# Ecosystem size and complexity are extrinsic drivers of food chain length in branching networks.

Justin P.F. Pomeranz[1]

Jacques C. Finlay[2]

Akira Terui[1]

[1]Department of Biology, University of North Carolina at Greensboro

[2] Department of Ecology, Evolution and Behavior, University of Minnesota

## Abstract

Understanding the drivers of food chain length in natural communities has intrigued ecologists since the publication of 'food cycles' by Elton in the early 20th century. Proposed drivers of food chain length have included extrinsic controls such as productivity, disturbance regime, and ecosystem size, as well as intrinsic factors including food web motifs. However, current theories have largely assumed simple, two-dimensional habitat architectures, and may not be adequate to predict food chain length in ecosystems which have a complex, branching structure. Here, we develop a spatially explicit theoretical model which provides an integrated framework for predicting food chain length in branching networks. We show food chain length responds independently to both ecosystem size and complexity, and that these responses are contingent upon other extrinsic and intrinsic controls. Our results show that accounting for ecosystem complexity is an important driver of food chain length and may reconcile inconsistent results from empirical studies of food chain length in river ecosystems.

## Introduction

The origin of variation in food chain length (FCL), the number of feeding links from the basal species to top predators, has intrigued ecologists for decades due to its importance in regulating top-down effects [1], energy flow[2], and contaminant concentrations in top predators that humans often consume[3].

Hypotheses regarding controls of food chain length can be organized into three broad categories. First, the resource availability hypothesis states that basal productivity controls food chain length because imperfect energy conversion through predator-prey interactions restricts energy available to higher trophic levels[4]. Second, the disturbance hypothesis (or the dynamical stability hypothesis) predicts longer food chains are vulnerable to environmental perturbations (i.e., the dynamical constraint). Thus, habitats with frequent and/or severe disturbance should support fewer trophic levels[5]. Lastly, the ecosystem size hypothesis posits that larger ecosystems should harbor longer food chains through a suite of mechanisms, including enhanced colonization rates[6] increased habitat heterogeneity[7] and surpassing spatial constraints inhibiting metapopulation dynamics.

Among the three hypotheses, the ecosystem size hypothesis has been consistently supported in both theoretical and empirical explorations, likely because it encompasses multiple mechanisms which underlie the positive association[8,9]. For example, large ecosystems have increased species richness (species-area relationship[8]), buffer against disturbances[9], and have greater numbers of immigrant sources in the ecosystem, harboring stable metapopulations of constituent species that make up long food chains [6,10,11]. Meanwhile, the theoretical basis for the other two hypotheses rest on implicit assumptions of linear food chains, which are often violated in natural systems. As such, the effects of disturbance and resource availability on food chain length have been found to be highly variable among ecosystem types and geographic regions [12–14].

More recently, food web ecologists have begun to recognize that FCL rarely responds to a single driver, but instead is determined by interactions of multiple factors. For example, Ward and McCann[15] showed that FCL was more responsive to increasing productivity in large ecosystems compared with small ecosystems, whereas FCL was more sensitive to size in low productivity systems . Such context dependent responses to extrinsic factors seem to be tightly coupled with the arrangement of trophic links within a food web, often referred to as food web motifs[16,17]. For example, omnivory (feeding at multiple trophic levels) is a strong intrinsic driver of food chain length [6,15,18,19]. When omnivory is absent in three-species communities, the trophic structure assumes a simple linear food chain with maximum length equal to 3 (**Figure 1**). However, as omnivory increases, FCL decreases proportionally as the top predator eats lower in the food web [15,20]. Theoretical work predicts that food webs are destabilized when omnivory is prevalent, leading to the loss of species and shorter food chains even in productive, large ecosystems. Conversely, weak omnivory stabilize the food web and increase persistence of species within a community, particularly at intermediate levels of productivity and high disturbance levels [20,21].

Omnivory is a dominant module in natural communities [22] but is often not accounted for in theoretical studies[20], but see [15]. Due to the context-dependency of FCL to multiple drivers, and the potential for interactive effects, a unified framework of multiple food chain drivers is needed to better understand when and where FCL will respond to these extrinsic and intrinsic factors.

While recent research efforts of food web ecologists have substantially advanced our understanding of food chain drivers, most of the previous work has been conducted in simple, two-dimensional systems, such as lakes[7,23,24] and oceanic islands [25,26]. However, many natural habitats have a high degree of spatial complexity that cannot be represented simply by ecosystem size. Branching ecosystems, for example, are ubiquitous yet overlooked landscape structures [27], in which geomorphological or biological processes form naturally fractal branching patterns, where the geometric arrangement of patches is similar across spatial scales[28] (e.g., rivers, trees, caves, mountain ranges). While we know that food web interactions are spatially influenced[29,30], the theory to account for it has lagged, particularly in branching habitat architectures. Here, we theoretically explore drivers of food chain length in rivers, an excellent example of branching ecosystems.

Ample evidence suggests that branching structure controls environmental heterogeneity and dispersal pathways of organisms within river ecosystems. For example, water and materials propagate downstream as small streams hierarchically join to form larger streams and rivers. Consequently, environmental signals are highly variable among tributaries while showing a strong within-tributary correlation e.g., carrying capacity, flood disturbance[31,32]. Confluences serve as merging points where branch-specific dynamics aggregate to form distinct environmental conditions in downstream habitat patches [33]. Habitat heterogeneity therefore increases with branching complexity (**Figure 1**), which we define here as the probability of branching per unit river distance. Hence, highly-branched river networks provide diverse habitats that buffer the impact of large-scale environmental fluctuations [34], broadening opportunities for species to recolonize habitat patches post-disturbance. Thus, ecosystem complexity in river networks may be a widespread yet overlooked extrinsic environmental driver which needs to be incorporated to understand food chain length.

We develop a model in which trophic community dynamics within branching habitats are simulated. We investigate extrinsic controls such as ecosystem size (number of patches), complexity (branching probability), disturbance regime and productivity level (carrying capacity). In addition, our simulation model incorporates intrinsic drivers of food chain length including food web motifs (degree of omnivory) and dispersal ability, allowing us to investigate the context dependency on drivers of food chain length

across extrinsic factors. We find that ecosystem size and complexity are both positively related to FCL, but the form of this relationship is dependent on other extrinsic and intrinsic factors. Trophic structure (degree of omnivory) and disturbance regime are both strong drivers of FCL on their own and have interactive effects. We also find that productivity and dispersal ability were both positively related to FCL. Our model is an important step in unifying drivers of FCL across extrinsic and intrinsic drivers and parsing their context-dependency in spatially complex ecosystems.

## Results

First, we depicted branching ecosystems as connected habitat patches, which local communities inhabit. Each local community is composed of basal (B), primary consumer (C), and top predator (P), and these constituent species disperse along the network corridors following exponential dispersal kernels. The arrangement and strength of trophic interactions can be varied based on the simulation parameters (food web motifs, **Figure 1**, see **Methods**). Periodically, local communities are impacted by disturbance, whose strength is highly correlated within a tributary. Therefore, our simulation framework resembles environmental dynamics of natural river systems.

We explored how food web motifs affected FCL across extrinsic factors. FCL is related to the degree of omnivory of the predator P (**Figure 1**). When P only consumes C, its trophic level is 3. However, as the proportion of B in P's diet increases, FCL decreases, reaching a minimum of 2 when B constitutes all of P's diet. We assume that P searches for both resources proportionally based on the conversion efficiency for each prey resource (i.e., prey quality) and the relative densities of each resource (see **Methods**). To control the level of omnivory, we varied the conversion efficiencies of both prey resources.

In addition to food web motifs, we also investigated the effects of species dispersal ability and disturbance regime. To simplify our simulations, we looked at two scenarios of dispersal ability. 1) Low dispersal ability, in which a small proportion of the population of all three species emigrate, and the likelihood of successful immigration to neighboring patches declines exponentially with increasing distance (i.e., short-distance dispersal). 2) High dispersal ability, in which a larger proportion of the population emigrates, and the likelihood of successful immigration to neighboring patches declines less-rapidly with increasing distance (i.e., long-distance dispersal; see **Methods**).

The effect of disturbance was also investigated by looking at low and high disturbance regimes. The low disturbance regime was defined as having a low probability of a relatively weak magnitude disturbance

occurring throughout the habitat network in each time step. This can be likened to spring-fed streams which are relatively stable through time but are subjected to occasional disturbances such as floods from high precipitation events. In contrast, we investigated a high disturbance regime in which the probability and magnitude of disturbances occurring throughout the habitat network was much higher. This scenario could represent "flashy" streams which are predominantly driven by surface runoff and precipitation events.

We examined interactive effects of extrinsic and intrinsic factors controlling food chain length. We simulated habitat networks over a gradient of both ecosystem size and complexity (branching probability). For each habitat network, we simulated trophic dynamics under two different regimes of disturbance (rare or frequent), productivity (low or high carrying capacity of the basal species), and dispersal ability (low and high), and three local food web structures (the degree of omnivory, **Figure 1**), for a total of 24 scenario combinations. We simulated meta-food web dynamics over 1200 time steps, including 200 steps for burn in, and yielded estimates of food chain length averaged in space and time (see **Methods**). Varying the disturbance regime resulted in markedly different results across ecosystem size and complexity. Thus, we present our results for each disturbance regime separately below.

**Low disturbance regime**

In ecosystem networks with a low disturbance regime (i.e., dynamical stability), FCL was largely determined by food web motifs (colored lines in **Figure 2**). Generally, FCL was negatively related to the degree of omnivory, with simple linear food chains (P only consumes C) having the longest FCL, and strong omnivory (P consumes very little C) having the shortest FCL. When carrying capacity and dispersal ability was low, there was a slight increase in FCL with the omnivorous motifs across increasing ecosystem size, and when dispersal ability was high, there was a moderate increase for the strong omnivory module (**Figure 2A**). In addition, there was a slight increase in FCL with high dispersal ability across both carrying capacities, but the effect was greater at low carrying capacity. However, when carrying capacity was increased, FCL was largely invariant across gradients of ecosystem size.

There was very little effect of ecosystem complexity on FCL in the low disturbance regime. In all scenarios, FCL did not vary in response to increasing ecosystem complexity, although there was a very slight decline for all three food web modules when carrying capacity and dispersal ability were low. FCL was predominantly determined by food web motifs and carrying capacity, with FCL being negatively related to omnivory, and positively related to carrying capacity. Like the results across ecosystem size,

there was a positive effect of high dispersal ability on FCL, and this effect was greater in ecosystems with low carrying capacity (interactive effect).

**High disturbance regime**

Ecosystem networks with a high disturbance regime had much different responses of FCL across both ecosystem size and complexity. FCL increased across both gradients under all combinations of carrying capacity and dispersal ability, and in most food web motifs (**Figure 3**). However, the strength of the relationship was variable for some motifs, and when dispersal ability and carrying capacity were low, and omnivory was absent, FCL was largely invariant across both gradients (top right panel in **Figure 3A and 3B**. Once again, FCL was largely determined by the degree of omnivory, with simple linear food chains generally having longer FCL, and weak IGP having the shortest FCL.

When dispersal ability and carrying capacity were low, however, a simple linear food chain structure resulted in the shortest FCL across both ecosystem size and complexity. It is likely that this simple motif with a few strong links was unstable and could not persist when exposed to frequent, high magnitude disturbances (e.g., supports dynamical stability hypothesis). In these environments, having prey choice (omnivory) appears to be a necessary requirement to maintain the persistence of species.

Increasing dispersal ability and carrying capacity had positive effects on FCL across both ecosystem size and complexity. Similar to the results in the low disturbance regime, increasing dispersal ability in ecosystems with low carrying capacity resulted in a large increase in FCL, suggesting an interactive effect.

## Discussion

The ecosystem size hypothesis of food chain length has received support from both theoretical and empirical work. Because ecosystem size encompasses several distinct mechanisms that could influence food chain structure, work so far has not elaborated the underlying drivers. Ecosystem complexity is an extrinsic factor which is scale-invariant and related to ecosystem size yet has not been incorporated in theoretical explorations of food chain length to date. Here, we show that food chain length responds positively and independently to increasing ecosystem size and complexity. This supports our hypothesis that scale-dependent theories of FCL are inadequate to correctly predict FCL in ecosystems with a complex structural geometry. However, the relationship between FCL and ecosystem complexity varied based on additional extrinsic and intrinsic factors, supporting the emerging paradigm that there is no

one single driver of FCL in ecosystems [15], but moreover that there is context dependency for which factor is the primary driver of the relationship. Although we focus on river networks here, we expect our results to be transferable to other branching structures, including mountain ranges and cave systems.

While increased disturbance intensity and frequency consistently reduced food chain length as predicted by dynamical stability hypothesis [5], more important findings in this study were evident from their synergistic effects with ecosystem structure. Our low disturbance regime scenario can be viewed as habitat networks having nearly homogeneous environments, in which long food chains were generally supported independent of ecosystem size and complexity and FCL was predominantly determined by food web motifs. However, under high disturbance regimes, only large and complex networks could provide diverse and heterogeneous habitats with potential refugia, allowing species to thrive and maintain long food chains. This result has important implications for ecosystem management because anthropogenic climate change is predicted to increase the frequency and magnitude of disturbances [35,36]. River networks are especially vulnerable due to their propensity to go dry during droughts[37] as well as susceptibility to more intense flood disturbance with increasing storm intensity[38]. Increased dispersal ability alleviated the reduction in FCL from increased disturbance regime in our simulation, and reduced the negative impacts of disturbance and climate change in empirical studies [39]. However, human activities have altered the habitat heterogeneity, connectivity, and dispersal ability of species within riverine networks [40,41]. In-stream habitat in rivers and streams have been homogenized and lateral connections to the floodplain have been severed by the removal of large woody debris, channelization and straightening of river corridors, and through flow regulation [42,43], and this has affected community assembly[44] and trophic structure [45].

Although dispersal has received little attention in food chain research, our simulation highlighted its importance in regulating food chain length. In particular, high dispersal was required for ecosystem size and complexity to have consistent positive effects on FCL under high disturbance regimes. This is probably because sufficient propagules are needed to successfully recolonize after disturbance and maintain long food chains in large and/or complex networks. Human activities that affect habitat connectivity and dispersal ability of species within riverine networks may compromise the buffering effects of ecosystem size and complexity against disturbance by limiting dispersing propagules. For example, dams and diversion have altered the connectivity between suitable habitat patches, or have completely blocked dispersal ability [46,47]. Light pollution has also negatively affected the ability of adult aquatic insects to disperse between habitat patches [48].

The context dependency of disturbance, productivity and dispersal has important implications for the conservation and management of freshwater ecosystems. This adds an important distinction in the discussion of the suitability of assigning protection or conservation status to regions. Current conservation schemes generally emphasize size, aiming to protect the largest areas possible. While this is certainly a useful attribute generally, designating regions with a high degree of variability, or many small patches, may be more beneficial than a large, homogenous area [49], and accounting for ecosystem complexity is particularly suited to ecosystem architectures in which branches have variable environmental conditions. Conservation and management activities should focus on activities that increase the habitat heterogeneity and connectivity of suitable habitat patches to increase resilience to predicted increases in disturbance frequency and magnitude.

The role of food web motifs on FCL largely followed our predictions (**Figure 1**). Specifically, FCL declined as the value of the resource B, and thus the proportion in the diet, to the predator P increased. Essentially, the relative qualitative values of the resources to the predator controlled the level of omnivory and FCL. Although FCL was greatest when the community was arranged as a simple linear food chain, this arrangement was not the most stable. Specifically, when disturbance regime was high, and productivity and dispersal ability were low, the observed FCL across networks was ~1 (although increasing dispersal ability and/or carrying capacity resulted in food chains of ~2.5-3). Therefore, in the absence of prey resource choice, strong, top-down trophic interactions were destabilizing, leading to the extirpation of consumer species at higher trophic levels across the networks.

Changes in community composition are occurring worldwide and in all habitats. Species extirpations and extinctions, coupled with species introductions and invasions are creating novel communities worldwide. In Lake Tahoe (USA), for example, replacement of the top fish predator and the introduction of a freshwater shrimp substantially altered the food web structure [50]. Understanding the interactions within these novel communities, and how this could affect food chain lengths is an important goal of contemporary ecologists. The strength of omnivory in these novel communities can be increased or decreased, depending on the introduced species and the extant species pool or community. For example, non-indigenous fishes in the Eastern Mediterranean Sea outcompete native species for high quality prey items, causing native species to increase consumption of non-preferred prey items[51]. Introductions of non-native fishes in the Colorado River basin have led to longer FCL and narrower niches (e.g., lower degree of omnivory) of native species[52]. Alternatively, the introduction of hyper-successful basal or primary consumer species can weaken or remove the omnivory link by becoming the

dominant prey resource. For example, nonnative New Zealand mud snail dominated the macroinvertebrate biomass in streams in Yellowstone National Park and likely affected food web structure[53]. Having a single prey resource dominate the community is akin to reducing the degree of omnivory (i.e., removing prey choice), which destabilizes the community, particularly under high disturbance regimes.

It is abundantly clear that ecosystem complexity in natural systems is the rule, not the exception. Early theoretical and empirical work has assumed simple habitat structure as a necessity. However, habitat heterogeneity, such as the presence of prey refugia, and dispersal between connected habitat patches allow communities to stabilize and persist through space and time [54–56]. Complex ecosystem structure is an extrinsic factor that can affect food chain length in communities. Incorporating a realistic understanding of ecosystem structural complexity could alter our perception of communities in naturally complex systems, potentially leading to paradigm shifts in macroecology.

## Methods

### Network Generation

We simulated branching networks of connected habitat patches as a series of nodes (Figure 1). The geometric arrangement of networks was controlled with two parameters: ecosystem size, the number of patches ($N_P$); and ecosystem complexity, branching probability ($P_B$). Nodes (habitat patches) were assigned to be a confluence (or upstream terminal node) with probability $P_B$, or an in-branch node with probability 1 - $P_B$. Branches (tributaries) are a series of connected nodes between confluences (or upstream terminal nodes). The number of patches within a branch, $q$, is a random variable drawn from a geometric distribution $n_{p,q} \sim Ge(P_b)$ (a geometric distribution is the discrete form of an exponential distribution). The number of branches $N_B$ is a random variable drawn from a Binomial distribution: $N_B \sim Binomial(N_P, P_B)$. This framework preserves the fractal nature of branching patterns. For each scenario described below, we generated 1000 networks across a gradient of ecosystem size and complexity by drawing a random variable from uniform distributions as: $N_P \sim Unif(5, 150)$ and $P_B \sim Unif(0.01, 0.99)$ ($N_p$ was rounded to the nearest integer before running simulations).

For each terminal branch (i.e., headwater branch), $h$, within the network, a disturbance value, $m_h$, was sampled. The distribution of $m_h$ values was sampled from a normal distribution in a logit scale. This was accomplished by transforming the proportional mean disturbance value, $\mu_m$, to the logit scale and sampling a value for each terminal node (headwater) as $m_h \sim$ Normal( logit( $\mu_m$), $\sigma^2_m$). Here, we varied

the mean source disturbance value, $\mu_m$ = (0.1, 0.9), based on disturbance regime (see **Simulations**) and set the variation of source disturbance as $\sigma^2_m = 2$. All the patches from terminal node moving downstream to a confluence node have identical disturbance values, $m_h$. At confluence nodes, the disturbance value takes a weighted mean based on the relative size of the two contributing branches, *s* (the number of upstream contributing patches): $m_{down} = \omega m_{1,up} + (1 - \omega)m_{2,up}$, where $m_{down}$ is the disturbance magnitude of the focal patch (the confluence), $m_{1,up}$ and $m_{2,up}$ are the disturbance magnitude of contributing tributaries, and $\omega = s_1/(s_1 + s_2)$. Disturbance value distributions had medians close to $\mu_m$ with long right or left tails (**Fig S1**), respectively, i.e., when $\mu_m = 0.1$ (e.g., a 10% reduction in species densities), most sites experienced low-magnitude disturbances and a few sites experienced high-magnitude disturbances. Likewise, when $\mu_m = 0.9$, most sites experienced high-magnitude disturbances with a few patches receiving low-magnitude disturbances. The low-magnitude patches can be viewed as refugia and a source to supply individuals to colonize impacted sites in subsequent timesteps. Example disturbance value distributions for simulation scenarios reported here are available in the SI. The R package *mcbrnet* was used to simulate the networks.

**Tri-trophic dynamic simulation model**

We simulated trophic community dynamics in branching river networks using a mixed-time model which models predation continuously and has discrete, synchronous reproduction. This is a life-history common to many lotic species, including macroinvertebrates and fish. Accounting for this stage structure is an important aspect which is often missed in continuous-time models[57]. We describe the model in detail below.

Our simulation model consists of three species communities; a basal species B which has positive logistic growth up to a carrying capacity, K; a primary consumer species, C which grows based on the consumption of B; and a predator P, which can consume both B and C. Let $N_{ix}(t)$ represent the population density of species *i* in patch *x* at time *t.* The model starts with some initial nonzero population densities for all three species in each patch. The initial population abundances for each species *i* in each patch *x* were independently sampled from a random Poisson distribution Poisson($\boldsymbol{\lambda}_{ix}$), where $\boldsymbol{\lambda}_{ix}$ for B, C and P was 0.8 * K, 0.5 *K and 0.25 * K, respectively. First, dispersal from all patches is accounted for. The number of emigrants, $E_{ix}(t)$for species *i* in from patch *x* at time *t*, is estimated as the product of dispersal probability, $P_{dispersal}$ and $N_{ix}(t)$. The dispersal probability was assumed to be constant for all three species in each simulation run. Then, the number of emigrants from patch *x* is divided into

immigrants, $I_{xy}(t)$ to each patch *y*. The immigration probability, $\xi_{xy}$ from patch *x* to *y* is calculated based on the habitat structure among patches according to:

$$\xi_{xy(t)} = \frac{\Sigma_{y,y\neq x} E_{xy}(t) e^{-\theta dxy}}{\Sigma_x \Sigma_{y,y\neq x} E_{xy}(t) e^{-\theta dxy}}$$

Where $d_{xy}$ is the distance between patch *x* and *y*, and θ controls the rate of decay.

The realized number of individuals $N_{ix}(t+1)$ (species $i$ at patch $x$ and time $t+1$) is given as:

$$N_{ix}(t+1) \sim Poisson(n_{ix}(t) + I_{ix}(t) - E_{ix}(t))$$

where $N_{ix}(t)$ is the number of individuals at time $t$,(i.e., before dispersal) $I_{ix}(t)$ the expected number of immigrants to patch $x$, and $E_{ix}(t)$ the expected number of emigrants from patch $x$. The realized discrete number of individuals is drawn from a Poisson distribution to account for stochasticity in demographic and dispersal processes. Note that a value of 0 is assigned to $N_{ix}(t+1)$ for any non-positive value of $\boldsymbol{\lambda}$ [58].

After dispersal, continuous predation occurs. The number of prey *i* eaten by consumer *j*, $W_{ij}$, in a season is calculated as:

$$W_{ij} = \frac{\alpha_{ij} N_i N_j}{\beta_{ij} N_j + N_i}$$

where $N_i$ and $N_j$ are resource and consumer densities, respectively and $\alpha_{ij}$ and $\beta_{ij}$ are coefficients describing the functional response as the ratio of resources, *i*, to consumers, *j*, increases. $\beta_{ij}$ describes how fast the predation curve reaches the asymptote of the Type II functional response, and $\alpha_{ij}/\beta_{ij}$ determines the location of the asymptote[57].

The predator species P, searches for and consumes both species in a patch. Its total effort is controlled by parameters $\alpha_{.p}$ and $\beta_{.p}$. The preference parameter $\delta$, controls the proportional search effort of species P for resource B and C:

$$W_{BP} = \delta(W_{ij})$$

$$W_{CP} = (1 - \delta)(W_{ij})$$

Predator preference for resource B is determined by the equation:

$$\delta = \frac{e_{BP} N_B}{e_{BP} N_B + e_{CP} N_C}$$

Where $e_{BP}$ and $e_{CP}$ are the conversion efficiencies of predator P for resource B and C, and $N_B$ and $N_C$ are resource densities, respectively. This formulation allows for flexible diet shifts of the predator based on realized energy gain (conversion efficiency times prey availability) from the prey species[59].

At each time-step within a patch, FCL was defined as 0 if no species were present, 1 if only B was present and 2 if only B + C or B + P were present. When all three species were present, FCL was calculated based on the proportion of each resource in P's diet ($\lambda_i$) according to:

$$FCL = (1 * \lambda_B + 2 * \lambda_C) + 1$$

When P only consumes B or C, FCL = 2 or 3, respectively. When P consumes both resources equally, FCL = 2.5. In each simulation, FCL was averaged across all patches and time-steps.

The number of individuals in each patch is reduced by the number consumed $W_{ij}$ (predation effect), as well as by a base survival probability, $s_0$:

$$N_B{}' = s_0(N_B - W_{BC} - W_{BP})$$
$$N_C{}' = s_0(N_C{}' - W_{CP})$$
$$P' = s_0 P$$

Finally, those individuals remaining reproduce according to the following:

$$N_{B,t+1} = (\frac{r_{max}}{1 + \upsilon N_B{}'} N_B{}')(1 - \varphi m_x)$$
$$N_{C,t+1} = (W_{BC} N_C{}')(1 - \varphi m_x)$$
$$N_{P,t+1} = (e_{BP} W_{BP} + e_{CP} W_{CP}) N_P{}'(1 - \varphi m_x)$$

where $r_{max}$ is the maximum per-capita reproduction rate of the basal species and $\upsilon = \frac{r_{max}-1}{K}$ determines the asymptotic relationship of population growth, where K is carrying capacity. In the absence of predation, population growth of the basal species follows a logistic growth relationship. For the consumer species (C and P) $e_{ij}$ is the efficiency of converting resource *i* into new consumer *j*. The last term in the equation accounts for the proportional mortality in the event of a disturbance, where $m_x$ = the disturbance value (see **Network Generation**) and $\varphi$ is an indicator variable of disturbance:

$$\varphi = 0, no\ disturbance$$

$$\varphi = 1, disturbance\ occurs$$

If species B is extirpated from a patch, both the other species were also assumed to be extirpated from that patch during that time step. Species B can colonize any open patch, whereas species C and P can only successfully colonize patches where B was already present (i.e., prey resource is necessary for successful establishment).

**Simulations**

We used 24 combinations of parameter sets (**SI Table**). The parameter sets include two environmental factors, and twelve ecological scenarios. For environmental scenarios, we used two parameter sets that controlled the frequency and magnitude of disturbances (low and high disturbance regime, respectively). In the low disturbance regime, the probability of a disturbance occurring at each time step was 0.0001 (i.e., a 10,000 year recurrence), and the mean disturbance magnitude was $\mu_m = 0.1$ (10% reduction in population densities). In contrast, the high disturbance regime had a probability of 0.01 (i.e., 100 year recurrence), and the mean disturbance magnitude was $\mu_m = 0.9$ (90% reduction in population densities[60])

For the ecological scenarios, we used two productivity levels, two dispersal abilities, and three levels of omnivory. Productivity was varied by setting the carrying capacity of the basal species, B (K = 50, 150). Dispersal ability varies the effective distance between patches and can control the probability that a patch is re-colonized post-disturbance. Two levels of dispersal ability were simulated. The high dispersal ability was set with parameters $P_{dispersal} = 0.1$ and $\theta = 0.1$ (i.e., high probability of individuals emigrating, and far distance traveled). The low dispersal scenario was set with parameters $P_{dispersal} = 0.01$ and $\theta = 1$ (i.e., low probability of dispersal, short dispersal distance). These values were chosen to represent extreme scenarios. Preliminary analyses using intermediate parameter sets had responses between these two scenarios.

Finally, we investigated three food web motifs (**Figure 1B**). Trophic interaction strengths can be controlled by varying the attack rate or conversion efficiency by the consumer P (Ward and McCann 2017). To simplify model simulations, we fixed the attack rate for both C and P (parameters $\alpha = 8$ and $\beta = 20$), and the conversion efficiency for C as $e_{bc} = 6$. We controlled the presence and strength of trophic interactions for the predator, P, by varying the conversion efficiency for both prey resources (i.e., varied prey resource quality). The scenario names and parameter levels were as follows: (1) “Food chain”: $e_{bp} = 0$, $e_{cp} = 4$; (2) “Strong IGP”: $e_{bp} = 2$, $e_{cp} = 4$; (3) “Weak IGP”: $e_{bp} = 4$, $e_{cp} = 2$.

Under each scenario, we simulate 1200 timesteps of trophic dynamics (including 200 steps of burn-in) in 1000 branching networks across a gradient of ecosystem size (number of habitat patches, from 5 to 150) and ecosystem complexity (branching probability, from 0.01 to 0.99). Simulations ran for 200 time steps as a burn-in period to minimize influences of initial conditions. The last 1000 timesteps were recorded to estimate the temporal average of food chain length across all patches within a network.

Values of simulation parameters are summarized in (SI). We performed simulations using R version 4.0.3. R functions for the generation of branching networks are provided as the R package *mcbrnet*. (available at https://github.com/aterui/mcbrnet). R functions for the trophic community simulations are provided in the R package *IGPtoy* (available at https://github.com/Jpomz/IGPtoy [*IGPtoy* will eventually be rolled into *mcbrnet* but leaving it here for now]). Codes for simulations, statistical analysis, figures, and tables are available at https://github.com/Jpomz/priv-igp-sens.

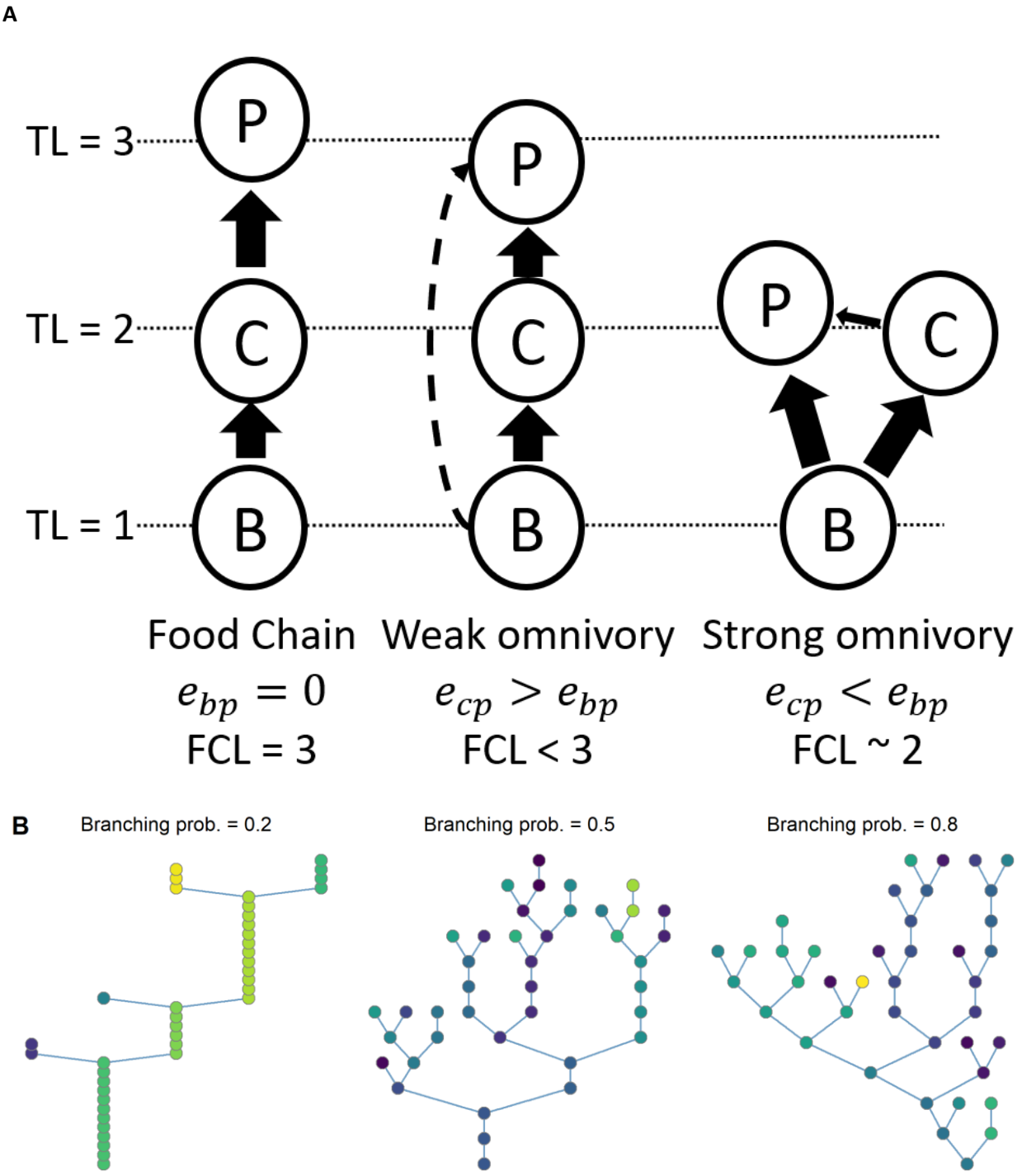


Figure 1. Food web motifs showing trophic interactions in a three-species community (A) and, examples of how branching probability changes network topology (B). In (A), B refers to a basal species whose population level is controlled by a density-dependent function. The consumer species, C, consumes B as a resource. The predator, P, is able to consume both B and C, and the structure and strength of these trophic interactions are a function of resource quality and relative densities (i.e., $e_{bp}$ and $e_{bc}$). The strength of trophic interactions involving P is represented by the thickness of the arrows (strength of the

B-C interaction is fixed), and food chain length varies in response to the proportion that each resource makes to the diet of P. On the far left, a simple linear food chain exists when the predator, P, does not consume the resource B. Moving to the right, weak omnivory exists when the resource C is much more valuable to P than the resource B is. Strong omnivory exists when resource B is more valuable to P than prey resource C. If C goes extinct, P only consumes B, and the FCL = 2. If P goes extinct C is the top species and FCL = 2. Likewise, when both C and P are extinct, FCL = 1, and if all three species go extinct FCL = 0. In (B), node color is reflective of the disturbance value calculated in network generation. Note that it is the same within branches, and changes at confluence nodes.

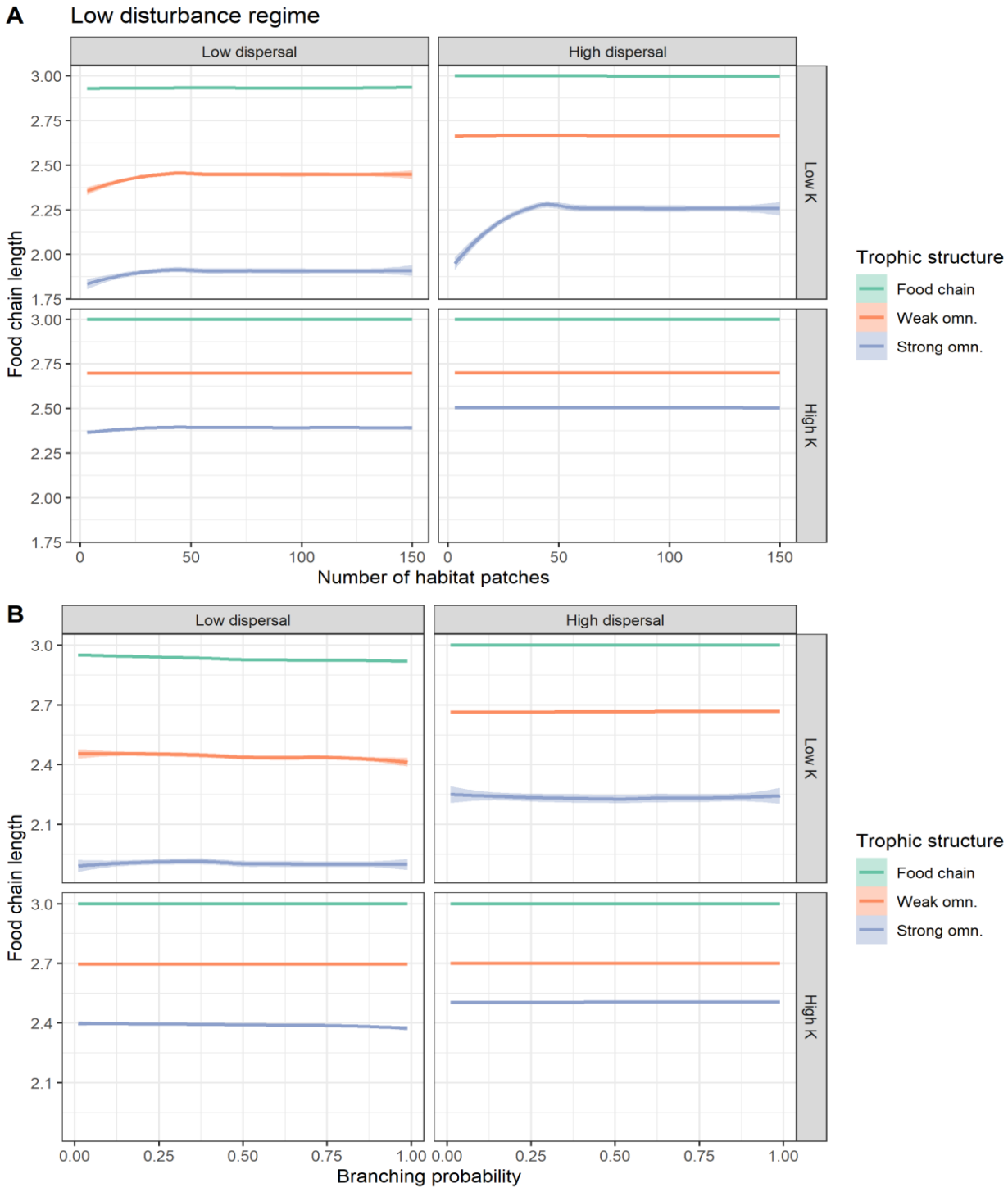


Figure 2. Food web structure is the primary driver of food chain length in low disturbance regimes. Food chain length is generally invariant across gradients of ecosystem size (A) and complexity (B). However, food chain length did respond positively to ecosystem size under some simulation scenarios (A, top panel). Lines are loess curves fitted to simulated data and are colored based on food web motif. Shaded ribbons are 95% confidence intervals (may be difficult to see). Parameters for low disturbance regime: probability disturbance = 0.0001, $\mu_m$ = 0.1. The low dispersal scenario had parameters $\theta$ = 1 and $P_{dispersal}$ =

0.01, and the high dispersal scenario had parameters $\theta = 0.1$ and $P_{dispersal} = 0.1$, respectively. Carrying capacity of the basal species, B, for the Low and High K = 50 and 150, respectively.

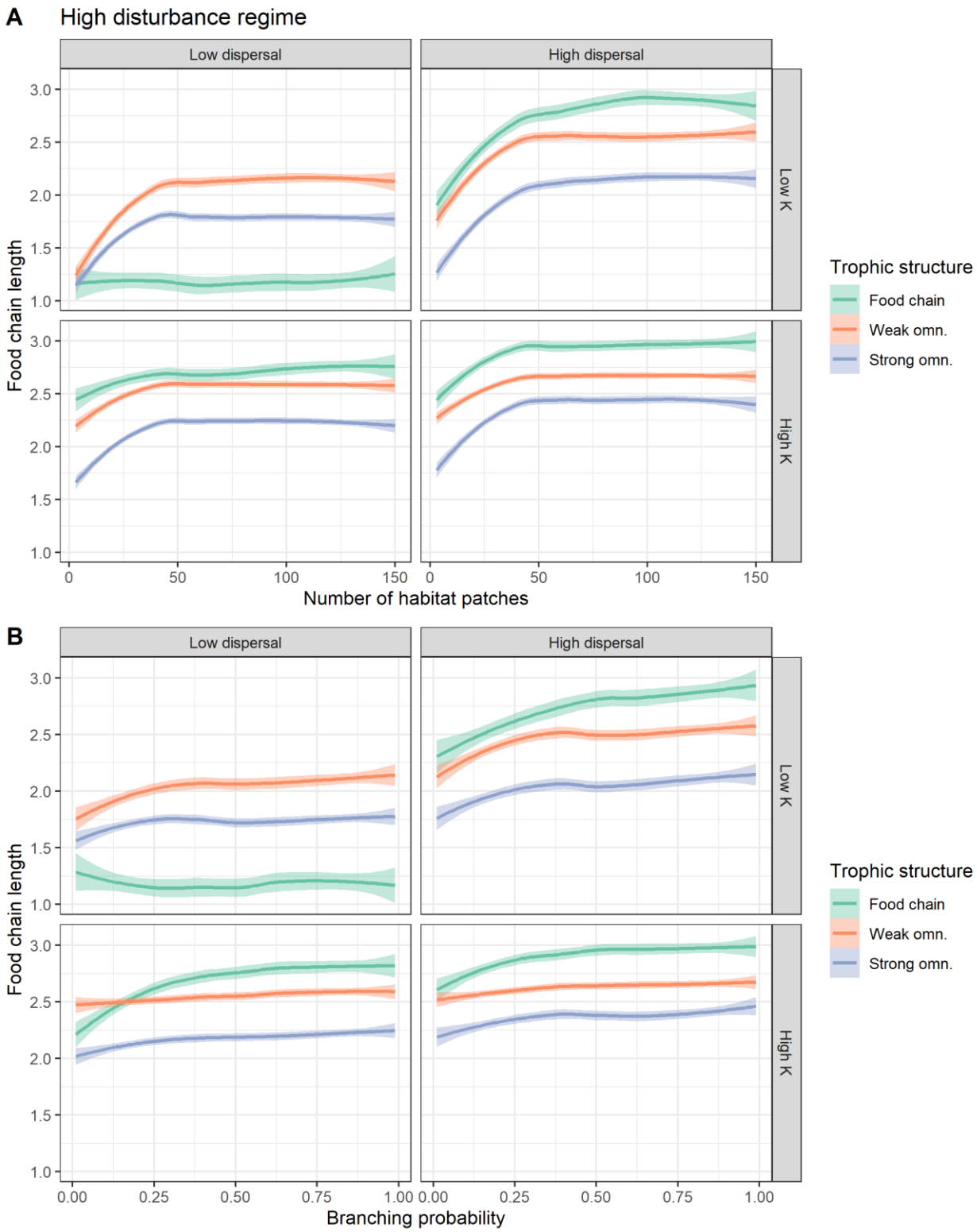


Figure 3. Ecosystem size (A) and complexity (B) increase food chain length under high disturbance regimes Lines are loess curves fitted to simulated data and are colored based on food web motif. Shaded ribbons are 95% confidence intervals (may be difficult to see). Parameters for low disturbance

regime: probability disturbance = 0.01, $\mu_m$ = 0.9. The low dispersal scenario had parameters $\theta$ = 1 and $P_{dispersal}$ = 0.01, and the high dispersal scenario had parameters $\theta$ = 0.1 and $P_{dispersal}$ = 0.1, respectively. Carrying capacity of the basal species, B, for the Low and High K = 50 and 150, respectively.

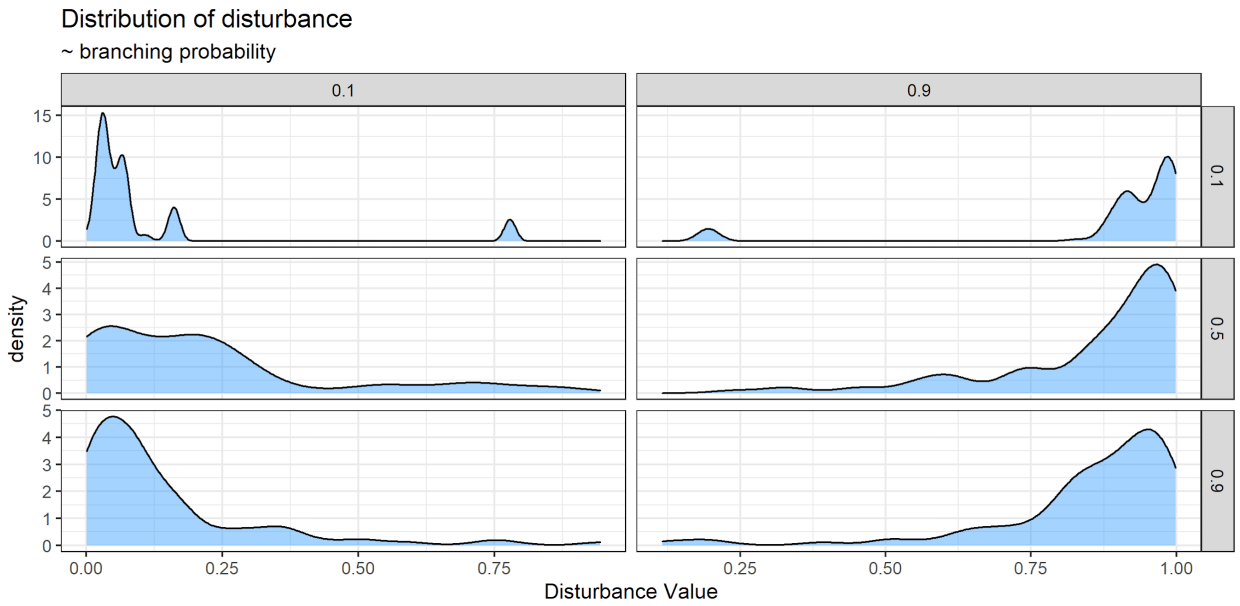


Figure S1. Example of disturbance value distributions (columns) across branching probabilities (rows) in simulations. Here, the disturbance value represents the proportional reduction in abundances in the event of a disturbance (i.e., a disturbance value of 0.9 means a 90% reduction in abundances). In the left column, the mean source disturbance parameter, $\mu_m$, was set to 0.1, and in the right column it was 0.9. In networks with a low branching probability (top row, $P_{branch} = 0.1$) there is very little variation in the disturbance values throughout the network. When branching probability is high (bottom row, $P_{branch} = 0.9$), the variability of disturbance values increases due to the confluence of more branches within the network. For these examples the number of patches was set to 100, and the $\sigma_m = 2$.

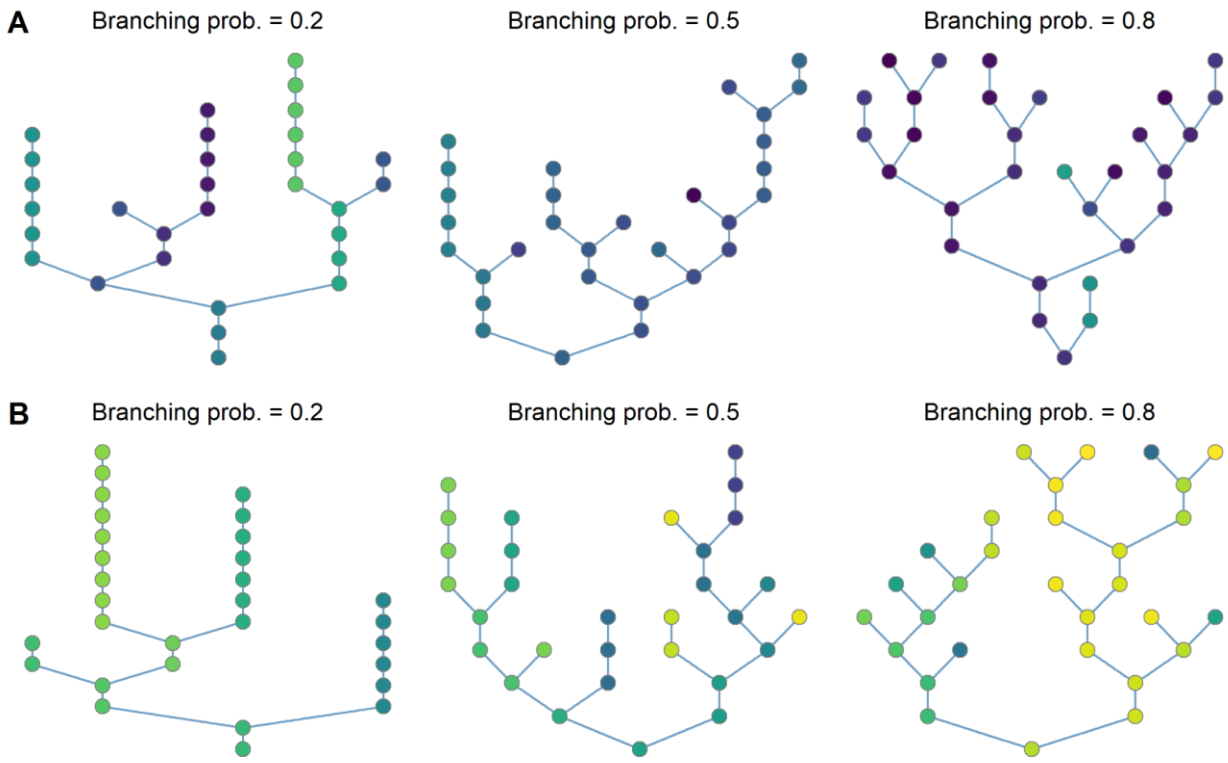


Figure S2. Branching networks with mean disturbance, $\mu_m$ = 0.1 (A) and $\mu_m$ = 0.9 (B). Nodes are color-coded with increasing disturbance values from dark blue to yellow. Note that the colors in row A have less variability than the colors in row B, and both rows have higher color variability with increasing network complexity (branching probability). This figure highlights the fact that networks in the low-disturbance regime (top row), had less habitat heterogeneity than the networks generated in the high-disturbance regime.

Table S1. Parameter values for main simulation. Not all possible combinations of all variables were used. See the main text for parameter combinations for specific scenarios.

| Parameter | Value | Interpretation |
|---|---|---|
| $\sigma_h$ | 2 | Variation of environmental value at headwater (source) nodes |
| $\sigma_l$ | 0.01 | Longitudinal variation of environmental value within branch |
| $\mu_m$ | 0.1, 0.9 | Mean disturbance magnitude (i.e., 0.1 = 10% reduction in population abundances) |
| $\sigma_m$ | 2 | Variation in disturbance magnitude |
| $P_{disturb}$ | 0.001, 0.01 | Probability of a disturbance occurring across the network at each timestep |
| $K_{base}$ | 50, 150 | Mean carrying capacity of the basal species, B |
| $R_{max}$ | 4 | Maximum reproductive rate of the basal species, B |
| $\theta$ | 0.1, 1 | Parameter controlling the rate of exponential decay in dispersal distance |
| $P_{dispersal}$ | 0.01, 0.1 | Probability of individuals dispersing in each time step |

| | | |
|---|---|---|
| $\alpha_C$, $\alpha_P$ | 8 | Parameter controlling the search effort of species C, and P, respectively |
| $\beta_C$, $\beta_P$ | 20 | Parameter controlling rate that consumer species reaches search effort asymptote. |
| | $\alpha_j/\beta_j$ | Location of the search effort asymptote for consumer species *j*. |
| $e_{BC}$ | 6 | Conversion efficiency of turning resource B into new consumer, C. |
| $e_{CP}$ | 2, 4 | Conversion efficiency of turning resource C into predator P. |
| $e_{BP}$ | 0, 2, 4 | Conversion efficiency of turning resource B into consumer P. |
| | $e_{ij}\alpha_j$ | Maximum reproductive rate of consumer *j*. |

Table S2. Parameter sets for ecological scenarios

| Scenario | Parameter Values |
|---|---|
| Food Chain | $e_{BC} = 6$, $e_{CP} = 4$, $e_{BP} = 0$ |
| Weak omnivory | $e_{BC} = 6$, $e_{CP} = 4$, $e_{BP} = 2$ |
| Strong omnivory | $e_{BC} = 6$, $e_{CP} = 2$, $e_{BP} = 4$ |
| Low Disturbance regime | $P_{disturb} = 0.0001$, $\mu_m = 0.1$ |
| High Disturbance | $P_{disturb} = 0.01$, $\mu_m = 0.9$ |
| Low Dispersal ability | $P_{dispersal} = 0.01$, $\theta = 1$ |
| High Dispersal ability | $P_{dispersal} = 0.1$, $\theta = 0.1$ |