or consumer link density Z_c, as that would raise questions such as how link strength is best measured, how the threshold f is best chosen, if it might need to be adjusted with system size, or if simply all links have to be counted ($f = 0$). Such questions (addressed in depth in electronic supplementary material, supplementary discussion S6.3) relate much more to what, by definition, should count as a 'trophic link' than to actual ecological phenomena. The power-law nature of the DPF reinforces these questions, because power laws are, in contrast to other kinds of functional relationships, 'scale free' [49]. This means that a characteristic scale of the independent variable (here r) is not implied in the functional relationship itself.

To sidestep the question how trophic links are to be counted (or weighted, see [8]), we wish to suggest reformulating the basic problem addressed here as a problem concerning relations between the distribution of trophic link strengths and local species richness (both, by specific measures). Specifically, we asked here if the DPF changes as species richness changes across the oceans. The answer we found was that, for fish and squid, stomach-contents data showed no significant changes, so implying a universal form: that of the power law, equation (3.2) (at least in the range of diet ratios $0.02 \lesssim r \lesssim 50$). These findings were anticipated by figure 2, and the subsequent statistical analyses fully confirmed them.

In electronic supplementary material, supplementary discussion S6.1, we detail a possible explanation for the universal power law; essentially that it arises from zooming into the upper tail of a diet-ratio distribution that is broad on a logarithmic scale (see [49] for a general discussion of power laws and mechanisms). But this explanation does not clarify why the exponent ν attains a universal value. Further study will be needed to understand the underlying mechanisms. This may be achieved by adapting theories that invoke the presence or absence of trophic-link to situations with broad distributions of link strengths on logarithmic scales. Good starting points might be theories examining limits to link density through limits to the stability [5], invadibility [50] or feasibility [51] of communities, and limits on the occurrence of—potentially destabilizing [52]—loops in interaction networks [53]. Generalizations of such theories will express the respective limits in terms of the DPF or similar characterizations of link-strength distributions.

Our characterization of the DPF is currently valid down to diet fractions of 0.02 and it is not yet clear whether this resolution is sufficient to distinguish between different theories for mechanisms controlling link density, or to judge what the implications of the observed DPF are for ecosystem properties such as stability, invadibility or feasibility. Only with the more general theories at hand will we be able to answer these questions. As an illustration, let us speculate that the diet fractions f_{ij} of a consumer j enter a relevant theoretical formula, e.g.

through

$$c_j^* = \frac{\sum_i f_{ij}^2}{\sum_i f_{ij}} = \sum_i f_{ij}^2, \tag{4.1}$$

a quantity closely related to Lloyd's [54] 'mean crowding' [55]. 'Niche breadth' *sensu* Levins [56] is simply the inverse $(c_j^*)^{-1}$ [55]. Invoking the probabilistic interpretation of the DPF [34], the expectation value of c_j^* is computed in electronic supplementary material, appendix S5 as

$$E c_j^* = - \int_0^1 f^2 Z_c'(f) \mathrm{d}f \tag{4.2}$$

(where $Z_c'(f) = \mathrm{d}Z_c(f)/\mathrm{d}f$). For power-law DPF following equation (3.2) with $0 < \nu < 1$, this integral evaluates to $E c_j^* = (1 - \nu)$. With the exponent $\nu = 0.54$ previously found, $E c_j^* = 0.46$. How much uncertainty does our inability to resolve the DPF for $f < 0.02$ add to this result? Consider the extreme and obviously counterfactual (figure 2) case that no diet fractions < 0.02 exist at all. The corresponding DPF is constant for $f \leq 0.02$, and requires a correction of normalization according to equation (3.1). With this DPF, formula (4.2) evaluates to $E c_j^* \approx 0.52$. The resulting uncertainty in $E c_j^*$ is 0.06 (plus the measurement uncertainty in ν of similar magnitude). This might be sufficiently small to distinguish between different theoretical models and predictions for dietary diversity in the oceans. We caution, however, that the DPF could enter theory through expressions very different from equation (4.2), perhaps giving more weight to small diet fractions.

(b) Separation of contributions to empirical DPF

The DPF and the diet-partitioning exponent ν are highly integrated summary statistics. To inform and constrain theories aimed at explaining their observed constancy, a separation of distinct contributions to these statistics can be helpful. This shall here be illustrated by two contrasting examples.

Define the consumer-specific DPF for a given consumer species as the number of this consumer's diet fractions larger than a threshold value f. The (proper) DPF $Z_c(f)$ is the arithmetic mean of all consumer-specific DPF over a community. Figure 4 displays approximations of the consumer-specific DPF of all consumers who contribute to dataset A. Immediately apparent is the large variability among consumers. To characterize this variability, we fixed the threshold diet ratio at $r = f/(1 - f) = 0.02$ (dashed line), and computed the cumulative distribution of the number of diet fractions greater than f, that is, of thresholded consumer *generality* [12]. The distribution is shown in the inset of figure 4. Following Camacho *et al.* [57], we normalized generality to its mean ($= Z_c(f)$). This allows a direct comparison with a heuristic generality distribution function that was shown by Camacho *et al.* [57] to describe binary food-web data well (dash-dotted line). Both distributions are characterized by a wide spread and a strong skew towards small generality. Apparently, the large variability of consumer-specific DPF reflects a corresponding phenomenon for binary food webs, presumably resulting from phylogenetic clustering [58–60]. The large variability among
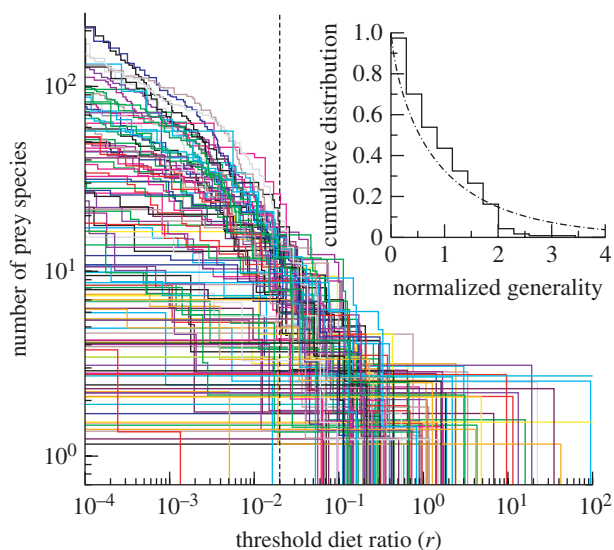
Figure 4. Consumer-specific DPF for dataset A (main panel) and distribution of generality (inset) for diet contributions larger than $r = 0.02$ (dashed vertical line). The dash-dotted line corresponds to the generality distribution proposed by Camacho *et al.* [57,62].

consumer-specific DPF, not found among community DPF, suggests that the mechanism regulating the diet-partitioning exponent ν operates rather at the community level than at species or individual level.

As a second example, figure 5 displays a time series of yearly diet-partitioning exponents ($\pm$s.e.) computed from datasets A and B. Each year over 1500 non-empty stomachs were sampled, so that, by figure 3b, biases owing to insufficient sampling are not expected. Even though overall deviations of the yearly slopes from the sample mean are larger than expected from the computed standard errors (χ^2-test, $p = 0.01$), these deviations are too small to be associated with any particular structure in the data. At the 95 per cent confidence level, the time series shows, despite the visual impression, no significant linear or quadratic trend, there is no significant year-to-year correlation and no individual data point deviates significantly from $\nu = 0.54$ or the sample mean after Bonferroni correction. We conclude that the current empirical accuracy is insufficient for identifying temporal variability in the diet-partitioning exponent ν.

(c) *Verification*

The regularities in DPF that we report should be scrutinized and subjected to empirical testing as widely as possible, especially since generalizations in comparative food-web studies have often later been shaken by comprehensive testing [7]. For marine communities, diet data have been collected at many sites throughout the world and over a considerable period from the beginning of the twentieth century onwards [38]. We would welcome more high-quality data of this kind becoming publicly accessible. Here we call for measurements of the DPF at marine biodiversity hotspots, where species richness in terms of S_{fish} can become many times larger than the largest values covered in this study [43]. We also need to better understand if, and how, the DPF changes when including consumers at lower trophic levels, where the diversity of both consumers and potential resources increases considerably. Preliminary observations indicate
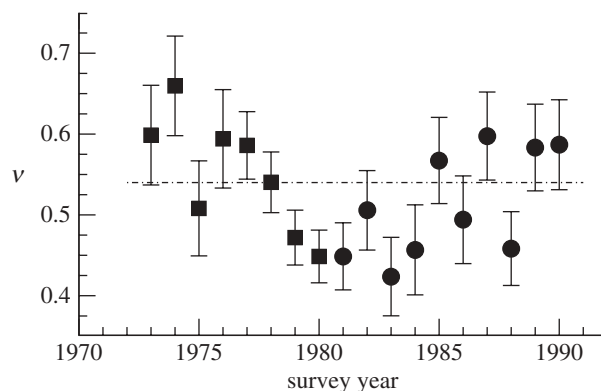
Figure 5. Diet-partitioning exponents ν ($\pm$s.e., see electronic supplementary material, appendix S3) for each survey year covered by datasets A and B. Filled squares, set A; filled circles, set B; dash-dotted line, $\nu = 0.54$.

that even among planktonic consumers, dietary specificity can be substantial [61].

5. CONCLUSIONS

We propose the DPF as a valuable new tool in community ecology. Since the DPF is defined as a community-level average, it can be estimated by averaging over subsamples. This allows investigating large communities without trade-offs in taxonomic resolution or concern about the empirical basis of link-assignment.

With the data available to us, we were unable to detect any significant deviations from a universal power law with exponent $\nu = 0.54(2)$ in the DPF of communities of fish and squid across the oceans, using five related but different statistical tests. Even though biodiversity varied fivefold over the sites considered, dietary diversity did not change noticeably. These findings are made even more remarkable by the fact that the datasets we used differ considerably in sampling methods and organisms.

Universal diet partitioning seems to reflect the working of a yet unidentified ecological mechanism structuring marine communities, which may or may not be related to community stability. Only future observations of deviations from DPF universality will tell how powerful this mechanism ultimately is. Findings by Townsend *et al.* [29] that link density decreases under ecosystem disturbance are noteworthy in this context.

REFERENCES

1 Odum, E. P. 1953 *Fundamentals of ecology*. Philadelphia, PA: Saunders.

2 MacArthur, R. H. 1955 Fluctuations of animal populations, and a measure of community stability. *Ecology* **36**, 533–536. (doi:10.2307/1929601)

3 Justus, J. 2008 Complexity, diversity, and stability. In *A companion to the philosophy of biology* (eds S. Sarkar & A. Plutynski), ch. 18, pp. 321–350. Malden, MA: Blackwell.

4 Gardner, M. R. & Ashby, W. R. 1970 Connectance of large (cybernetic) systems: critical values for stability. *Nature* **228**, 784. (doi:10.1038/228784a0)

5 May, R. M. 1972 Will a large complex system be stable? *Nature* **238**, 413–414. (doi:10.1038/238413a0)

6 Dunne, J. A. 2005 The network structure of food webs. In *Ecological networks: linking structure to dynamics in food webs* (ed. J. A. D. Mercedes Pascual), Santa Fe Institute Studies in the Sciences of Complexity Proceedings, ch. 2, pp. 27–86. Cary, NC: Oxford University Press.

7 Bersier, L.-F. 2007 A history of the study of ecological networks. In *Biological networks* (ed. F. Képès), ch. 11, pp. 365–421. Hackensack, NJ: World Scientific.

8 Banašek-Richter, C. *et al.* 2009 Complexity in quantitative food webs. *Ecology* **90**, 1470–1477. (doi:10.1890/08-2207.1)

9 Chen, X. & Cohen, J. 2001 Global stability, local stability and permanence in model food webs. *J. Theor. Biol.* **212**, 223–235. (doi:10.1006/jtbi.2001.2370)

10 Rejmánek, M. & Stary, P. 1979 Connectance in real biotic communities and critical values for stability of model ecosystems. *Nature* **280**, 311–313. (doi:10.1038/280311a0)

11 Yodzis, P. 1980 The connectance of real ecosystems. *Nature* **284**, 544–545. (doi:10.1038/284544a0)

12 Schoener, T. W. 1989 Food webs from the small to the large. *Ecology* **70**, 1559–1589. (doi:10.2307/1938088)

13 Winemiller, K. O. 1990 Spatial and temporal variation in tropical fish trophic networks. *Ecol. Monogr.* **60**, 331–367. (doi:10.2307/1943061)

14 Martinez, N. D. 1992 Constant connectance in community food webs. *Am. Nat.* **139**, 1208–1218. (doi:10.1086/285382)

15 Schmid-Araya, J. M., Schmid, P. E., Robertson, A., Winterbottom, J., Gjerløv, C. & Hildrew, A. G. 2002 Connectance in stream food webs. *J. Anim. Ecol.* **71**, 1056–1062. (doi:10.1046/j.1365-2656.2002.00668.x)

16 McCann, K., Hastings, A. & Huxel, G. R. 1998 Weak trophic interactions and the balance of nature. *Nature* **395**, 794–798. (doi:10.1038/27427)

17 Pelletier, J. D. 2000 Are large complex ecosystems more unstable? A theoretical reassessment with predator switching. *Math. Biosci.* **163**, 91–96. (doi:10.1016/S0025-5564(99)00054-1)

18 Jansen, V. & Kokkoris, G. 2003 Complexity and stability revisited. *Ecol. Lett.* **6**, 498–502. (doi:10.1046/j.1461-0248.2003.00464.x)

19 Kondoh, M. 2003 Foraging adaptation and the relationship between food-web complexity and stability. *Science* **299**, 1388–1391. (doi:10.1126/science.1079154)

20 Brose, U., Williams, R. J. & Martinez, N. D. 2006 Allometric scaling enhances stability in complex food webs. *Ecol. Lett.* **9**, 1228–1236. (doi:10.1111/j.1461-0248.2006.00978.x)

21 Neutel, A.-M., Heesterbeek, J. A. P., van de Koppel, J., Hoenderboom, G., Vos, A., Kaldeway, C., Berendse, F. & de Ruiter, P. C. 2007 Reconciling complexity with stability in naturally assembling food webs. *Nature* **449**, 599–602. (doi:10.1038/nature06154)

22 Uchida, S. & Drossel, B. 2007 Relation between complexity and stability in food webs with adaptive behavior. *J. Theor. Biol.* **247**, 713–722. (doi:10.1016/j.jtbi.2007.04.019)

23 Otto, S. B., Rall, B. C. & Brose, U. 2007 Allometric degree distributions facilitate food-web stability. *Nature* **450**, 1226–1229. (doi:10.1038/nature06359)

24 Garcia-Domingo, J. L. & Saldana, J. 2008 Effects of heterogeneous interaction strengths on food web complexity. *OIKOS* **117**, 336–343. (doi:10.1111/j.2007.0030-1299.16261.x)

25 Rezende, E. L., Albert, E. M., Fortuna, M. A. & Bascompte, J. 2009 Compartments in a marine food web associated with phylogeny, body mass, and habitat structure. *Ecol. Lett.* **12**, 779–88. (doi:10.1111/j.1461-0248.2009.01327.x)

26 Paine, R. T. 1980 Food webs: linkage, interaction strength and community infrastructure. *J. Anim. Ecol.* **49**, 667–685.

27 Martinez, N. D. 1991 Artifacts or attributes? Effects of resolution on the Little Rock Lake food web. *Ecol. Monogr.* **61**, 367–392. (doi:10.2307/2937047)

28 Sugihara, G., Bersier, L.-F. & Schoenly, K. 1997 Effects of taxonomic and trophic aggregation on food web properties. *Oecologia* **112**, 272–284. (doi:10.1007/s004420050310)

29 Townsend, C. R., Thompson, R. M., McIntosh, A. R., Kilroy, C., Edwards, E. & Scarsbrook, M. R. 1998 Disturbance, resource supply, and food-web architecture in streams. *Ecol. Lett.* **1**, 200–209. (doi:10.1046/j.1461-0248.1998.00039.x)

30 Winemiller, K. O., Pinaka, E. R., Vitt, L. J. & Joern, A. 2001 Food web laws or niche theory? Six independent empirical tests. *Am. Nat.* **158**, 193–199. (doi:10.1086/321315)

31 Arim, M. & Jaksic, F. M. 2005 Productivity and food web structure: association between productivity and link richness among top predators. *J. Anim. Ecol.* **74**, 31–40. (doi:10.1111/j.1365-2656.2004.00894.x)

32 Kenny, D. & Loehle, C. 1991 Are food webs randomly connected? *Ecology* **72**, 1794–1799. (doi:10.2307/1940978)

33 Ulanowicz, R. E. & Wolff, W. F. 1991 Ecosystem flow networks: loaded dice? *Math. Biosci.* **103**, 45–68. (doi:10.1016/0025-5564(91)90090-6)

34 Rossberg, A. G., Yanagi, K., Amemiya, T. & Itoh, K. 2006 Estimating trophic link density from quantitative but incomplete diet data. *J. Theor. Biol.* **243**, 261–272. (doi:10.1016/j.jtbi.2006.06.019)

35 Drossel, B., McKane, A. J. & Quince, C. 2004 The impact of nonlinear functional responses on the long-term evolution of food web structure. *J. Theor. Biol.* **229**, 539–548. (doi:10.1016/j.jtbi.2004.04.033)

36 Rountree, R. A. 2001 Diets of NW Atlantic fishes and squid. See http://www.fishecology.org/diets/summary.htm. Accessed April 30, 2005.

37 Satoh, K., Yokawa, K., Saito, H., Matsunaga, H., Okamoto, H. & Uozumi, Y. 2004 Preliminary stomach contents analysis of pelagic fish collected by Shoyo-Maru 2002 research cruise in the Atlantic Ocean. *Col. Vol. Sci. Pap. ICCAT* **56**, 1096–1114.

38 Pinnegar, J. K. & Stafford, R. 2007 DAPSTOM—an integrated database & portal for fish stomach records. Final report, Centre for Environment, Fisheries & Aquaculture Science, Lowestoft, UK. See http://www.cefas.co.uk/dapstom (datasets ELMA01-94, CIROL03-92).

39 Bachok, Z., Mansor, M. & Noordin, R. 2004 Diet composition and food habits of demersal and pelagic marine fishes from Terengganu waters, east coast of Peninsular Malaysia. *NAGA, WorldFish Center Q.* **27**, 41–47.

40 Livingston, P. A. 2005 Monitoring and modeling predator/prey relationships. North Pacific Research Board

Final Report (grant no. R0305), Alaska Fisheries Science Center. See http://doc.nprb.org/web/03_prjs/.

41 Froese, R. & Pauly, D. 2009 FishBase. See http://www.fishbase.org (version July 2009).

42 OBIS. 2009 Ocean biogeographic information system. See http://www.iobis.org.

43 Kaschner, K. *et al.* 2008 Aquamaps: Predicted range maps for aquatic species. See http://www.aquamaps.org (version May 2008, accessed 28 February 2009).

44 Cheung, W., Alder, J., Karpouzi, V., Watson, R., Lam, V., Day, C., Kaschner, K. & Pauly, D. 2005 Patterns of species richness in the high seas. Technical report no. 20, Secretariat of the Convention on Biological Diversity, Montreal.

45 Mora, C., Derek, P. T. & Myers, R. A. 2008 The completeness of taxonomic inventories for describing the global diversity and distribution of marine fishes. *Proc. R. Soc. B* **275**, 149–155. (doi:10.1098/rspb.2007.1315)

46 Pinnegar, J. K., Blanchard, J. L., Steve Mackinson, R. D. S. & Duplisea, D. E. 2005 Aggregation and removal of weak-links in food-web models: system stability and recovery from disturbance. *Ecol. Model.* **184**, 229–248. (doi:10.1016/j.ecolmodel.2004.09.003)

47 Hill, R. C., Griffiths, W. E. & Lim, G. C. 2008 *Principles of econometrics*, 3rd edn. New York, NY: Wiley.

48 Jonsson, T., Cohen, J. E. & Carpenter, S. R. 2005 Food webs, body size, and species abundance in ecological community description. *Adv. Ecol. Res.* **36**, 1–84. (doi:10.1016/S0065-2504(05)36001-6)

49 Newman, M. E. J. 2005 Power laws, Pareto distributions and Zipf's law. *Contemp. Phys.* **46**, 323–351. (doi:10.1080/00107510500052444)

50 Chesson, P. 1994 Multispecies competition in variable environments. *Theor. Popul. Biol.* **45**, 227–276. (doi:10.1006/tpbi.1994.1013)

51 Bastolla, U., Lässig, M., Manrubia, S. C. & Valleriani, A. 2005 Biodiversity in model ecosystems, I: coexistence conditions for competing species. *J. Theor. Biol.* **235**, 521–230. (doi:10.1016/j.jtbi.2005.02.005)

52 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. (doi:10.1126/science.1068326)

53 Cohen, J. E. & Newman, C. M. 1985 A stochastic theory of community food webs: I. Models and aggregated data. *Proc. R. Lond. Soc. B* **224**, 421–448. (doi:10.1098/rspb.1985.0042)

54 Lloyd, M. 1967 Mean crowding. *J. Anim. Ecol.* **36**, 1–30. (doi:10.2307/3012)

55 Iwao, S. 1976 A note on the related concepts 'mean crowding' and 'mean concentration'. *Res. Popul. Ecol.* **17**, 240–242. (doi:10.1007/BF02530775)

56 Levins, R. 1968 *Evolution in changing environments.* Princeton, NJ: Princeton University Press.

57 Camacho, J., Guimerà, R. & Amaral, L. A. N. 2002 Robust patterns in food web structure. *Phys. Rev. Lett.* **88**, 228 102. (doi:10.1103/PhysRevLett.88.228102)

58 Laird, S. & Jensen, H. J. 2006 A non-growth network model with exponential and 1/k scale-free degree distributions. *Europhys. Lett.* **76**, 710–716. (doi:10.1209/epl/i2006-10319-x)

59 Rossberg, A. G., Matsuda, H., Amemiya, T. & Itoh, K. 2006 Food webs: experts consuming families of experts. *J. Theor. Biol.* **241**, 552–563. (doi:10.1016/j.jtbi.2005.12.021)

60 Rossberg, A. G., Matsuda, H., Amemiya, T. & Itoh, K. 2006 Some properties of the speciation model for food-web structure—mechanisms for degree distributions and intervality. *J. Theor. Biol.* **238**, 401–415. (doi:10.1016/j.jtbi.2005.05.025)

61 de Figueiredo, G. M., Nash, R. D. & Montagnes, D. J. 2007 Do protozoa contribute significantly to the diet of larval fish in the Irish Sea? *J. Mar. Biol. Ass. UK* **87**, 843–850. (doi:10.1017/S002531540705713X)

62 Camacho, J., Guimerà, R. & Amaral, L. A. N. 2002 Analytical solution of a model for complex food webs. *Phys. Rev. E* **65**, p030 901R. (doi:10.1103/PhysRevE.65.030901)

Supporting Appendices
to
**Universal power-law diet partitioning by marine fish and squid with surprising
stability-diversity implications**
by A. G. Rossberg, K. D. Farnsworth, K. Satoh, and J. K. Pinnegar

S1   Details of calculation of diet partitioning function

To correct for incomplete taxonomic resolution of resource species in diet data, we employ the
assumption that the probability that a species is resolved in the recorded diet of a consumer is
independent of the size of this species' contribution to the diet (Rossberg *et al.*, 2006). The as-
sumption is plausible since indiscernibility of species (one of the most frequently mentioned reasons
in the literature for lumping species into unresolved diet categories), is unrelated to diet fraction.
The number of unresolved resource species is then estimated from the number of resources that
were resolved in the diet tables. If, for example, 3/4 of diets are resolved to species level, and
these contain an average of 6 items contributing more than 1% to consumer diets, then we expect
that among the remaining unresolved 1/4 of diets there are another $6 \div (3/4) \times (1/4) = 2$ resource
species per consumer above the 1% threshold on average, hence the total consumer link density is
$6 \times (3/4 + 1/4) \div (3/4) = 6 \div (3/4) = 8$. Since diet tables generally list diets of taxonomic species,
taxonomic resolution of consumers is not an issue.

    This idea leads to the estimation procedure for the DPF used here (for the derivation, see
Rossberg *et al.*, 2006): From stomach-contents data (marine diet-tables), the fraction that each
diet item contributes to the diet of each consumer species was estimated by dividing the total
amount (mass of volume) found of this item in stomachs of a given consumer species by the total
amount of stomach content of this consumer species analyzed (excluding abiotoc matter). A list
$\{f_i\}$ was compiled of the numerical values of all the diet fractions for all the consumers in the
community, without regard for the identities of resources and consumers. Only those diet items
that were resolved to species level were included in this list (discarding those that were not). The
list was placed in rank order, such that $f_1 \geq f_2 \geq \ldots \geq f_n$. Then, with i the index of the last
fraction greater or equal to f, the function $\hat{Z}_c(f)$ defined by

$$\hat{Z}_c(f) = \frac{i}{\sum_{k=1}^{n} f_k} \tag{S1}$$

was used as an estimator for the DPF (Rossberg *et al.*, 2006). It has been shown that $\hat{Z}_c(f)$
approaches $Z_c(f)$ if the number of consumer species included in the analysis and the sum $\sum_k f_k$
over all resolved diet fractions are sufficiently large. For an analytical calculation of the error of
$\hat{Z}_c(f)$ in predicting $Z_c(f)$, see Rossberg *et al.* (2006).

    Consumer-specific DPF (figure 4 main text) were approximated as $(n + u) \times c$, where n is the
number of the consumer's species-resolved diet fractions $\geq f = r/(1 + r)$, u is a random number
uniformly distributed between -0.5 and 0.5, added to reduce overlaps among curves (but $u = 0$ for
$n = 0$), and c is a correction for unresolved diet items, identical for all consumers, chosen such that
the average over all consumer-specific DPF equals the comunity DPF, here curve A in figure 2.

    To compute the distribution of normalized generality (inset in figure 4), we set $u = 0$. The value
of c cancels out.

S2 Sources of diet data

The diet data used in this study came from a variety of sources. The two largest and best-resolved diet tables (sets A and B) describe gravimetric (A) and volumetric (B) diet fractions based of fish and squid sampled in annual bottom-trawl surveys between Cape Fear, North Carolina and Nova Scotia in 1973-80 (A, 29032 non-empty stomachs) and 1981-90 (B, 50027 non-empty stomachs) (Rountree, 2001). Notwithstanding the non-overlapping time intervals and the fact that experimental protocols and the focus of the analyses differed substantially (Rountree, 2001), one should caution that the two sets might not be fully independent. (Of 767 resolved prey and 146 predator species in set A, 641 prey and 53 predators are unique. Similarly in set B, of the 216 prey and 117 predators, 90 prey and 24 predators are unique.) Sets C and D represent mass fractions from stomach samples of large pelagic fish caught in longline operations in the open North (C, 188 non-empty stomachs) and Tropical (D, 357 non-empty stomachs) Atlantic Ocean from June to October 2002 (Satoh *et al.*, 2004). In set E gravimetric data from 5599 non-empty stomachs of commercial fish collected in two CEFAS otter trawl research cruises in the central and northern North Sea in 1992 and 1994 are combined (Pinnegar & Stafford, 2007). The gravimetric data-set F was sampled from 1439 stomachs (of which about 30% were empty) of pelagic and demersal coastal fish from the South China Sea and the east coast of Malaysia, landed in Kuala Terengganu, Malaysia between 1993 and 1994 (Bachok *et al.*, 2004). Finally, set G was obtained by gravimetric analyses of 17688 non-empty stomachs sampled on the western Bering Sea continental shelf, including the slope, on NMFS bottom and midwater surveys and commercial fishing vessels (Livingston, 2005).

S3 Curve fitting

Estimates of ν given in Tab. 3 are maximum likelihood fits making use of the known covariance structure of errors when estimating the DPF $Z_c(f)$ by Eq. (S1) and assuming errors in $\ln Z_c$ to be normal. Only data in the range $0.02 < f < 0.98$ is used in order to reduce biases due to unreported small diet fractions. The data in this range is sufficient to distinguish power laws from other functional relationships (Tab. 3).

The covariance of (sampling) errors of $\ln \hat{Z}_c(f)$ for different values of f can be estimated by the same method as was used in by Rossberg *et al.* (2006) to derive an expression for the standard (sampling) error of estimator (S1). Here we only state the result. Let a theoretical relation $Z_c(f)$ be given (as here for the purpose of curve fitting) and let $\theta(z)$ denote its inverse, i.e., $\theta(Z_c(f)) = f$. Then, for a DPF $\hat{Z}_c(f)$ estimated by Eq. (S1), one has for any two θ_1, θ_2 between 0 and 1, with abbreviations $\hat{Z}_1 := \hat{Z}_c(\theta_1)$, $\hat{Z}_2 := \hat{Z}_c(\theta_2)$, $Z_1 := Z_c(\theta_1)$, $Z_2 := Z_c(\theta_2)$,

$$
\begin{aligned}
\operatorname{cov}(\ln \hat{Z}_1, \ln \hat{Z}_2) &\approx \frac{\operatorname{cov}(\hat{Z}_1, \hat{Z}_2)}{Z_1 Z_2} \\
&\approx \frac{1}{S_c Z_1 Z_2} \int_0^{Z_1} \int_0^{Z_2} j\left[\min\left(\frac{z_1}{z_2}, \frac{z_2}{z_1}\right)\right] dz_1 \, dz_2 \\
&\quad + \frac{1}{p S_c \max(Z_1, Z_2)} \\
&\quad + \frac{(1-p)}{p S_c}\left[\int_0^\infty \theta(z)^2 \, dz - \frac{\int_0^{Z_1} \theta(z) \, dz}{Z_1} - \frac{\int_0^{Z_2} \theta(z) \, dz}{Z_2}\right],
\end{aligned}
$$

(S2)

where S_c is the number of consumer species included in the analysis, $p = S_c^{-1} \sum_{k=1}^{n} f_k$, and

$$j(x) := \frac{1}{4x^2} \int_0^\infty u E_1\left(\frac{u}{2x}\right) \times \left(u + \frac{4}{2 - u \exp(u/2) E_1(u/2)}\right) du. \tag{S3}$$

$E_1(x) := \int_x^\infty t^{-1} \exp(-t) dt$ denotes the exponential integral function. The function $j(x)$ can be approximated as

$$j(x) = 1 + 3.734x - 0.625x^2, \tag{S4}$$

which is, for $0 \leq x \leq 1$, accurate to about 1%. Equation (S2) describes the effect of sampling errors, that is, errors in estimating $Z_c(f)$ due to the fact that not all consumer species in the food web are sampled, and not all diet items of the sampled consumers are resolved to species level. It does not contain contributions from errors in measuring the f_i, since these have been shown to be comparatively small (Rossberg *et al.*, 2006).

To estimate the exponent ν, the likelihood of the empirical $\hat{Z}_c(f)$ was maximized with respect to ν, assuming true DPF as given by Eq. (3.2), and errors of $\ln \hat{Z}_c(f)$ that follow a multivariate normal distribution with a covariance matrix given by Eq. (S2). Integrals over $\theta(z)$ in Eq. (S2) were evaluated numerically. The other functional relationships listed in Tab. 3 were fitted accordingly, with the normalization given by Eq. (3.1). The known covariances were taken into account also when computing χ^2 and AIC statistics for goodness of fit.

The error estimates for ν in Tab. 3 are Cramér-Rao lower bounds, where the same normal approximation as above was used to compute the Fisher information (see Kay, 1993). We note that lower bounds are the conservative choice with respect to the hypothesis that ν is constant across communities.

Obviously, neighbouring points of $\hat{Z}_c(f)$ are highly correlated with each other. This can lead to numerical instabilities when inverting the covariance matrix for the computation of likelihoods. In order to enhance stability, successive data points of the empirical DPF $\hat{Z}_c(f)$ were substituted by their averages if the corresponding diet ratios differed by less then $q = 20\%$. As a result, d.o.f. in Tab. 3 can be considerably smaller than the number of resolved resources. It was confirmed that results are robust with respect to the particular choice of q.

S4 Details of richness estimation

All data sources we used for richness estimates have specific strength. Here, details of the methods used to obtain estimates of S_{fish} are described and discussed.

FishBase counts of fish species by Ecosystem have the advantages of representing direct observations. A major disadvantage is that ecosystem boundaries in FishBase do not always match the boundaries of study areas considered here. The "South China Sea" in FishBase, for example, does not cover the port of Kuala Terengganu where data-set F was collected. Yet, except for the open North and Tropical Atlantic (sets C and D), approximate correspondences can be established. To complement the FishBase data, counts of fish species from the OBIS occurrence observation database for lattice squares corresponding to study areas C and D were obtained (set C: 679 species in $30°N - 43°N \times 25°W - 60°W$; set D: 179 species in the union of $10°S - 10°N \times 20°W - 30°W$ and $5°N - 15°N \times 30°W - 45°W$).

AquaMaps (Kaschner *et al.*, 2008) provides, among others, estimates of the richness of Actinopterygii (spiny rayed fishes) at high spacial resolution. From these, average richnesses over

the study areas were computed (these averages depend only weakly on precise delineations of areas). AquaMaps estimates species-occurrence probabilities from environmental envelopes, and richness estimates count all species with over 50% probability of occurrence. Technically, this involves interpolations and extrapolations of occurrence data, and the same caveats apply as for any kind of extrapolation. Yet, these estimates reproduce the known latitudinal gradient in biodiversity and local variations, e.g., near biodiversity hotspots (Kaschner *et al.*, 2008). Estimates of richness from AquaMaps have the advantages of allowing arbitrary delineations of study areas while—by working with environmental envelopes—being more robust to data gaps in spacial sampling than OBIS data.

S5 Computation of $\mathbb{E}\overset{*}{c}_j$ from the DPF

We compute the expectation value of $\overset{*}{c}_j$, defined by Eq. (4.2). Denote by $p(f)\,\Delta f$ the probability that a randomly chosen consumer has a diet fraction in a narrow range $[f, f + \Delta f]$. If Δf is very small, there will be at most one such diet fraction in this range. The expectation value of the sum of the square of all diet fractions in this range can therefore be written as $f^2 p(f)\Delta f$. We obtain the expectation value of the sum of all squared diet fractions of a consumer in the range $[0, 1]$ by summing $f^2 p(f)\Delta f$ over this range, that is

$$\begin{aligned} \mathbb{E}\overset{*}{c}_j &= \mathbb{E} \sum_i f_{ij}^2 \\ &= \int_0^1 f^2 p(f) df. \end{aligned} \tag{S5}$$

Rossberg *et al.* (2006) showed that $p(f) = -Z_c'(f)$, with the prime denoting the derivative. Thus

$$\mathbb{E}\overset{*}{c}_j = -\int_0^1 f^2 Z_c'(f) df.$$

References

Bachok, Z., Mansor, M. & Noordin, R. 2004 Diet composition and food habits of demersal and pelagic marine fishes from Terengganu waters, east coast of Peninsular Malaysia. *NAGA, World-Fish Center Quarterly*, **27**(3), 41–47.

Kaschner, K., Ready, J. S., Agbayani, E., Rius, J., K. Kesner-Reyes, P. D. E., South, A. B., S. O. Kullander, T. R., Close, C. H., Watson, R. *et al.* 2008 Aquamaps: Predicted range maps for aquatic species. World Wide Web electronic publication, www.aquamaps.org. Version 05/2008, accessed 28/02/2009.

Kay, S. M. 1993 *Fundamentals of statistical signal processing: Estimation theory*, chap. 3, p. 47. Prentice Hall.

Livingston, P. A. 2005 Monitoring and modeling predator/prey relationships. North Pacific Research Board Final Report (grant number: R0305), Alaska Fisheries Science Center. Data available at `http://doc.nprb.org/web/03_prjs/`.

Pinnegar, J. K. & Stafford, R. 2007 DAPSTOM - An integrated database & portal for fish stomach records. Final report, Centre for Environment, Fisheries & Aquaculture Science, Lowestoft, UK. `http://www.cefas.co.uk/dapstom`, datasets ELMA01-94, CIROL03-92.

Rossberg, A. G., Yanagi, K., Amemiya, T. & Itoh, K. 2006 Estimating trophic link density from quantitative but incomplete diet data. *J. Theor. Biol.*, **243**(2), 261–272.

Rountree, R. A. 2001 Diets of NW Atlantic fishes and squid. `http://www.fishecology.org/diets/summary.htm`. Accessed Apr 30, 2005.

Satoh, K., Yokawa, K., Saito, H., Matsunaga, H., Okamoto, H. & Uozumi, Y. 2004 Preliminary stomach contents analysis of pelagic fish collected by Shoyo-Maru 2002 research cruise in the Atlantic Ocean. *Col. Vol. Sci. Pap. ICCAT*, **56**(3), 1096–1114.

Supporting Material
for
Universal power-law diet partitioning by marine fish and squid with surprising stability-diversity implications
by A. G. Rossberg, K. D. Farnsworth, K. Satoh, and J. K. Pinnegar

S6 Supplementary Discussion

In the following we address several concerns readers might have in relation to the results presented in the main text. Part of the discussion here provides additional support to some steps in our main argument, even though the conclusions of the main text do not require this. Specifically, the observation of a power-law DPF does not hinge on its reproducibility under certain model assumptions. The main purpose of this supporting discussion is to clarifying perceived incompatibilities between our conclusions and other views widely held. To address a broad variety of points that may come to the reader's mind and, at the same time, offer the maximum possible orientation, we follow the highly appraised example by Horwich (1998) and formulated this discussion in a question-answer format.

S6.1 Questions concerning the power-law form of the DPF

Is the power law form of the DPF, equation (3.2)**, plausible beyond the mere fact that it fits empirical data?** Yes. A simple stochastic model can explain why DPF follow approximate power laws for communities containing many potential resource species and being characterized by complex trophic interactions. Assume that, in order for a consumer to forage on a resource, several specific conditions need to be satisfied, and that the abundance-independent interaction strength or *affinity* a_{ij} of consumer j to resource i is given by a product of indices for the degree to which each of these conditions is met. This product structure of affinities follows, for example, from a model for affinities (there termed "trophic interaction strengths") by Rossberg *et al.* (2010). For randomly chosen i and j, each of the indices contributing to a_{ij} becomes a random number. (Symbols used in this discussion are summarized in Tab. S1.) If the number of conditions to be met is large, a_{ij} then follows, by virtue of the central limit theorem, an approximate log-normal distributions. Such a distribution of affinities was observed by Goldwasser & Roughgarden (1993) for an island food web. Biomass densities B_i of species are also known to follow approximate log-normal distributions. Let the average contribution of i to the diet of j be given by $m_{ij} = a_{ij}B_iF_j$, with the factor F_j describing, e.g., saturation of the consumer. The diet contributions m_{ij} are then, up to the factor F_j, which cancels out when computing r_{ij} or f_{ij}, log-normally distributed. Appendix A below analyzes the DPF resulting from consumers foraging on a set of S_r potential resources, with resource i contributing an amount

$$m_{ij} = \exp(\sigma\xi_{ij})F_j \tag{S6}$$

to the diet of consumer j, where ξ_{ij} are independent, standard-normal random numbers and the parameter σ controls the spread of the log-normally distributed *availabilities* $a_{ij}B_i$ $(= \exp(\sigma\xi_{ij}))$. It is shown that, over the reliable range in r, the DPF follows a power law $Z_c \propto r^{-\nu}$ with

$$\nu \approx \sigma^{-1}\sqrt{2\ln S_r}, \tag{S7}$$

1

Table S1.   List of symbols used in Supplementary Discussion

Symbol	Meaning
β	Slope of linear relation
$\hat{\beta}$	Regression slope
ν	Exponent for scaling of Z_c with r
σ	std $\ln a_{ij} B_i$
σ_a	std $\ln a_{ij}$
σ_B	std $\ln B_i$
ξ_i, ξ_{ij}	Standard normal random variables
a_{ij}	Affinity of j to resource i
$a_{ij} B_i$	Availability of i to j
B_i	Biomass density of i
f_{ij}	Diet fraction: relative contribution of i to j's diet
f	Diet-fraction threshold
F_j	Feeding intensity of j
q	Relative error of species-richness data
m_{ij}	Contribution of i to j's diet
n	Sampled stomachs per consumer
r	Diet-ratio threshold $= f/(1-f)$
r_{xy}	Correlation coefficient
S_r	Resource species richness
S_{fish}	Fish species richness
s_x, s_y, s_ν	Sample standard deviations of x, y, ν
$\overline{x}, \overline{y}, \overline{\nu}$	Sample means of x, y, ν
x	Some relative measure of diversity
y	Some characterization of link density
Z_c	Links per consumer
$Z_c(f)$	Diet partitioning function (DPF)
Z_c	Links per consumer
Z_{sample}	Sampled links per consumer "without threshold"

provided the spread σ and the number of potential resources S_r are both large. Equation (S7) was confirmed in simulations of this log-normal model (figure S1, solid lines). An ecological interpretation of equation (S7) is offered in response to the next question.

A method to directly compute accurate DPF for model (S6) and similar models numerically, valid for small and large σ and S_r and over the full range in r, is described in Appendix B below. The method is demonstrated in figure S1 (dotted lines) and used below to evaluate model (S6) for small σ and S_r.

Why don't you try to estimate S_r or σ based on ν and equation (S7) as a plausibility check?   Equation (S7) suggests an ecological interpretation that links the scope of an empirical investigation to the specificity of diet choice within this scope and the exponent ν characterizing the DPF. This interpretation implies that, while scope and specificity, quantified by S_r and σ,
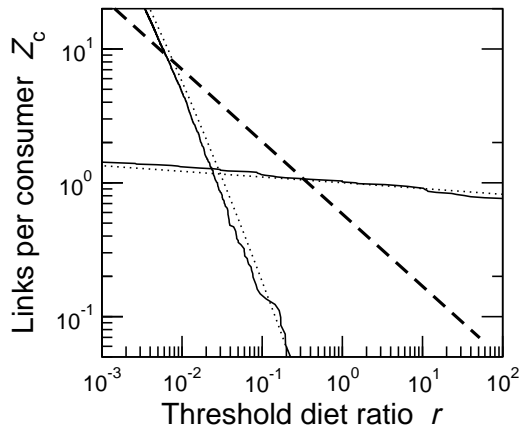
Figure S1. Theoretical diet partitioning functions. Solid lines are simulation results for the log-normal model [Flat: $S = 1000$, $\sigma = 75$, and 80 consumer species; slope corresponds to exponent $\nu = 0.057$, compared to 0.050 predicted by equation (S7). Steep: $S = 10^6$, $\sigma = 3.3$, 80 consumers; $\nu = 1.34$ (1.59 predicted)]; dotted lines are corresponding accurate numerical results (Appendix B); the dashed line is as in figure 2, for comparison.

depend on the observation protocol, ν is an inherent system property. To prepare discussion of this interpretation, we first argue quantitatively that the dominating contribution to the spread σ of availabilities $a_{ij}B_j$ entering equation (S7) is the spread of affinities (a_{ij}) rather than that of biomass densities (B_j). Since the spread of affinities is not directly accessible, it is here estimated indirectly *via* equation (S7).

Observations by Petchey *et al.* (2004) for small freshwater ecosystems suggest that species richness decreases by approximately a factor of three with each trophic level. Assuming, conservatively, a richness of resources $S_r = 3S_{\text{fish}}$, equation (S7) predicts a spread in availabilities $a_{ij}B_j$ of $\sigma = \text{std} \ln(a_{ij}B_j) \approx 6.3$ for low-diversity marine ecosystems with $S_{\text{fish}} \approx 100$ and $\nu \approx 0.54$. Since S_r is comparatively small, the accurate numerical evaluation of the log-normal model (S6) differs somewhat from equation (S7), rather suggesting a spread of affinities $\sigma \approx 4.3$. The spread of biomasses B_i among species is surprisingly rarely quoted in the literature, but graphical data (Cyr, 2000; Connolly *et al.*, 2005) indicate that $\sigma_B = \text{std} \ln B_j = 2$ at most. Thus, ignoring correlations between a_{ij} and B_j, so that $\sigma^2 = \sigma_a^2 + \sigma_B^2$, we obtain $\sigma_a = \text{std} \ln a_{ij} = (4.3^2 - 2^2)^{1/2} = 3.8$ as a lower bound for the spread in affinities. Since $\sigma_a^2 = 14.4$ is considerably larger than $\sigma_B^2 = 4$ even by these extremely conservative estimates, the total spread in availabilities $a_{ij}B_j$ is dominated by the spread in a_{ij}. As a caveat, the effect of B_i could be somewhat enhanced by prey switching behaviour (Greenwood & Elton, 1979). Our conclusion that availabilities are dominated by affinities rather than biomasses is consistent with the general observation that prey biomasses are bad predictors of prey intake (Winemiller, 1990; Pinnegar *et al.*, 2003; Quince *et al.*, 2005; de Figueiredo *et al.*, 2007).

Natural communities, especially marine ones, are rarely sharply bounded. General macroecological phenomena, such as the form of the DPF, should not depend on the observation protocol delimiting a community. Hence, when measuring the *scope* of analysis by of the number S_r of potential resource species included, and when measuring the *specificity* by which consumers choose among these resources by σ_a^2, equation (S7) implies a relation between the scope and the specificity

3

of diet choice of the form $\sigma_a^2 = 2\nu^{-2} \ln S_r - \sigma_B^2$ for a given set of consumers whose diet partitioning exponent is ν, under the assumption of uncorrelated affinities and biomasses. Consider, for example, two empirical or theoretical analyses of diet choice, one being restricted to a narrow scope covering S_r potential resource species, the other having a wider scope covering $S'_r > S_r$ species. The narrow and the wider scope, respectively, may be given, for example by the sunlit epipelagic zone and the full water column; or a specific taxonomic group and the full range of species inhabiting a prescribed volume; or a narrow and a wider range in the resources' adult body masses. Assume that, for the given set of consumers considered, the realized diet is nearly completely contained in the narrow scope, such that inclusion of diets from the wider scope will not change ν. Define $u = S'_r/S_r$. Equation (S7) then predicts that, correspondingly, the specificity σ_a^2 of diet choice from the extended scope will be larger than from the narrow scope by an amount $2\nu^{-2} \ln u$. This can be interpreted intuitively in terms of an addition of criteria for establishing consumer-resource links (e.g., a match of depth or prey size), which corresponds, in the picture developed in response to the previous question, to multiplying further indices depending on resources, consumers, and the environment to the affinities a_{ij}, and thus increasing the variance σ_a^2 of $\ln a_{ij}$ over consumer-resource pairs.

In conclusion, depending on the scope taken into consideration for resources, the value of the specificity σ_a^2 will change, at least as long as the smallest scope considered contains already all relevant resources. The numerical values of σ_a and of the closely related spread σ of availabilities therefore do not have much ecological significance. More important are the DPF itself and its characterization in terms of the exponent ν.

The curves of the empirical DPF plotted in figure 2 can be approximated by power laws over a certain range, but for $r < 0.01$ there are clear deviations. Doesn't this speak against power laws? The log-normal model (S6) for diet contributions introduced above does in fact imply that the power law holds only over a limited range. For very small f, the consumer link density Z_c in this model converges to the total number S_r of available resources. This implies that $Z_c(0)$ increases linearly with S_r. This model property is not necessarily unrealistic. Since even objects of doubtful nutritious value such as plastic bags, pieces of fiberglass, or rubber bands are occasionally found in fish stomachs (e.g., Rountree, 2001), it would be no surprise if, over sufficiently long times, bits and pieces of most species will ultimately be consumed by some individuals of any other or the same species.

Yet, the leveling-off of empirical curves in figure 2 for low diet ratios is not necessarily related to deviations from power laws by the true DPF of the study systems. The leveling-off can simply be the result of incomplete sampling. Appendix C below develops a primitive sampling theory, which assumes, simplistically, that prey size equals stomach size, that each non-empty stomach contains exactly one species, and that exactly n non-empty stomachs per consumer are sampled. The theory makes predictions Z_{sample} for the value $\hat{Z}_c(0)$ that empirical curves as in figure 2 would reach for very small r. Two different results are obtained: First, an integral formula for Z_{sample}, equation (S20), is derived, which is then approximated to $Z_{\text{sample}} \approx Z_c(1/n)$. Predicted Z_{sample} and observed $\hat{Z}_c(0)$ of the seven empirical data-sets considered here, assuming true DPF of power-law form with $\nu = 0.54$, are plotted in figure S2 against the number of non-empty stomachs per consumer n in data-sets A-G. Both, predictions by equation (S20) and observations exhibit a clear positive trend with increasing n consistent with the approximation $Z_{\text{sample}} \approx Z_c(1/n) \propto n^\nu$. Observed values $\hat{Z}_c(0)$ tend to be rather larger than predicted Z_{sample} than smaller. This underestimation of empirical $\hat{Z}_c(0)$ by the sampling theory appears attributable to the fact that non-empty stomachs often contain more
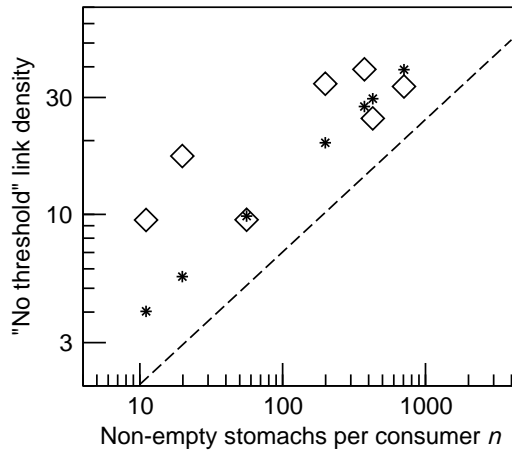
4

Figure S2. Empirical and theoretical mean numbers of distinct diet species found after sampling n non-empty stomachs per consumer. Diamonds: empirical $\hat{Z}_c(0)$ for data-sets A-G, stars: Z_{sample} by equation (S20), dashed line: $Z_c(1/n)$. Theoretical values are based on power-law DPF with $\nu = 0.54$.

than the single prey item assumed in Appendix C. If, on the other hand, true DPF $Z_c(f)$ would flatten off shortly below $r = 0.02$, deviating from power laws, one would expect observed $\hat{Z}_c(0)$ to be lower than the Z_{sample} values predicted for power-law DPF. Thus, the empirical curves rather speak against deviations of true DPF from power laws for r shortly below 0.02.

Link & Almeida (2000), cited by Link (2002), report that in their marine stomach sampling survey "The number of stomachs [of a predator] examined versus the number of prey items observed generally indicated that an asymptote was reached between 500 and 1,000 stomachs." Doesn't this speak in favor of well defined link absence and presence, contradicting the idea of power-law diet partitioning? Looking at the 15 sampling curves presented by Link & Almeida (2000) in support of this statement, one sees that, while all curves start steeply and then become flatter, some with one or more pronounced bends, none of the curves ever reaches saturation, except perhaps for that of Yellowtail Flounder. Even after examining over 15,000 stomach of Spiny Dogfish and finding over 350 distinct prey items, new prey items are still encountered by Link & Almeida (2000). Discounting for the bends, the overall shapes of the curves are in good qualitative and, by order of magnitude, semi-quantitative agreement with the primitive sampling formula derived in Appendix C, $Z_{\text{sample}} \approx Z_c(1/n) \propto n^{0.54}$. The data by Link & Almeida (2000) support our picture that $Z_c(f)$ approaches a large value of the order of species richness when $f \to 0$, i.e., when even the rarest prey items are counted. Similar conclusions can be drawn regarding the sampling curves by Schmid-Araya *et al.* (2002*a*) for the gut contents of aquatic macroinvertebrates.

The fitted values you find for the power-law exponent ν appear to differ from what one would get from simple regressions of $\log Z_c$ against $\log r$ in figure 2, or from drawing lines by hand. Where does this difference come from? Our fitting procedure takes the known covariance structure of empirical DPF into account. For large r this method puts emphasis

on fitting absolute values, for small r rather on fitting the slopes, because there the variance from the denominator in equation (S1) dominates.

S6.2 Questions concerning the dependence of the DPF on species richness

Linear regression can establish the existence of a—possibly indirect—relationship between two variables, but it can never establish the absence of relations. How can you conclude that ν does not depend on $S_{\mathbf{fish}}$? The structure and nature of our argument is similar to that employed in establishing the "constancy" of other constants of nature. For example, there is broad scientific agreement that the speed of light (in vacuum) is constant, wherever it is measured on earth. [Evidence for this fact is so strong that since 1983 the speed of light is used to define units of length indirectly in terms of units of time. We disregard this complication here, thus describing the situation before 1983.] This "absence of dependence on location" was established by measuring the speed of light at different locations, estimating bounds on the accuracy of each measurement and verifying that the measured speeds were equal within the measurement errors. The accurate statement of this result is therefore to say that the speed of light is constant within measurement errors, thereby quantifying the errors. As a second step, assuming a constant speed of light, one can then deduce an optimal value for the speed of light (plus an error estimate) by combining the available measurements. Along an analogous route, the value $\bar{\nu} = 0.54(2)$ was computed here. Of course, such analyses become stronger the more high-accuracy measurements they incorporate. In this respect, our current results leave room for improvements. Yet, they are non-trivial and informative, as is argued in response to the next question.

Any independence arguments can be taken one step further when a specific hypothesis for the nature of a suspected dependence exists. The scientific nature of this additional step shall again be illustrated using the example of the speed of light. Assume that some speculative physical theory implied that the speed of light changed slightly with the latitude at which it is measured. It would then be natural to re-analyze existing measurements to search for such an effect, despite the fact that previous analyses had established that the speed of light is constant within measurement errors. Why this? The restriction to one particular explanatory variable increases the statistical power of the analysis. Thus, even though the speed of light was found to be constant within measurement errors, a comparison of these measurements with the latitude where they were taken could still reveal a statistically significant correlation. On the other hand, if such a correlation cannot be established, the statistical bounds on conceivable remaining, minute dependencies on latitude can become smaller than the errors of individual measurements, thus further constraining the parameter range of the theory in question. We performed an analogous analysis to search of a dependence of ν on $S_{\mathbf{fish}}$. No dependence could be found.

Mora _et al._ (2008) demonstrated large data gaps in inventories of marine fish. Isn't it possible that insufficient accuracy of your species richness ($S_{\mathbf{fish}}$) estimates masks a dependence of ν on $S_{\mathbf{fish}}$? Consider first the richness estimates based on FishBase+OBIS. The analysis by Mora _et al._ (2008), which is based on essentially the same raw data, estimates completeness of species inventories for geographic lattice squares of different sizes by extrapolations of sampling curves. While many parts of the oceans appear to be poorly sampled, the study sites considered here are largely located in the better-sampled areas. Mora _et al._ (2008) estimate that taxonomic inventories of marine fish for the North West Atlantic Shelf (data sets A, B), the Open
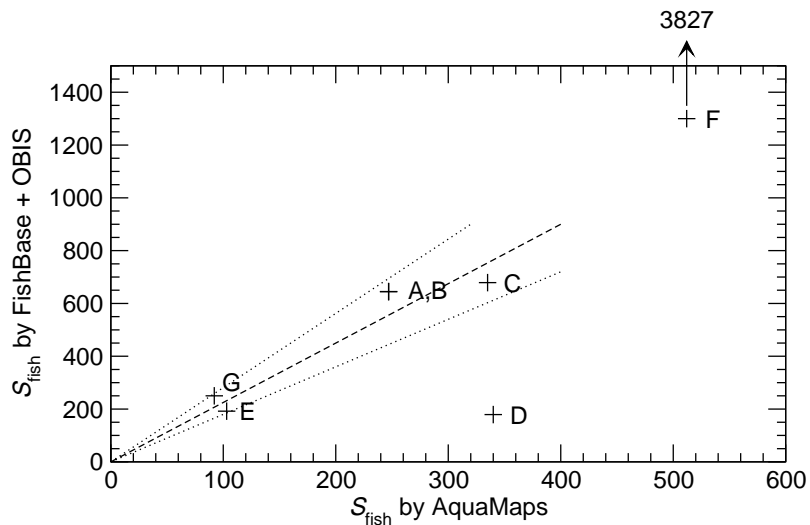
Figure S3. Comparison of estimated relative species richnesses by AquaMaps and FishBase+OBIS. Characters A to G indicate study areas (Tab. 2). The dashed line has a slope of 2.25; the two dotted lines have slopes of 2.25×0.8 and $2.25/0.8$, respectively; all lines go through the origin. The analysis shows that relative richness estimates by AquaMaps and FishBase+OBIS are consistent with each other to within 20%, except for the deviations for study sites D and F discussed in *Materials and methods: Species richness*.

North Atlantic (set C) and the Eastern Bering Sea (set G) are to 78% or more complete at 36° resolution, and to more than 78% complete in the North Sea (set E) and the South China Sea around Kuala Terengganu (set F) at 18° resolution. Sampling in the 18° lattice square to the east of Kuala Terengganu (also part of the South China Sea) is estimated to be at least to 67% complete. The estimated taxonomic completeness in the area of the Open Tropical Atlantic corresponding to set D is above 67% at 36° resolution. One has to caution that these figures are estimates and true completeness may differ in either direction. Yet, this data suggests that the relative errors of our estimates of fish richness are on the order of $20 - 30\%$ compared to exhaustive sampling. Since these errors are one-sided, the relative errors of using fish richness as estimates of relative richness across study sites can be expected to be even smaller; and relative richness is what matters most here.

Since the methodology underlying AquaMaps is more complicated than a simple accumulation of inventories, the accuracy of richness estimates derived thereof is more difficult to asses by direct methods. In figure S3 we compare richness estimates by AquaMaps with those from FishBase+OBIS. AquaMaps estimates are consistently by a factor of about 2.25 smaller than FishBase+OBIS estimates, except for sets D and F, which we discussed in *Material and methods: Species richness* (main text). For sets A, B, C, G, and E, deviations from this linear relationship are smaller than the $20 - 30\%$ one would expect from the considerations above. We conclude that, with qualifications for sets D and F, $20 - 30\%$ is a conservative bound for the relative errors of relative richness estimates by both AquaMaps and FishBase+OBIS.

Applying the standard propagation-of-errors formula to the formula for regression slopes, and taking into account that sets A and B refer to the same location, we computed that a relative error of $(q \times 100)\%$ in richness estimates leads to an additional uncertainty of $0.00030\,q$ in the regression

7

slope of ν *vs.* S_{fish} using AquaMaps, and of $0.000038\,q$ using FishBase+OBIS with set D, and of $0.000025\,q$ without it. Even for a conservative $q = 0.3$, these uncertainties only slightly increase the total uncertainty of the slopes. For AquaMaps, for example, the estimated standard error of the slope increases from 2.7×10^{-4} to $[2.7^2 + (3.0 \times 0.3)^2]^{1/2} \times 10^{-4} = 2.8 \times 10^{-4}$. For FishBase+OBIS, the estimated standard error increases from 3.6×10^{-5} to 3.7×10^{-5} when including set D, and has no significant numerical effect when excluding set D.

Diversity among the species forming a community might be poorly characterized by S_{fish}. Wouldn't is be possible that regression of link density against a more appropriate characterization of diversity reveals a pronounced dependence of the former on the latter? A general mathematical argument suggests that this is unlikely. Let x be some measure of diversity and y some characterization of trophic link density. Simple linear regression of k empirical data-points (x_i, y_i) $(i = 1, \ldots, k)$ yields two distinct characterizations of the dependence of y on x: the sample correlation coefficient r_{xy} and an estimate $\hat{\beta}$ of the slope β of an assumed linear dependence of y on x. Crucial for the validation of theories relating diversity to link density (or complexity) is not the value r_{xy}, but the value of $\hat{\beta}$, possibly expressed as the elasticity $\hat{\beta}\overline{x}/\overline{y}$ at the sample means $\overline{x}$ and $\overline{y}$ (Hill *et al.*, 2008), to obtain a dimensionless quantity: Provided a linear fit is justified at all, observation of a strong correlation between x and y, i.e. $|r_{xy}| \approx 1$, but of a small slope, i.e., $|\hat{\beta}\overline{x}/\overline{y}| \ll 1$, would support theories predicting that y is essentially independent of x, because, as x is varied in observations, y changes only little. The strong correlation between x and y, even when statistically significant, would signify only a numerically small perturbation of the non-dependence of y on x. In ordinary least-square regression, $\hat{\beta}$ and r_{xy} are related as

$$\hat{\beta} = r_{xy}\frac{s_y}{s_x}, \tag{S8}$$

where s_x and s_y are the sample standard deviations of x and y. Since r_{xy} is bounded to lie between -1 and 1, a small value of s_y/s_x alone already implies a small value of $|\hat{\beta}|$. When the sample ecosystems included in the analysis are a mix of low-, intermediate-, and high-diversity systems, so that the sample standard deviation s_x is of the order of magnitude of the sample mean $\overline{x}$, smallness of $s_y/\overline{y}$ alone is sufficient to ensure smallness of $|\hat{\beta}\overline{x}/\overline{y}| = |r_{xy}(s_y/\overline{y})/(s_x/\overline{x})|$, for any value of r_{xy} and independent of the specific values x_i. Of course, improvements in the accuracy of diversity measures x_i over applying simple rules of thumb will improve estimates of β. But even rough diversity measures can yield reliable results when the empirical coefficient of variation of the characterization of link density, $s_y/\overline{y}$, is small. This is the situation encountered here. The small value of $s_\nu/\overline{\nu} = 0.06$ found leaves little room for a dependence of ν on any measure of diversity. From these considerations, the weak effect of errors in S_{fish} on regression slopes, computed in response to the previous question, is also easily understood.

Couldn't the narrow range of values for ν you find be an artifact of the definition of the DPF? In what sense is the range narrow? From our model explaining power law DPF, it follows that, in principle, DPF with arbitrary values of $\nu > 0$ can be obtained. Figure S1 demonstrates DPF with $\nu = 0.057$ and $\nu = 1.34$. Compared to this range, the width of the 95% confidence interval for ν in figure 3c of about 0.2 is indeed narrow. Narrow, compared to 1, are also the error bounds obtained for the elasticity of the dependence of ν on species richness.

Comparisons of the current findings with relations between link density and diversity proposed
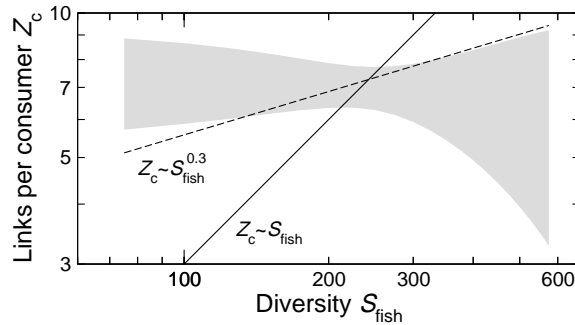
Figure S4. The 95% confidence interval for consumer link densities for the diet fraction threshold $f = 0.01$ (shaded) in comparison with steep (solid) and flat (dashed) power-law scaling relations proposed in the literature. Both axes have logarithmic scales.

in the literature can also be considered. However, it is not immediately clear how the diet-fraction threshold f should be chosen in such comparisons. Speculatively setting $f = 0.01$, independent of S_{fish}, figure S4 maps the 95% confidence region for ν (figure 3c) onto corresponding consumer link densities via equation (3.2). As illustrated in figure S4, power-law relations of the form $Z_c \propto S_{\text{fish}}^{\alpha}$ with $\alpha = 0.3$ (Schmid-Araya *et al.*, 2002*b*) are extremely unlikely in this picture, and for larger α (Dunne, 2005; Martinez, 1992) clearly fall outside the confidence region, independent of the choice of the coefficient of proportionality (which shifts the straight lines in figure S4 up and down).

S6.3   Questions concerning thresholding of trophic links

Why don't you simply count all trophic links? We are unaware of clear empirical demonstrations of a lower bound for the strength of trophic links, or of a bound below which weak links become highly unlikely (reports of saturating stomach/gut sampling curves were discussed above). This may be so because no such bounds exist, or because they exist but are too low to be observable without sampling nearly every stomach in a population. In both cases, we cannot observe the weakest links, and hence cannot count them.

Besides these epistemic constraints, it is conceivable—if not likely—that the sharp ontological distinction between present and absent trophic links cannot be upheld when trophic link strengths are spread out widely on logarithmic scales.

Finally, it is at least a plausible conjecture that strong trophic links are ecologically more important than weak ones, and that an investigation of strong links alone provides some ecologically useful information. See, however, the next question.

Does it make sense to consider only strong links in your analysis, even though weak links are known to be important for food-web dynamics? Weak links may be important, but perhaps not that important: For the theoretical models of community stability and dynamics that we are aware of, a very weak link, i.e., the limit (link strength) $\rightarrow 0$, is equivalent to an absent link. Even in the scenarios displaying stabilizing effects of weak links described by McCann *et al.* (1998), links weaker than a few percent by the relative measure of McCann *et al.* have negligible effects on dynamics.

9

Yet, it is conceivable in principle that for some ecological mechanisms the combined effect of all weak trophic links in a community is stronger or comparable to the combined effect of the strong links that we consider here. Deciding whether this is the case or not will require, among others, knowledge of the statistical distribution of link strengths. Our work contributes to understanding this distribution. For the reasons given above, the empirical characterization of this distribution towards weaker links must remain incomplete here.

When introducing a diet-fraction threshold for counting link, shouldn't one lower this threshold with increasing community size to allow for larger dietary diversity in more species rich communities? Since, historically, the diversity-stability-complexity debate had build on the ideas of well-defined link absence and presence, it remained largely mute on the question if a link-strength threshold should be kept fixed of be adjusted (see, however, Kartasche *et al.*, 2009). By expressing results directly in terms of the DPF, the current approach remains open to both options. Our proposal to formulate contributions to the diversity-stability-complexity debate directly in terms of link-strength distributions such as the DPF, rather than in terms of link counts, would make this question obsolete.

Since generalist consumers partition their diets into more and therefore smaller fractions than specialists, shouldn't the diet-fraction threshold be lowered for generalists compared to that for specialists when counting the number of their prey species? Is the DPF a suitable characterization of dietary diversity at all? While we sympathize with the sentiment behind this suggestion, it is not immediately practicable. An obvious technical difficulty would be to consistently determine generality prior to setting the thresholds for counting links. The absolute and relative degree of generality of consumers will depend on the specific measure of generality used.

Many conceivable measures of generality can be defined in terms of the set of a consumer's diet fractions, that is, the consumer's diet partitioning (e.g., Levins, 1968; Bersier *et al.*, 2002). A consumer's diet partitioning, in turn, is fully described by the consumer-specific diet partitioning function (§b, main text). Ideally, dietary diversity in a community would therefore be characterized by the full set of consumer-specific DPF (figure 4). The (proper) DPF $Z_c(f)$ considered here is a summary statistic for this set, obtained by averaging over all consumer-specific DPF (in practice, only over a selected sub-set). Other summary statistics will have their own strengths and weaknesses, but none will convey all the information contained in the full set of consumer-specific DPF.

This perspective on the DPF stresses that the whole function $Z_c(f)$ is the statistic of interest, rather than its value at particular points f. Comparisons of DPF at specific points f, implied by above question, are not part of our main argument.

In fact, assuming power laws, any characterization of the DPF by the consumer link density $Z_c(f)$ at a fixed threshold f would be ambivalent, because $Z_c(f)$, as given by equation (3.2), is an unimodal function of the parameter ν. For $f = 0.01$, for example, it has a maximum $Z_c(f) = 9.4$ at $\nu = 0.76$.

Diet fractions as small as 0.001 can and have been empirically determined (Baird & Ulanowicz, 1989; Goldwasser & Roughgarden, 1993). Why are only diet fractions larger than 0.02 considered in your analysis? Relevant for our analysis is not the question if some diet fractions of a given size can be found, but if most of them can be found (precisely: if

the probability of finding them is not constrained by their size). By the arguments in Appendix C, exhaustive sampling of diet fractions of size f requires on the order of $n \approx 1/f$ non-empty stomach samples per consumer. Applying this criterion to our data, we find that for sets C and D, with $n = 11$ and 20, i.e., $1/n \approx 0.09$ and 0.05, the lower bound 0.02 might rather be too low than too high. This may partially explain the observed weak dependence of measured diet-partitioning exponents ν on sampling effort. We chose the threshold as low as 0.02 in order to capture more information from other data-sets, and because more precise calculations (stars in figure S2) and the saturation points in figure 2 suggest that the estimate $n \approx 1/f$ is a bit too pessimistic.

There may be practical limits to what is feasible, but since the lower-bound $f = 0.02$ you choose for the fitting range does not even capture the mean diet fraction of some highly generalist consumers, isn't it, nevertheless, too high? It is easily seen that the mean diet fraction of a consumer is simply the inverse of the number of its prey species. This measure therefore hinges on a sharp distinction between presence and absence of trophic links. But such a distinction appears difficult, if not impossible, for the epistemic and ontological reasons given above. The "mean diet fraction" is therefore not sufficiently well defined to provide orientation for setting a diet-fraction threshold.

A Analytic Approximation of the DPF for independent, log-normally distributed intakes

We derive equation (S7). Consider a system with S_r potential resource species and chose one consumer j at random. Assume the intakes m_{ij} of species i by this consumer to follow, up to a constant factor, independent, identical log-normal distributions $\ln m_{ij} \sim \mathcal{N}(\mu, \sigma)$, that is, all m_{ij} have the cumulative distribution function (c.d.f.) $P_M(m) := P[m_{ij} < m] = \Phi((\ln(m) - \mu)/\sigma)$, where $\Phi(x) = [1 + \mathrm{erf}(x/\sqrt{2})]/2$ is the cumulative standard normal distribution. Diet ratios are defined as

$$r_{ij} = \frac{m_{ij}}{\sum_{k \neq i} m_{kj}}, \tag{S9}$$

with k going from 1 to S_r, excluding i. Since the effect of μ cancels out in this expression, we can set $\mu = 0$ without loss of generality. Numerator and denominator of r_{ij} are statistically independent. For large σ, the fat tail of the log-normal distribution allows the distribution of the denominator in equation (S9) to be approximated by assuming that the sum is dominated by the largest term. In this case, the sum is smaller than m if an only if all addends are $< m$:

$$P\left[\sum_{k \neq i} m_{kj} < m\right] \approx \prod_{k \neq i} P[m_{kj} < m] = [\Phi(\ln(m)/\sigma)]^{S_r - 1}. \tag{S10}$$

The c.d.f. of the logarithm of the sum $Y := \ln \sum_{k \neq i} m_{kj}$ can be approximated as

$$\begin{aligned} P_Y(y) &:= P[Y < y] = [\Phi(y/\sigma)]^{S_r - 1} \\ &= [1 - (1 - \Phi(y/\sigma))]^{S_r - 1} \approx \exp\left[-(1 - \Phi(y/\sigma))(S_r - 1)\right], \end{aligned} \tag{S11}$$

11

provided $(1 - \Phi(y/\sigma))$ is small (i.e., for sufficiently large y/σ). It is not difficult to see that for large S_r the function $P_Y(y)$ makes a swift transition from 0 to 1 at some values $y > 0$. To estimate the location of this transition we seek the inflection point y_0 given by $P_Y''(y_0) = 0$ (primes here and below denote derivatives). This yields the condition

$$(S_r - 1) \exp\left[-\frac{y_0^2}{2\sigma^2}\right] = \frac{\sqrt{2\pi} y_0}{\sigma} \tag{S12}$$

for y_0. If S_r is sufficiently large, the approximation that $\ln(y_0/\sigma)$ is small and can be ignored leads to an estimate

$$y_0 \approx \sigma\sqrt{2\ln S_r}. \tag{S13}$$

y_0 increases as the square root of the logarithm of S_r. For sufficiently large S_r, estimate (S13) justifies the assumption made for equation (S11) that $(1 - \Phi(y/\sigma))$ is small. In the following, we shall, however, work directly with the more precise equation (S12). A straightforward calculation shows that equation (S12) implies that the width of the region where $P_Y(y)$ goes from 0 to 1, estimated as $|P_Y'(y_0)/P_Y'''(y_0)|^{1/2} = \sigma^2/\sqrt{\sigma^2 + y_0^2}$, becomes narrower as S_r, and hence y_0 increases. This implies that for large S_r the mean of Y is approximately y_0, and the width of its distribution is considerably narrower than the width σ of the distribution of $\ln m_{ij}$. One can therefore approximate

$$X := \ln r_{ij} = \ln m_{ij} - Y \approx \ln m_{ij} - y_0. \tag{S14}$$

Thus,

$$P_X(x) := P[X < x] \approx P_M\left(\exp(x + y_0)\right) = \Phi((x + y_0)/\sigma). \tag{S15}$$

Now note that, since $Z_c(f)$ is the expected number of diet fractions larger than f (Rossberg *et al.*, 2006), the probability that one of the S_r diet fractions f_{ij}, chosen at random, is larger than f is $Z_c(f)/S_r$. In other words, the DPF equals S_r times the complementary cumulative distribution of f_{ij}. With $x := \ln r = \ln[f/(1-f)]$,

$$Z_c = S_r\left[1 - P_X(x)\right] = \frac{S_r}{2}\left[1 - \mathrm{erf}\left(\frac{x + y_0}{\sqrt{2\sigma^2}}\right)\right]. \tag{S16}$$

Using the well-known approximation $1 - \mathrm{erf}(u) \approx \pi^{-1/2} u^{-1} e^{-u^2}$ for large u, one obtains

$$Z_c \approx \frac{\sigma S_r}{\sqrt{2\pi}\,(x + y_0)} \exp\left[-\frac{(x + y_0)^2}{2\sigma^2}\right]. \tag{S17}$$

The order of magnitude of the empirically accessible range of diet fraction is limited (fractions smaller than 1% are typically not fully resolved). For large S_r ($\approx S_r - 1$), one may therefore assume $|x| \ll y_0$, which leads, with help of equation (S12), to

$$Z_c \approx \exp\left[-\frac{x y_0}{\sigma^2}\right]. \tag{S18}$$

12

Substituting $x = \ln r$, this becomes the power law

$$Z_c \approx r^{-\nu} \tag{S19}$$

with $\nu = y_0/\sigma^2$ or, employing approximation (S13), $\nu \approx \sigma^{-1}\sqrt{2 \ln S_r}$. This result is verified numerically in figure S1. To satisfy normalization condition (3.1), a factor $(\sin \pi\nu)/\pi\nu$ of order of magnitude one needs to be inserted into equation (S19).

B   Numerical Computation of the DPF for independent, log-normally distributed intakes

An accurate numerical computation of DPF as $Z_c(f) = S_r \left(1 - P[r_{ij} < f/(1 - f)]\right)$ (see equation (S16)) is easy once the cumulative distribution of the random variable r_{ij} defined by equation (S9) is known. We recall that the distribution of the sum of independent random variables is the convolution of their distributions. The convolution of the $S_r - 1$ independent log-normal distributions necessary to obtain the distribution of the denominator in equation (S9) is readily computed using a method described by Rossberg (2008). The distribution of the full quotient can then be obtained by a convolution of two distributions on logarithmic scales, as becomes clear from rewriting $\log r_{ij} = \log m_{ij} - \log \sum_{k \neq i} m_{kj}$.

C   Sampling Theory

The following primitive sampling theory links the DPF-approach to empirical studies using stomach sampling and a presence/absence model of trophic links.

Assume that prey size equals stomach size and that each non-empty sampled stomach contains exactly one species. Assume further that exactly n non-empty stomachs are sampled from each consumer species considered, and that all stomach samples are statistically independent.

The probability of finding species i in a non-empty stomach of species j is then f_{ij}, the prey species' diet fraction. The probability of not finding it there is $1 - f_{ij}$, and the probability of finding it in none of the n stomachs sampled is $[1 - f_{ij}]^n$. If f_{ij} is of the order of magnitude of 1 (i.e., i contributes a sizable proportion to the diet of j), this probability will be nearly zero for reasonably large value of n. In the opposite case (small f_{ij}), the probability can be approximated as $[\exp(-f_{ij})]^n = \exp(-nf_{ij})$, which is also near zero for f_{ij} of order one. The probability of actually finding i among n non-empty stochmachs of j is therefore $1 - \exp(-n f_{ij})$.

DPF theory implies that, if $Z_c(f)$ is the DPF of a community, the probability that there is a species contributing a diet fraction lying in the interval $[f; f + \Delta f]$ to the diet of a randomly chosen consumer from this community is $-Z_c'(f)\Delta f$, provided Δf is small (Rossberg $et\ al.$, 2006). ($Z_c'(f)$ denotes $dZ_c(f)/df$). The probability that a random consumer has a resource species with a diet fraction of this size and this species is found among n non-empty stomach samples is then $-Z_c'(f)[1 - \exp(-nf)]\Delta f$. The expected number of different resource species to be found among n stomachs of a random consumer is the sum of these probabilities over all diet-fraction sizes

$$Z_{\text{sample}} = -\int_0^1 Z_c'(f)\left[1 - \exp(-nf)\right] df. \tag{S20}$$

13

This is the expected empirical consumer link density "without threshold".

Integration by parts yields

$$Z_{\text{sample}} = -Z_c(f)\left[1 - \exp(-nf)\right]\Big|_{f=0}^{1} + \int_0^1 Z_c(f)n\exp(-nf)df. \tag{S21}$$

The $f = 0$ boundary term always vanishes and the $f = 1$ term vanishes if $Z_c(1) = 0$, that is, if no consumer feeds on a single resource alone, which is often a plausible assumption, e.g., in the case of fish. With the change of variables $f = \exp(u)$, the remaining integral can be re-written as

$$Z_{\text{sample}} = \int_{-\infty}^{0} Z_c(e^u)n\exp(-ne^u)e^u du. \tag{S22}$$

This integral can be evaluated approximately by noting that $n\exp(-ne^u)e^u$ is close to zero along most of the u axis, but attains a maximum at $u = -\ln n$. At the maximum, $n\exp(-ne^u)e^u = e^{-1}$ is of order one. If $Z_c(e^u)$ does not change much over the region where $n\exp(-ne^u)e^u$ is of order one, one can therefore approximate $Z_c(e^u)$ in Eq. (S22) by the constant $Z_c(e^{-\ln n}) = Z_c(1/n)$ and approximate sample consumer link density as

$$\begin{aligned}
Z_{\text{sample}} &\approx Z_c(1/n)\int_{-\infty}^{0} n\exp(-ne^u)e^u du \\
&\approx Z_c(1/n),
\end{aligned} \tag{S23}$$

where the last step again assumes that the number n of samples per consumer species is reasonably large (so that $\exp(-n) \approx 0$). Thus, sampling n non-empty stomachs per consumer approximately corresponds to sampling with a diet-fraction threshold of $f = 1/n$. With $r \approx f$, this implies $Z_{\text{sample}} \propto n^\nu$ for power law DPF with exponent ν.

The range in u over which $n\exp(-ne^u)e^u$ is of order one is itself of order one. Power-law DPF $Z_c(e^u)$ can actually change considerably over this range, which is why equation (S23) is just a crude approximation.

References

Baird, D. & Ulanowicz, R. E. 1989 The seasonal dynamics of the Chesapeake Bay ecosystem. *Ecol. Monog.*, **59**, 329–364.

Bersier, L.-F., Banašek-Richter, C. & Cattin, M.-F. 2002 Quantitative descriptions of food-web matrices. *Ecology*, **83**(9), 2394–2407.

Connolly, S. R., Hughes, T. P., Bellwood, D. R. & Karlson, R. H. 2005 Community structure of corals and reef fishes at multiple scales. *Science*, **309**, 1363–1365.

Cyr, H. 2000 The allometry of population density and inter-annual variability. In *Scaling in biology* (eds J. H. Brown & G. B. West), pp. 267–295. Oxford, UK: Oxford University Press.

de Figueiredo, G. M., Nash, R. D. & Montagnes, D. J. 2007 Do protozoa contribute significantly to the diet of larval fish in the Irish Sea? *J. Mar. Biol. Ass. U.K.*, **87**, 843–850.

Dunne, J. A. 2005 The network structure of food webs. In *Ecological networks: Linking structure to dynamics in food webs* (ed. J. A. D. Mercedes Pascual), Santa Fe Institute Studies in the Sciences of Complexity Proceedings, chap. 2, pp. 27–86. USA: Oxford University Press.

Goldwasser, L. & Roughgarden, J. 1993 Construction and analysis of a large Caribbean food web. *Ecology*, **74**, 1216–1233.

Greenwood, J. J. D. & Elton, R. A. 1979 Analysing experiments on frequency-dependent selection by predators. *J. Anim. Ecol.*, **48**, 721–737.

Hill, R. C., Griffiths, W. E. & Lim, G. C. 2008 *Principles of econometrics*. New York: Wiley, 3rd edn.

Horwich, P. 1998 *Truth*. Oxford University Press, 2nd edn.

Kartasche, B., Guill, C. & Drossel, B. 2009 Positive complexity-stability relations in food web models without foraging adaptation. *J. Theor. Biol.*, **259**(1), 12–23.

Levins, R. 1968 *Evolution in changing environments*. Princeton, NJ: Princeton University Press.

Link, J. 2002 Does food web theory work for marine ecosystems? *Marine Ecology Progress Series*, **230**, 1–9.

Link, J. S. & Almeida, F. P. 2000 An overview and history of the food web dynamics program of the Northeast Fisheries Science Center, Woods Hole, Massachusetts. NOAA Tech Memo NMFS-NE-159, National Oceanic and Atmospheric Administration.

Martinez, N. D. 1992 Constant connectance in community food webs. *Am. Nat.*, **139**, 1208–1218.

McCann, K., Hastings, A. & Huxel, G. R. 1998 Weak trophic interactions and the balance of nature. *Nature*, **395**(794–798).

Mora, C., Derek P, T. & Myers, R. A. 2008 The completeness of taxonomic inventories for describing the global diversity and distribution of marine fishes. *Proc. R. Soc. B*, **275**, 149–155.

Petchey, O. L., Downing, A. L., Mittelbach, G. G., Persson, L., Steiner, C. F., Warren, P. H. & Woodward, G. 2004 Species loss and the structure and functioning of multitrophic aquatic systems. *OIKOS*, **104**, 467–478.

Pinnegar, J. K., Trenkel, V. M., Tidd, A. N., Dawson, W. A. & Dubuit, M. H. 2003 Does diet in Celtic Sea fishes reflect prey availability? *J. Fish Biol.*, **63 (Supplement A)**, 197–212.

Quince, C., Higgs, P. G. & McKane, A. J. 2005 Topological structure and interaction strengths in model food webs. *Ecological Modelling*, **187**, 389–412. Doi:10.1016/j.ecolmodel.2004.12.018.

Rossberg, A. G. 2008 Laplace transforms of probability distributions and their inversions are easy on logarithmic scales. *J. Appl. Prob.*, **45**(2).

Rossberg, A. G., Brännström, Å. & Dieckmann, U. 2010 How trophic interaction strength depends on traits — A conceptual framework for representing multidimensional trophic niche spaces. *Theor. Ecol.*, **3**(1), 13–24. (doi:10.1007/s12080-009-0049-1)

Rossberg, A. G., Yanagi, K., Amemiya, T. & Itoh, K. 2006 Estimating trophic link density from quantitative but incomplete diet data. *J. Theor. Biol.*, **243**(2), 261–272.

Rountree, R. A. 2001 Diets of NW Atlantic fishes and squid. `http://www.fishecology.org/diets/summary.htm`. Accessed Apr 30, 2005.

Schmid-Araya, J. M., Hildrew, A. G., Robertson, A., Schmid, P. E. & Winterbottom, J. 2002*a* The importance of meiofauna in food webs: Evidence from an acid stream. *Ecology*, **83**(5), 1271–1285.

Schmid-Araya, J. M., Schmid, P. E., Robertson, A., Winterbottom, J., Gjerløev, C. & Hildrew, A. G. 2002*b* Connectance in stream food webs. *J. Anim. Ecol.*, **71**, 1056–1062.

Winemiller, K. O. 1990 Spatial and temporal variation in tropical fish trophic networks. *Ecol. Monog.*, **60**, 331–367.