nature ecology & evolution

SUPPLEMENTARY INFORMATION
https://doi.org/10.1038/s41559-018-0517-3

The spatial scaling of species interaction networks

Nuria Galiana[1], Miguel Lurgi[1], Bernat Claramunt-López[2,3], Marie-Josée Fortin[4], Shawn Leroux[5], Kevin Cazelles[6,7,8], Dominique Gravel[9] and José M. Montoya[1]*

[1]Ecological Networks and Global Change Group, Theoretical and Experimental Ecology Station, CNRS and Paul Sabatier University, Moulis, France. [2]CREAF, Edifici Ciències, UAB, Bellaterra, Catalunya, Spain. [3]Ecology Department, BABVE, Bellaterra, Catalunya, Spain. [4]Department of Ecology & Evolutionary Biology, University of Toronto, Toronto, Ontario, Canada. [5]Department of Biology, Memorial University, St. John's, Newfoundland and Labrador, Canada. [6]Département de Biologie, Chimie et Géographie, Université du Québec à Rimouski, Quebec, Canada. [7]Quebec Center for Biodiversity Science, Montreal, Quebec, Canada. [8]UMR MARBEC, Université de Montpellier, Montpellier, France. [9]Département de Biologie, Université de Sherbrooke, Sherbrooke, Quebec, Canada. *e-mail: josemaria.montoyateran@sete.cnrs.fr

SUPPLEMENTARY MATERIAL OF THE SPATIAL SCALING OF SPECIES INTERACTION NETWORKS

Núria Galiana[1], Miguel Lurgi[1], Bernat Claramunt[2,3], Marie-Josée Fortin[4], Shawn Leroux[5], Kevin Cazelles[6,7,8], Dominique Gravel[9], José M. Montoya[1]

Index

- **Supplementary Methods:**
    - **Regional food webs (metawebs)** (Page 2)
    - **Local food web assembly models** (Page 2)
    - **Network properties** (Page 6)
    - **Beta-diversity metrics (**Page 7)
    - **Statistical analyses** (Page 8)
- **Supplementary Table 1:** Description of the metawebs (Page 10)
- **Supplementary Table 2:** Statistical results for Species-Area Relationships (SARs) across trophic levels (Page 11)
- **Supplementary Table 3:** Fit of each network property of each model to a bounded exponential function (Page 12)
- **Supplementary Figure 1:** Network-Area Relationships (NARs) for the *Trophic Sampling* model (Page 13)
- **Supplementary Figure 2:** Decomposition of the intermediate trophic level (Page 14)
- **Supplementary Figure 3:** NARs for the *Trophic Theory of Island Biogeography* model (Page 15)
- **Supplementary Figure 4:** Dispersal - beta diversity relationship for the *Trophic meta-community* model (Page 16)
- **Supplementary Figure 5:** NARs for the *Trophic meta-community* model (Page 17)
- **Supplementary Figure 6:** Comparison of NARs across models (Page 18)
- **Supplementary Figure 7:** Comparison of NARs with and without trophic constraint for each model (Page 19)
- **Supplementary Figure 8:** Comparison of NARs with different connectances in the metaweb (Page 20)

Supplementary Methods

We explore the effect of each process on the spatial scaling of food web structure with three models of multi-trophic community assembly. They all consider the same regional food webs (or metawebs) from which species are selected to conform local communities, and the local assembly processes of the three models follows the same trophic constraint, i.e. consumers need to have at least one resource in the local community. In addition, we have a non-trophic constrained version of each of the models presented below. Supplementary Figure 7 provides the comparison between the two versions, which shows that most of the differences occur at small areas where the effect of the trophic constraint is stronger. We then analysed several network properties at different spatial scales, which allowed us to characterize a suite of NARs. In this section, we first explain how we generated the regional food webs, we then present each community assembly model, and we finally explain the network properties analysed, the beta-diversity metrics used, and the statistical analyses conducted.

Regional food webs (Metawebs)

We used the niche model (Williams & Martinez 2000) to generate food webs that were used as the regional pool of species for the three models. The niche model requires two input parameters: the number of species (S), and network connectance (*C*) defined as the fraction of realised links among the possible links (i.e. $C=L/S^2$). This model describes trophic niche occupancy between consumers and resources along a resource axis. It generates model food webs that approximate well the central tendencies and the variability of a number of empirical food web properties (Williams & Martinez 2000; Dunne et al. 2002; Stouffer et al. 2005).

We performed one hundred replicates of each of the three models using that same number of distinct regional food webs generated using the niche model with $S = 200$ and $C = 0.056$ as input parameters. The stochastic nature of the niche model ensures variability across the replicate networks. The value for *C* was assigned following the power decay of *C* as *S* increases found in empirical networks ($C = \sigma S^{-\lambda}$, where $\sigma = 0.8$ and $\lambda = 0.5$; Cohen & Briand 1984; Montoya & Solé 2003; Ings et al. 2009). The specific values of S and C, however, did not affect qualitatively our results (See Supplementary Figure 8). For a full description of the regional network (or metaweb) properties see Supplementary Table 1.

Local food web assembly models

a) Trophic Sampling model

This model randomly samples networks of different sizes from the 200-species regional food web (i.e., the metaweb) described above. This is not a spatial model, and as such there are no occupancy dynamics (i.e., there are no spatial processes involved in the local community assembly process) (Box 1). The assembly process was exclusively constrained by the requirement that each consumer species needed to have at least one prey in the local community. Thus, for each area we randomly selected samples from the regional food web with the number of species required and we checked that all selected consumers have at least one prey. If the trophic requirement was not satisfied, the subsample selected was excluded from the analyses. To obtain the number of species for a given area, we used $S = kA^z$, with k = 10 and z = 0.27. Therefore, to sample different number of species we generated a range of area sizes that covers the entire spectrum of species available in the pool. The value of z used (z = 0.27) was derived from a meta-analysis performed by Drakare *et al.* (2006), which is widely regarded as the best empirical estimate for terrestrial systems. The value of k corresponds to the smallest local community. We used the power function since it is one of the best known and most common reported relationships in the literature (Arrhenius 1921; Tjørve 2003; Scheiner 2003). The power law has been shown to be the central phase of a triphasic curve: SAR is concave at local scales, approximately linear at regional scales, and finally convex at continental scales (Rosindell & Cornell 2007). Therefore, the mechanisms, expectations, and results presented in our theory would reflect local-to-regional scale dynamics.

The *Trophic Sampling model* can be considered as a baseline reference for NARs, where the increase of species richness with area is constrained by the simplest food web assembly process: the occurrence of a predator is mediated by the presence of one of its prey. Even though the processes tested with this model (i.e., link-area scaling and SARs dependency on trophic level) are not exclusive of this assembly process, we are interested in the patterns emerging from the model with the simplest assumptions.

b) Trophic Theory of Island Biogeography model

The trophic theory of island biogeography (*TTIB*; Gravel et al. 2011) predicts the structure of food webs on islands connected through immigration to a mainland, which in turn contains a large regional species pool. The structure of the local assembly of species and their interactions is defined by the equilibrium between colonisation and extinction dynamics. It considers the additional constraint of consumers requiring at least one prey on the island to colonize and persist. The ratio between colonisation and extinction (for a fixed colonisation rate) is considered a proxy of geographical area; the smaller the ratio the smaller the area (MacArthur & Wilson 1967; Gravel et al. 2011).

In this model, we assumed species from the regional metaweb could colonise the island with a fixed colonisation probability (*c*=0.2) and could go extinct with an increasing extinction probability (*e* ={0.01, 0.05, 0.1, 0.2, 0.3, 0.4, 0.5}). These parameter values allowed the exploration of a range of local communities that correspond to islands of different area (i.e., area size in this model is given by the *c/e* ratio). For each combination of parameters (extinction, colonisation) and for each of our regional networks, we ran simulations for a total of 1000 time steps, which is enough to reach the equilibrium between colonisation and extinction dynamics and, therefore, to determine the number of species for a given island (see Gravel et al. 2011 for details). This is a discrete time one-patch species occupancy model, in which the occupancies of all species are determined by the trade-off between *c* and *e* plus prey availability. To calculate the final species configuration for a given combination of *c* and *e*, in each time step of the simulation we sequentially went over each species in the network and stochastically determine its extinction or colonisation based on the corresponding rates and whether there are resources present for that species. Secondary extinctions were calculated after primary extinctions occurred. We checked for all species that lost their resources and repeated this check until no further extinctions were found. Network properties were computed for resulting networks at the end of the simulation (i.e., time step 1000).

The difference between this *TTIB* model and the *Trophic Sampling* model lays on the complexity of the assembly process. The stochastic nature of the community assembly enforced by the *TTIB* model, which is the result of colonisation-extinction dynamics, favours the persistence over time of generalist consumers. Specialist consumers that depend on a single resource are more prone to become secondarily extinct given the trophic constraint: if their only resource goes stochastically extinct, they go extinct too. In contrast, the *Trophic Sampling* model only searches for possible configurations of a given number of species where every consumer needs to have a resource, without subjecting the selected community to additional stochastic extinctions.

c) *Trophic Meta-community model*

We used a multi-trophic meta-community model (Box 1) to determine the effects of spatial variability in species distribution (i.e., beta-diversity) on NARs. We used different levels of dispersal among local communities to generate beta-diversity in space. Seventy-five local communities (patches) were randomly placed on a 2-dimensional Euclidean space of 1 unit of length on each dimension and the distances between patches were calculated. Patch pairs were considered to be connected if the distance between them was less than or equal to 0.3, because it provided a good balance between local species richness and spatial heterogeneity in species composition. Connections between patches were the same for all species and were bidirectional and homogeneous, i.e., dispersal probabilities were equal across every pair of connected patches, and there was no dispersal between non-spatially connected patches. A different (randomly generated) network of patches was used for each replicate.

We used the same regional networks as in the previous models. At each time step, species could arrive to each patch from the regional pool independently of each other given a fixed colonisation probability (c = 0.1) and could go locally extinct with a fixed probability (e = 0.4). We chose these values for c and e because they allowed for a local community size that yielded a good trade-off between species richness at the local and regional scales. Additionally, at each time step all species present in a given patch could disperse to a connected patch according to a given dispersal rate. We ran two versions of this model using two different values for the dispersal rate d={0.01, 0.1}. This allowed us to test the effect of different levels of dispersal on the spatial scaling of network structure. Within each version of the model, dispersal rates were constant across all species. As for the *TTIB* model, stochastic colonisation-extinction dynamics were dependent on the trophic constraint (i.e., on the availability of resources for each consumer present). Similarly, dispersal followed the same trophic constraint. That is, consumers could only disperse to those connected patches where there was at least one of their resources. Simulations were run for 1000 time steps to achieve an equilibrium state in terms of species composition. Similarly to the *TTIB*, this is a discrete time patch occupancy model, only this time it is multi-patch. At each time step, extinction, colonisation, and dispersal dynamics are calculated sequentially for every patch and species. Occupancy time series were thus created by calculating extinction and colonisation independently at each local patch and then calculating the dispersal probabilities across neighbouring patches. Network properties were again analysed (as for the *TTIB*) over the food web resulting at the end of the simulation.

The spatial scale was determined by the spatial aggregation of several local communities in an ever-increasing way from 1 (a single local community) to 75 (maximum number of communities; the regional scale). That is, the largest spatial scale was the complete aggregation of all local communities, which yielded the total set of species of the regional species pool if every species persisted at the regional scale after the simulation. The aggregation procedure was based on neighbouring aggregation, whereby local communities were aggregated based on the distance to their neighbours. One local community was selected randomly as the starting point. This process was repeated 75 times selecting every local community as a different starting point. It is important to notice that in this model communities were assembled at the local scale (i.e., in each local patch) while the spatial scale of sampling increased with the aggregation of different local patches. Therefore, the spatial scale of sampling differs from the scale at which the assembly takes place. In contrast, in the two previous models the scale of sampling is the scale at which the assembly process occurs. This difference is key to understand the components of the spatial scale that each model captures. While the *Trophic Sampling* and the TTIB consider 'complete' areas (i.e., area = island area), the *meta-community* model considers a sub-area of a larger region. Therefore, while the first two models capture the effect of alpha diversity, the latter also captures the effect of beta-diversity.

After the generation of different food webs for each of the three models, the corresponding area was rescaled to fall in the range between 0 and 1, where 0 is the smallest local scale and 1 is the largest regional scale, to allow model comparison.

Network properties

We classified NARs into three categories: vertical diversity, complexity, and modules.

Vertical diversity properties included the fraction of omnivory (fraction of species feeding from more than one trophic level; *O*), mean food chain length (mean number of links needed to get from the basal species to the top species, restricted to a bottom-top direction, with paths from more than 1 basal species being averaged; *MFCL*)*,* fraction of basal species (species without prey; *B*), fraction of intermediate species (species with predators and prey; *I*) and fraction of top species (species with predators; *T*). Notice that the fraction of intermediate species includes both primary and secondary consumers. For consistency with the literature on food web research, we keep the analysis on the fraction of intermediate species as a whole. However, we also analyse individually primary consumers, secondary consumers and omnivorous species that form the intermediate trophic level.

Network complexity properties included the number of species (*S*), number of links (*L*), connectance ($C=L/S^2$), links per species (*L/S*), mean indegree (average number of prey items per predator; also known as mean generality) and mean outdegree (average number of predators per prey species; also known as mean vulnerability).

Module properties included network modularity (*M*), which indicates the presence of densely linked groups within a network (Newman & Girvan 2003), and three-species trophic motifs as building blocks of complex networks (Milo et al. 2002; Bascompte & Melian 2005; Camacho et al. 2007). The trophic motifs have been used as an indication of the modules that are structurally over- or under-represented in food webs. We calculated the proportion of (i) tri-trophic food chains, (ii) apparent competition, and (iii) exploitative competition motifs across spatial scales.

Additionally, to assess the extent to which more generalist consumers are better colonisers in smaller areas, we calculated the distribution of the diet breadth of those consumers present in a given patch. To do so, we considered their potential diet breadth (i.e., indegree or generality), defined as the number of resource species they consume within the regional metaweb. We call this the *potential indegree distribution*. Additionally, we calculated the *realized indegree distribution* within each island. This property reflects the consequences of the preferential selection of generalist consumers on the structure of the realized local network. For the realized indegree distributions, we calculated the cumulative probability at each spatial scale Pc(k), for $\geq k$, where Pc(k) is the probability a species has k or less prey in the network.

Beta-diversity metrics

In order to assess the level of dissimilarity on species composition among local communities (i.e., species *beta-diversity*) depending on the level of dispersal, we used the metrics developed by Baselga et al., (2017), available in 'betapart' package in R.

For accounting for the total beta-diversity we used the multiple-site Sorensen index:

$$\beta_{SOR} = \frac{\left[\sum_{i<j} \min(b_{ij}, b_{ji})\right] + \left[\sum_{i<j} \max(b_{ij}, b_{ji})\right]}{2\left[\sum_i S_i - S_T\right] + \left[\sum_{i<j} \min(b_{ij}, b_{ji})\right] + \left[\sum_{i<j} \max(b_{ij}, b_{ji})\right]}$$

where S_i is the total number of species in site i, S_T is the total number of species in all sites considered together and b_{ij}, b_{ji} are the number of species exclusive to sites i and j, respectively, when compared by pairs.

Additionally, we calculated the turnover of interactions computing the number of interactions in common between two local communities instead of the number of species. In our models, interactions turnover is determined by the absence of one or the two interacting species (i.e. species beta-diversity) because interactions are determined by the metaweb. That is, if two species co-occur locally and they have a link in the metaweb they will always interact locally. No differences were observed when comparing the patterns for species turnover and interactions turnover.

Statistical analyses

Mean and standard error of the mean among the 100 replicates of each of the simulation experiments were calculated for each food web property and plotted using generalised additive models (GAM) with the R package 'ggplot2' (Wickham 2009). GLMs were used to analyse differences in SARs depending on species trophic level and to analyse the variation of the number of links with species richness for the *Trophic Sampling* model.

We fitted bounded exponential curves ($f(x) = L(1 - he^{-gx})$; where *L* represents the curve's maximum value, *g* the steepness of the curve and *h* determines the fraction of *L* where the curve starts) using nonlinear least squares (NLS) with the 'nls' function in R. Growth exponents were used to quantify the rate of growth of network properties with area.

References

Arrhenius, O. (1921). Species and area. *J. Ecol.*, 9(1), 95-99.

Bascompte, J., & Melián, C. J. (2005). Simple trophic modules for complex food webs. *Ecology*, 86(11), 2868-2873.

Baselga, A., Orme, D., Villeger, S., De Bortoli, J., & Leprieur, F. (2017). Betapart: Partitioning Beta Diversity into Turnover and Nestedness Components. R package version 1.4-1.

Camacho, J., Stouffer, D. B., & Amaral, L. A. N. (2007). Quantitative analysis of the local structure of food webs. *J. Theoret. Biol.*, 246(2), 260-268.

Cohen, J. E., & Briand, F. (1984). Trophic links of community food webs. *Proc. Natl. Acad. Sci.*, 81(13), 4105-4109.

Drakare, S., Lennon, J. J., & Hillebrand, H. (2006). The imprint of the geographical, evolutionary and ecological context on species–area relationships. *Ecol. Lett.*, 9(2), 215-227.

Dunne, J. A., Williams, R. J., & Martinez, N. D. (2002). Network structure and biodiversity loss in food webs: robustness increases with connectance. *Ecol. Lett.*, 5(4), 558-567.

Gravel, D., Massol, F., Canard, E., Mouillot, D., & Mouquet, N. (2011). Trophic theory of island biogeography. *Ecol. Lett.*, 14(10), 1010–6.

Ings, T. C., Montoya, J. M., Bascompte, J., Blüthgen, N., Brown, L., Dormann, C. F. et al. (2009). Review: Ecological networks–beyond food webs. *J. Anim. Ecol.*, 78(1), 253-269.

MacArthur, R.H. & Wilson, E.O. (1967). The Theory of Island Biogeography. Princeton University Press, Princeton, NJ.

Milo, R., Shen-Orr, S., Itzkovitz, S., Kashtan, N., Chklovskii, D., & Alon, U. (2002). Network motifs: simple building blocks of complex networks. *Science*, 298(5594), 824-827.

Montoya, J. M., & Solé, R. V. (2003). Topological properties of food webs: from real data to community assembly models. *Oikos*, 102(3), 614-622.

Newman, M. E., & Girvan, M. (2003). Mixing patterns and community structure in networks. In *Stat. Mech. Com. Net.* (pp. 66-87). Springer Berlin Heidelberg.

R Development Core Team. (2013). R: A Language and Environment for Statistical Computing. Vienna: R Foundation for Statistical Computing.

Rosindell, J., & Cornell, S. J. (2007). Species–area relationships from a spatially explicit neutral model in an infinite landscape. *Ecol. Lett*., *10*(7), 586-595.

Scheiner, S. M. (2003). Six types of species-area curves. *Glob. Ecol. Biogeogr*., 12(6), 441-447.

Stouffer, D. B., Camacho, J., Guimera, R., Ng, C. A., & Nunes Amaral, L. A. (2005). Quantitative patterns in the structure of model and empirical food webs. *Ecology*, 86(5), 1301-1311.

Tjørve, E. (2003). Shapes and functions of species–area curves: a review of possible models. *J. Biogeogr*., 30(6), 827-835.

Wickham, H. (2009). ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag, New York.

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

Supplementary Table 1. Description of the metawebs used in the theoretical models. A 100 different metawebs were used.

Property	Mean(Standard deviation)
Number of Species	200(0)
Connectance	0.056(0.004)
Number of links	2209.3(160.9)
Links per species	11.04(0.80)
Indegree	12.48(0.88)
Outdegree	11.63(0.99)
Mean food chain length	12.26(1.08)
Proportion of basal	0.12(0.02)
Proportion of intermediate	0.83(0.03)
Proportion of top	0.05(0.03)
Omnivory	0.86(0.05)
Modularity	0.37(0.02)

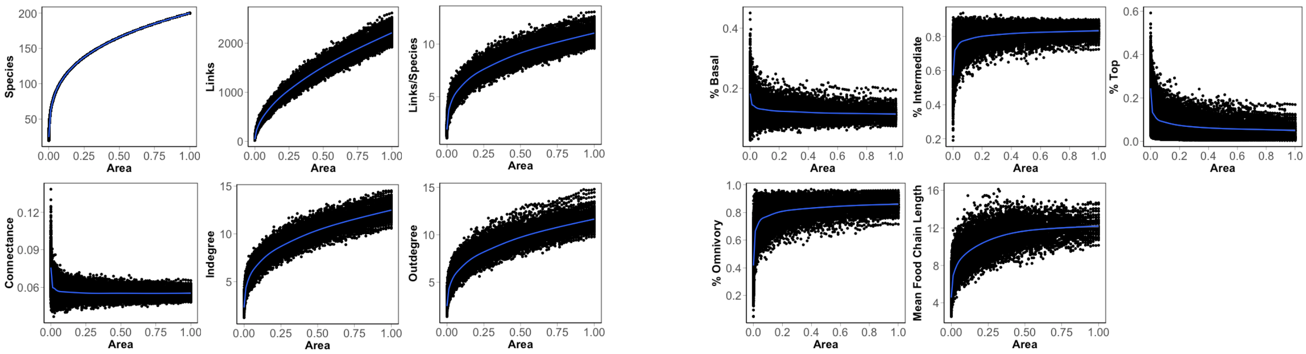
Supplementary Table 2. GLM results for the differences in SAR depending on trophic level. Estimates, T values and statistical significances are indicated (*** for p-values<0.001; ** for p-values<0.01; * for p-values<0.05) for the slope and intercepts of the relationships. Basal was set as reference level, that is, values for Intermediate and Top show the deviation from Basal values.

		Estimate	T value
Basal	Intercept	1.38	605.98 ***
	Slope	0.10	72.55 ***
Intermediate	Intercept	0.84	262.19 ***
	Slope	0.23	112.16 ***
Top	Intercept	-0.48	-149.55 ***
	Slope	-0.03	-17.07 ***

Supplementary Table 3. Fit of each network property of each model to a bounded exponential function $(f(x) = L(1 - he^{-gx})$, where *L* represents the curve's maximum value, *g* the steepness of the curve and *h* determines the fraction of *L* where the curve starts. We used nonlinear least squares (NLS) with the 'nls' function in R.

Network property	Model	t value			Estimate		
		L	*g*	*h*	*L*	*g*	*h*
Species	Trophic Sampling	918.9	219.2	1043.2	183.2	5.13	0.75
	TTIB	358.50	79.42	397.80	181.2	10.7	0.95
	Dispersal 0.01	1020.2	136.6	260.1	173.8	6.5	0.78
	Dispersal 0.1	2897.9	89.95	162.82	193.7	34.2	0.41
Links	Trophic Sampling	359.9	169.6	1821.3	2350	2.1	0.93
	TTIB	250.04	67.18	412.37	1925	5.8	1
	Dispersal 0.01	461.76	88.45	182.21	1842	5.2	0.95
	Dispersal 0.1	996.27	44.37	80.33	2108	34.1	0.6
Links/species	Trophic Sampling	602.3	143.6	702.4	10.19	4.82	0.72
	TTIB	240.44	49.63	240.17	10.06	12.3	0.86
	Dispersal 0.01	788.49	79.78	139.45	10.30	9.5	0.74
	Dispersal 0.1	1129.01	23.06	50.38	10.86	69.9	0.37
Connectance	Trophic Sampling	1002.65	27.72	-62.63	0.055	558.8	-0.30
	TTIB	63.577	8.524	-18.49	0.057	58.3	-0.62
	Dispersal 0.01	629.03	38.06	-68.66	0.060	7.9	-0.39
	Dispersal 0.1	987.55	15.08	-25.44	0.056	15.7	-0.15
Indegree	Trophic Sampling	616.1	130.8	609.8	11.57	5.05	0.66
	TTIB	245.31	47.92	232.22	11.43	11.85	0.81
	Dispersal 0.01	851.01	74.91	128.24	11.87	11.6	0.72
	Dispersal 0.1	1174.37	20.55	49.15	12.33	85.7	0.35
Outdegree	Trophic Sampling	554.1	126.7	609.6	10.75	4.89	0.70
	TTIB	215.38	40.95	97.23	10.70	12.03	0.79
	Dispersal 0.01	683.20	66.31	116.24	10.91	9.39	0.70
	Dispersal 0.1	968.08	18.74	40.71	11.45	68.9	0.34
MFCL	Trophic Sampling	570.71	89.41	346.62	11.37	10.15	0.57
	TTIB	170.33	31.17	136.32	11.57	18.91	0.72
	Dispersal 0.01	592.62	49.65	88.41	11.53	9.93	0.53
	Dispersal 0.1	657.603	9.904	25.29	12.01	82.16	0.23
% Basal	Trophic Sampling	272.61	68.35	-134.8	0.14	25.7	-1.35
	TTIB	41.75	10.93	-20.12	0.12	42.69	-1.19
	Dispersal 0.01	246.28	39.01	-73.75	0.13	5.87	-0.85
	Dispersal 0.1	407.72	11.20	-18.91	0.11	13.36	-0.24
% Intermediate	Trophic Sampling	1061.69	70.92	218.02	0.78	30.5	0.37
	TTIB	160.61	22.07	62.11	0.80	67.46	0.64
	Dispersal 0.01	1143.57	39.52	67.72	0.80	12.16	0.28
	Dispersal 0.1	1711.599	9.424	16.655	0.82	29.61	0.07
% Top	Trophic Sampling	121.40	30.38	-60.53	0.066	44.22	-1.56
	TTIB	16.90	12.62	-12.85	0.065	46.96	-3.4
	Dispersal 0.01	110.17	21.74	37.67	0.05	42.84	-2.94
	Dispersal 0.1	126.538	3.149	-9.380	0.05	122.9	-0.64
% Omnivory	Trophic Sampling	908.34	62.52	179.81	0.81	67.64	0.47
	TTIB	152.92	24.95	74.05	0.82	56.45	0.75
	Dispersal 0.01	1042.21	36.27	64.94	0.84	31.67	0.45
	Dispersal 0.1	1403.842	7.065	16.186	0.85	78.74	0.09

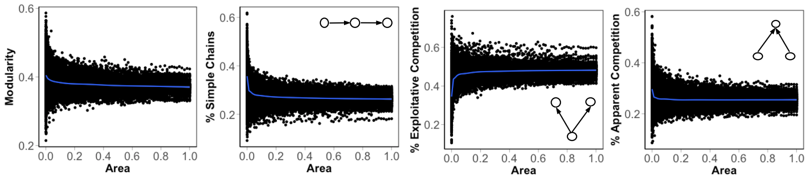
Supplementary Figure 1. Network-Area relationships (NAR) for the Trophic Sampling model. Area values close to 0 correspond to local communities and values close to 1 correspond to regional communities. Properties giving information about food web a) complexity, b) vertical diversity and c) motifs and modularity. Notice that the Species-Area relationship shown in (a) is given by $S = kA^z$, with k = 10 and z = 0.27. Simple representations of network motifs are shown in the corresponding panels; Nodes represent species and arrows trophic interactions from prey to predators.



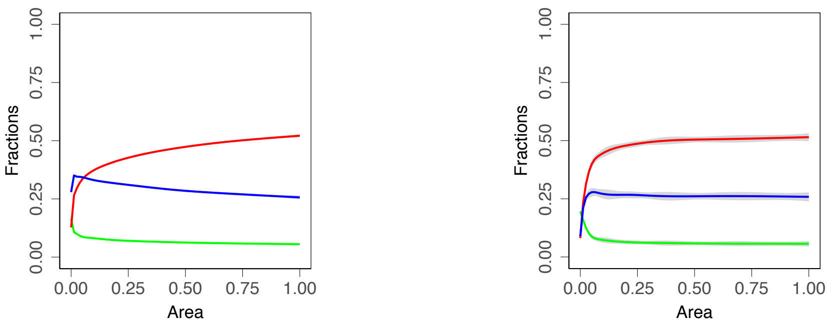
a) Complexity

b) Vertical diversity

c) Motifs and modularity



Supplementary Figure 2. Decomposition of the intermediate trophic level. Relationship of the fraction of species of each group with area for (a) the *Trophic Sampling* model, (b) the *Trophic Theory of Island Biogeography* model, (c) the *Trophic metacommunity* model with low dispersal (d=0.01) and (d) high dispersal (d=0.1). Green: herbivores species (i.e. species only consuming basal species); Blue: carnivores (i.e. species only preying on other consumer species); Red: omnivorous species (i.e. species consuming both basal and consumer species). It is important to notice that all these groups only refer to those species included in the intermediate trophic level, it does not consider top species.

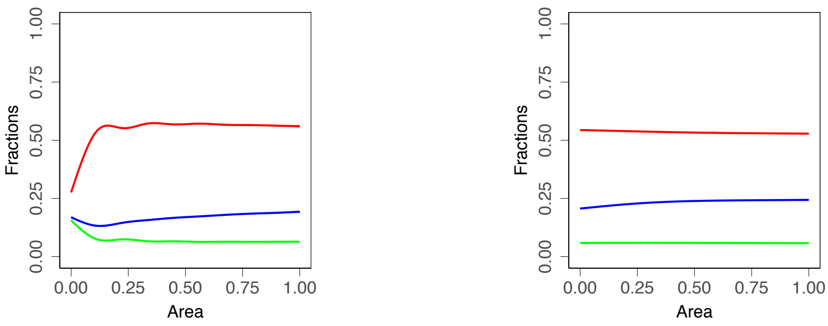


a) Trophic Sampling

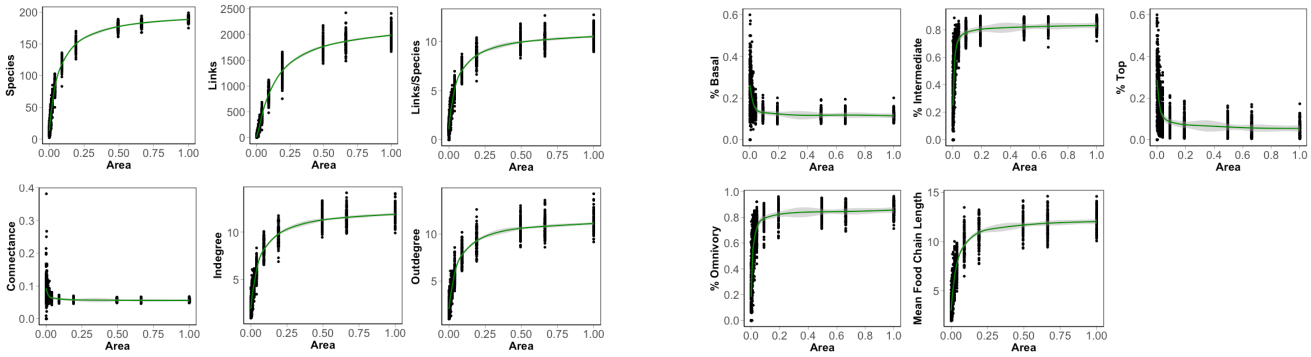
b) Trophic Theory of Island Biogeographgy

c) Trophic Metacommunity Low dispersal

d) Trophic Metacommunity high dispersal



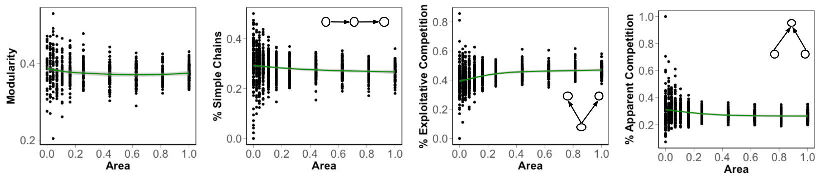
Supplementary Figure 3. Network-Area relationships (NAR) for the TTIB model. Area values close to 0 correspond to local communities and values close to 1 correspond to regional communities. Properties giving information about food web a) complexity, b) vertical diversity and c) motifs and modularity. Simple representations of network motifs are shown in the corresponding panels; Nodes represent species and arrows trophic interactions from prey to predators. Shaded areas show 95% confidence intervals.



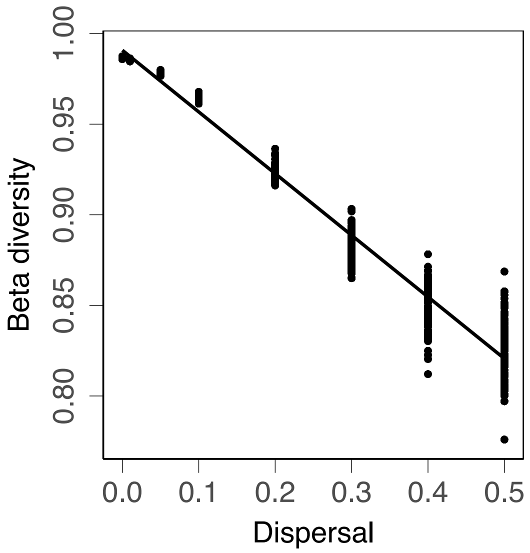
a) Complexity

b) Vertical diversity

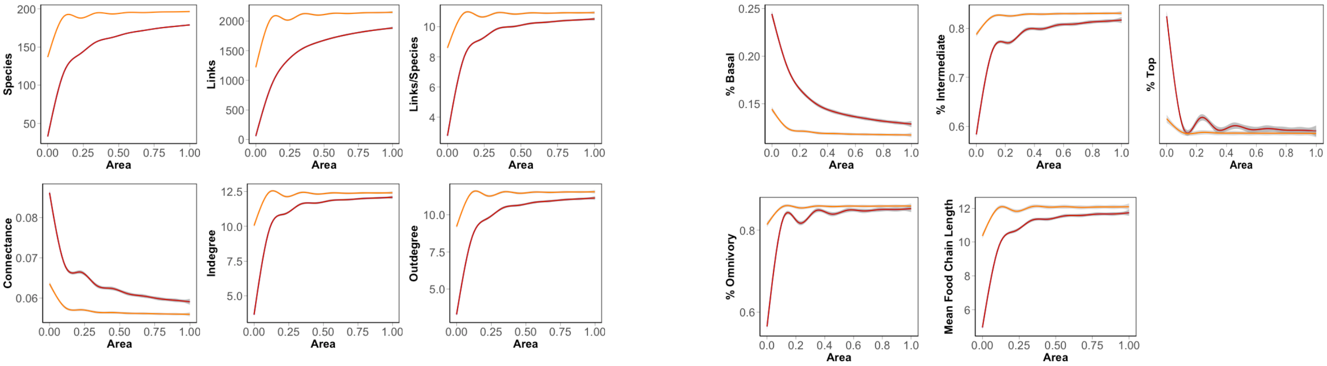
c) Motifs and modularity



Supplementary Figure 4. Dispersal - beta diversity relationship for the *Trophic meta-community* model. Increasing species dispersal rates increases the similarity between patches. Note that we analysed these patterns for eight different dispersal rates d ={0, 0.01, 0.05, 0.1, 0.2, 0.3, 0.4, 0.5}.



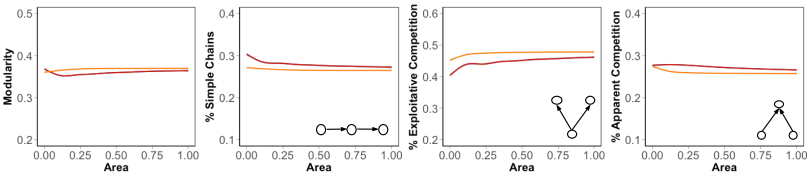
Supplementary Figure 5. Network-Area relationships (NAR) for the trophic meta-community model. Area values close to 0 correspond to local communities and values close to 1 correspond to regional communities. Properties giving information about food web a) complexity, b) vertical diversity and c) motifs and modularity. Red line: 0.01 dispersal rate; Orange line: 0.1 dispersal rate. Simple representations of network motifs are shown in the corresponding panels; Nodes represent species and arrows trophic interactions from prey to predators. Shaded areas show 95% confidence intervals.



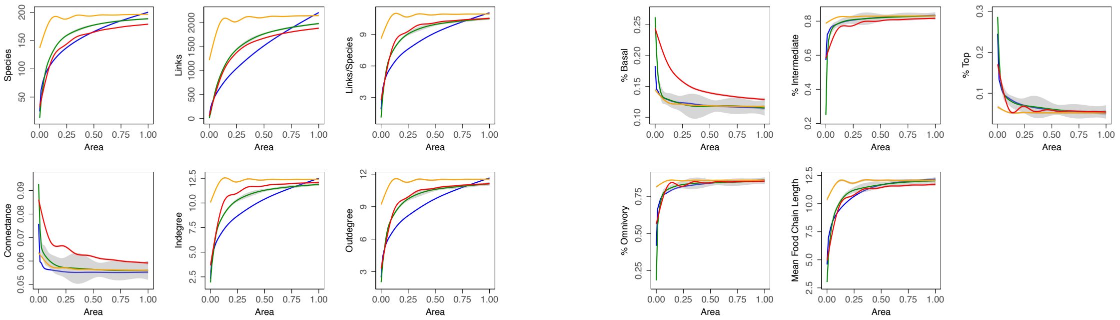
a) Complexity

b) Vertical diversity

c) Motifs and modularity



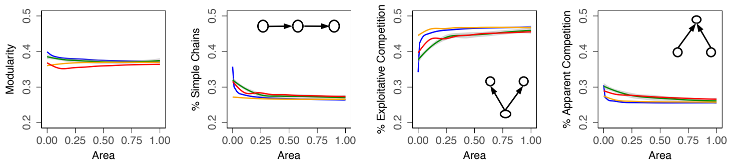
Supplementary Figure 6. Comparison of Network-Area relationships (NAR) across models. Area values close to 0 correspond to local communities and values close to 1 correspond to regional communities. Properties giving information about food web a) complexity, b) vertical diversity and c) motifs and modularity. Blue: *Trophic Sampling* model; Green: *Trophic Theory of Island Biogeography* model; Red: *Trophic meta-community* model with 0.01 dispersal rate; Orange: *Trophic meta-community* model with 0.1 dispersal rate. Simple representations of network motifs are shown in the corresponding panels; Nodes represent species and arrows trophic interactions from prey to predators. Shaded areas show 95% confidence intervals.



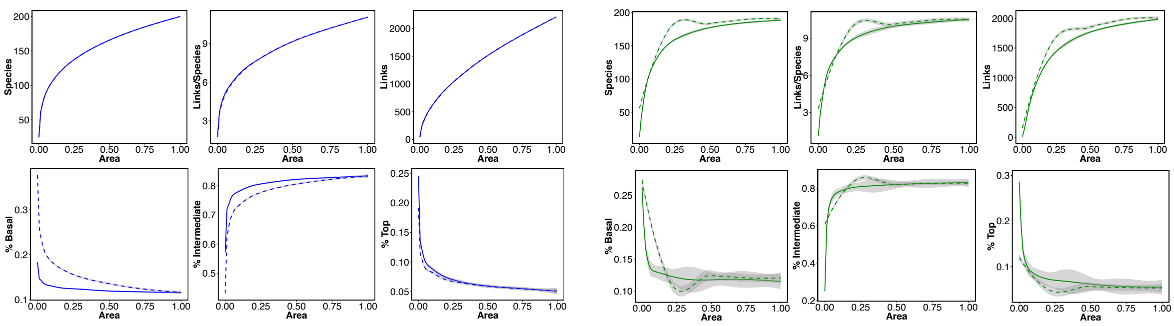
a) Complexity

b) Vertical diversity

c) Motifs and modularity



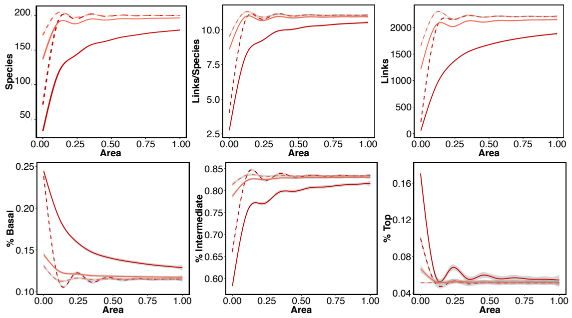
Supplementary Figure 7. Comparison of Network-Area relationships (NAR) for the models with trophic constraint (solid line) and without the trophic constraint (dashed line). Area values close to 0 correspond to local communities and values close to 1 correspond to regional communities. a) NARs for the *Trophic Sampling* model, b) NARs for the *Trophic Theory of Island Biogeography* model and c) NARs for the *Trophic Metacommunity* model (Red line: 0.01 dispersal rate; Orange line: 0.1 dispersal rate). Notice that the non-constraint TTIB in (b) corresponds to the classic Theory of Island Biogeography model. Notice that the Species-Area relationship for the *Trophic Sampling* model shown in (a) is given by $S = kA^z$, with *k* = 10 and *z* = 0.27. Shaded areas show 95% confidence intervals.



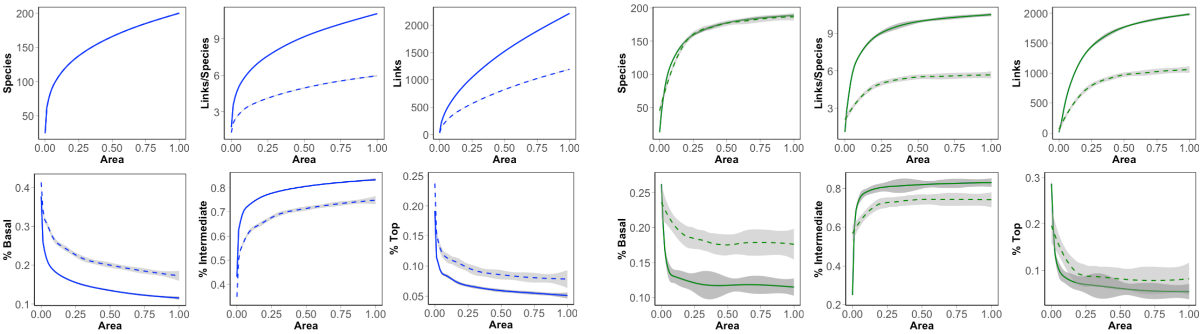
a) Trophic Sampling

b) Trophic Theory of Island Biogeography

c) Trophic Metacommunity



Supplementary Figure 8. Comparison of Network-Area relationships (NAR) for the models considering metawebs with different connectances; C= 0.056 (solid line) and C=0.03(dashed line). Area values close to 0 correspond to local communities and values close to 1 correspond to regional communities. a) NARs for the *Trophic Sampling* model, b) NARs for the *Trophic Theory of Island Biogeography* model and c) NARs for the *Trophic metacommunity* model (Red line: 0.01 dispersal rate; Orange line: 0.1 dispersal rate). Notice that the Species-Area relationship for the *Trophic Sampling* model shown in (a) is given by $S = kA^z$, with k = 10 and z = 0.27. Shaded areas show 95% confidence intervals.



a) Trophic Sampling

b) Trophic Theory of Island Biogeography

c) Trophic Metacommunity