ECOLOGY LETTERS

Towards a mechanistic understanding of variation in aquatic food chain length

Journal:	*Ecology Letters*
Manuscript ID	ELE-00668-2023.R2
Manuscript Type:	Letter
Date Submitted by the Author:	29-Aug-2023

Complete List of Authors:	Guo, Guanming; Jiangxi Normal University, Key Laboratory of Poyang Lake Wetland and Watershed Research Barabas, Gyorgy; Linkopings universitet, IFM; Takimoto, Gaku; The University of Tokyo, Graduate School of Agricultural and Life Sciences; Bearup, Daniel; University of Kent, School of Mathematics, Statistics and Actuarial Sciences Fagan, William; University of Maryland, Biology Chen, Dongdong; Chengdu Institute of Biology, Chinese Academy of Sciences Liao, Jinbao; Jiangxi Normal University, Key Laboratory of Poyang Lake Wetland and Watershed Research; Yunnan University

Towards a mechanistic understanding of variation in aquatic food chain length

Guanming Guo[1], György Barabás[2, 3], Gaku Takimoto[4], Daniel Bearup[5], William F. Fagan[6], Dongdong Chen[7], Jinbao Liao[1,8*]

[1]Key Laboratory of Poyang Lake Wetland and Watershed Research, School of Geography and Environment, Jiangxi Normal University, Ziyang Road 99, 330022 Nanchang, China

[2]Division of Theoretical Biology, Department IFM. Linköping University, SE-58183 Linköping, Sweden

[3]Institute of Evolution, Centre for Ecological Research. Konkoly-Thege Miklós út 29-33, H-1121 Budapest, Hungary

[4]Graduate School of Agricultural and Life Sciences, The University of Tokyo, 1-1-1 Yayoi, Bunkyo-ku, Tokyo 113-8657, Japan

[5]School of Mathematics, Statistics and Actuarial Sciences, University of Kent, Parkwood Road, Canterbury CT2 7FS UK

[6]Department of Biology, University of Maryland, College Park, Maryland 20742 USA

[7]CAS Key Laboratory of Mountain Ecological Restoration and Bioresource Utilization & Ecological Restoration Biodiversity Conservation Key Laboratory of Sichuan Province, Chengdu Institute of Biology, Chinese Academy of Sciences, Chengdu, China

[8]Centre for Invasion Biology, Institute of Biodiversity, School of Ecology and Environmental Science, Yunnan University, 650500 Kunming, China

Emails: G.G. (ggming1990@163.com); G.B. (dysordys@protonmail.com)
G.T. (gakut@es.a.u-tokyo.ac.jp); D.B. (d.bearup@kent.ac.uk)
W.F.F. (bfagan@umd.edu); D.C. (chendd@cib.ac.cn)

*__Corresponding author__: Dr. Jinbao Liao (jinbaoliao@163.com)
Address: Ziyang Road 99, 330022 Nanchang, Jiangxi Province, China.

Short running title: Determinants of food chain length

Article type: Letters **Number of Abstract: 137** **Number of words: 4981**

Number of references: 56 **Number of tables: 1** **Number of figures: 5**

Keywords: Food chain length; multiple environmental drivers; ecosystem size; competition-colonization tradeoff; resource productivity; disturbance.

Author contributions: Conceptualization: JL; Methodology: GG, GB, JL; Investigation: GG, JL, GB, DC; Visualization: GG, GB, JL; Supervision: JL; Writing—original draft: JL; Revisions—review & editing: GB, GT, DB, WFF.

Data availability statement

Empirical datasets are available on Dryad: https://doi.org/10.5061/dryad.jsxksn0df

Both statistical and simulation codes are available on Zenodo:

https://doi.org/10.5281/zenodo.8297624

Abstract

Ecologists have long sought to understand variation in food chain length (FCL) among natural ecosystems. Various drivers of FCL, including ecosystem size, resource productivity and disturbance, have been hypothesized. However, when results are aggregated across existing empirical studies from aquatic ecosystems, we observe mixed FCL responses to these drivers. To understand this variability, we develop a unified competition-colonization framework for complex food webs incorporating all of these drivers. With competition-colonization tradeoffs among basal species, our model predicts that increasing ecosystem size generally results in a monotonic increase in FCL, while FCL displays non-linear, oscillatory responses to resource productivity or disturbance in large ecosystems featuring little disturbance or high productivity. Interestingly, such complex responses mirror patterns in empirical data. Therefore, this study offers a novel mechanistic explanation for observed variations in aquatic FCL driven by multiple environmental factors.

Introduction

Food chain length (FCL), i.e., the maximum trophic position among all members of a food web (Post & Takimoto, 2007), is an important characteristic of ecological communities. It influences: ecosystem resilience/stability, by altering the organization of trophic interactions (Pimm & Lawton, 1977; Post et al., 2000); key ecosystem functions, such as nutrient cycling (Pace et al., 1999; McIntyre et al., 2007), primary productivity and atmospheric carbon exchange (Schindler et al., 1997); and ecosystem health, by adjusting the bioaccumulation of contaminants in top predators (Kidd et al., 1995). Given the central role played by FCL in these processes, it is important to understand what determines FCL.

Currently, multiple potential drivers of FCL have been hypothesized, making significant progress in understanding variation in FCL (Post, 2007; Takimoto et al., 2008, 2012). Among these drivers, the contributions of ecosystem size, resource productivity and disturbance to FCL have received the most attention. Various theories have been developed to explain the mechanisms by which each of these drivers affects FCL. The *ecosystem-size hypothesis* proposes that larger ecosystems should have longer food chains, simply because they can provide greater habitat availability and suitability for top predators (Holt, 1996; Takimoto et al., 2008). The *resource-productivity hypothesis* posits that FCL is ultimately constrained by the efficiency with which energy is transferred between trophic levels. Thus, longer food chains should occur in more productive systems where basal energy supply is greater (Post, 2002a; Takimoto et al., 2012, 2013). Similarly, the *productive-space hypothesis*

predicts that increasing both ecosystem size and resource productivity will increase FCL, as both can result in an increase in the resource base of the community (Doi et al., 2009; Young et al., 2013). Finally, the *disturbance hypothesis* predicts that disturbance will reduce FCL, as long food chains are more fragile in environments subject to more disturbance (Post, 2002a; McHugh et al., 2010).

Despite these advances, it appears that these drivers provide only a limited explanation of the variation in FCL observed empirically (Post, 2002a; Sabo et al., 2010). Several empirical studies have documented no strong effect of ecosystem size (Warfe et al., 2013; Young et al., 2013) and resource productivity (Hairston & Hairston, 1993; Post, 2002a) on FCL. Given that the *productive-space hypothesis* is based on these two factors (Doi et al., 2009), the mixed evidence for them also applies to it (Spencer & Warren, 1996; Vander Zanden & Rasmussen, 1999). Furthermore, there is no strong empirical evidence to directly support the idea that disturbance could limit FCL through dynamical constraints (Townsend et al., 1998). Thus, no particular hypothesis has received universal empirical support even within a single ecosystem type, and what exactly the relationship between these drivers and FCL is remains a topic of debate.

To reconcile the inconsistency between theoretical hypotheses and empirical observations, many mechanisms have been proposed, such as vertical energetic constraints (Arim et al., 2016; Ward & McCann, 2017), adaptive foraging (Kondoh & Ninomiya, 2009), intraguild predation (IGP; Takimoto et al., 2012), and regional metacommunity dynamics (Calcagno et al., 2011; Häussler et al., 2020). However,

these mechanistic models have often used relatively simple trophic modules to represent complex food webs (except Kondoh & Ninomiya, 2009), so that FCL varies via the addition, subtraction, or omnivory change of a few species on different trophic levels (Takimoto et al., 2012; Ward & McCann, 2017). By focusing on single drivers, these studies also have not assessed how drivers may interact, and their relative importance, when influencing FCL in more realistic, complex food webs. Thus, a systematic mechanistic understanding of observed variations in FCL driven by multiple environmental factors remains lacking, especially in complex trophic systems.

More importantly, the aforementioned mechanistic models have overlooked a significant dynamic process: spatial competition. In many aquatic ecosystems, species richness is primarily driven by competition processes for resources (Sun et al., 1988; Callaway & Josselyn, 1992; Huisman et al., 1999; Cardinale et al., 2009). A recent competition-colonization (C-C) model (Li et al., 2020) on the disturbance-diversity relationship has shown that disturbance and C-C tradeoffs can interact to facilitate different subsets of competitors to coexist (Liao et al., 2022). This mechanism produces variation in the effect of the environmental driver. Thus, we hypothesize that interactions between environmental drivers and C-C tradeoffs among basal species in complex food webs could be responsible for the mixed responses observed in FCL via bottom-up control.

In this study, we undertake a meta-analysis on an empirical dataset compiled from diverse aquatic ecosystems, to check the precise form of the relationship

between FCL and multiple drivers (including ecosystem size, resource productivity and disturbance). Subsequently, we develop a site-occupancy dynamic framework for complex trophic structures in the context of aquatic ecosystems, to provide a possible mechanistic explanation for the empirical analysis. Based on the above hypothesis, our framework considers the C-C tradeoff among basal species, and incorporates all of these drivers. Note that this framework is particularly relevant to aquatic ecosystems, which are often subject to strong spatial constraints. In addition, there is ample evidence for displacement competition among resource species. Examples include competition for light among periphyton or phytoplankton (Huisman et al., 1999; Cardinale et al., 2009), the inhibitory effect of water hyacinth on algae (Sun et al., 1988), and invasive smooth cordgrass competing with local aquatic plants (Callaway & Josselyn, 1992). These resource species can grow in streams, ponds, reservoirs, rivers and lakes, documenting the ubiquity of displacement competition in various aquatic ecosystems.

Methods

Meta-analysis of empirical studies

We reviewed empirical studies that have explicitly tested the relationships between FCL and one or more potential drivers. We screened the relevant literature from Web of Science (1950-present), using keywords related to the environmental variables in combination with those related to the response variable (FCL). Empirical studies pertaining to aquatic ecosystems (including lakes, rivers, ponds, reservoirs, streams, wetlands and freshwater everglades) were selected, while other ecosystems (terrestrial

or microbial) were excluded. The reported drivers were aggregated into three main variables: ecosystem size (including 'lake/pond volume', 'drainage/cross-sectional/watershed area' and 'stream/lake size'), resource productivity (including 'primary production/productivity', 'energy gradients', 'nutrient fertility/concentration', and 'primary biomass'), and disturbance (including 'a multivariate disturbance index', 'dynamic stability', 'average disturbance intensity', and 'flow/discharge variation'). We included only studies that explicitly recorded actual FCL or maximum trophic position estimated from complex food webs, excluding studies evaluating FCL only by the presence or absence of top or intermediate predators. With these restrictions, we found 30 relevant papers (compared to 13 papers in Takimoto & Post, 2013), many of which tested more than one potential driver of FCL. The resulting dataset (available at Dryad: https://doi.org/10.5061/dryad.jsxksn0df) consisted of 36 unique cases relating to ecosystem size, 36 unique cases relating to resource productivity, and 16 unique cases relating to disturbance. Where response values were not reported, we digitized points from graphs.

We quantify the effect of variation in each driver on FCL for each unique case using the log response ratio $LRR = \ln\left(\overline{Y}_e/\overline{Y}_c\right)$, where $\overline{Y}_e$ is the mean response under the experimental condition, and $\overline{Y}_c$ is the mean response under the control condition (Hedges et al., 1999). In each comparison, we take the smallest value of the environmental variable as the control condition while the largest value as the experimental. Thus, negative effect sizes ($LRRs$) indicate negative response of FCL to

these drivers, while positive effect sizes indicate the reverse. In particular, the *LRR*, as a unitless measure, is an appropriate effect size for cross-study comparisons (Hedges et al., 1999). Additionally, we record the levels (X_e) of these variables corresponding to each response value (Y_e) in each case, so that we can analyze the change in FCL along environmental gradients using linear regressions instead of solely by the deviation from *LRR*=0 (Figure 1). Both simple and multiple regressions are applied to separately test the individual and interactive effects of these variables, which are logarithmically transformed ($logX_e$) for normality (Supporting Information S1; R code available at Zenodo: https://doi.org/10.5281/zenodo.8297624). Note that there are 46 cases available for multiple regression analysis, which were not tested in Takimoto and Post (2013).

Theoretical framework for complex trophic systems

In the context of aquatic ecosystems (illustrated in Figure 2), we consider a well-mixed system with size S, representing the proportion of habitat sites available for species colonization. We assume that each habitat site can accommodate one individual for each species, and species can disperse randomly across the entire system. Thus, the population size of a species is given by the fraction of habitat sites it occupies (site occupancy).

We characterize the site-occupancy dynamics for basal species by incorporating spatial competition and multiple drivers. Following previous models (Tilman, 1994), we assume that basal species cannot coexist within a habitat site, thus competition can occur only through displacement of a resident species by a superior competitor

(*competitive displacement*). Furthermore, increasing resource productivity is assumed to enhance the colonization rate of all basal species by scaling them with a unitless factor R, similar to previous models (Kondoh, 2001; Worm et al., 2002). Therefore, we model the site-occupancy dynamics of basal species i (P_i – the site occupancy of basal species i) as

$$\frac{dP_i}{dt} = \underbrace{c_i^P P_i R\left(S - \sum_{j=1}^{n_P} P_j\right)}_{Colonization} \underbrace{- e_i^P P_i}_{Mortality} + R\underbrace{\sum_{j=1}^{n_P}\left(c_i^P P_i H_{ij} P_j - c_j^P P_j H_{ji} P_i\right)}_{Competitive\ displacement} + \underbrace{P_i f(t,D,T)}_{Disturbance} - \underbrace{P_i \sum_{k=1}^{n_A} \theta_{ik}\mu_{ik}A_k}_{Predation}, \quad (1)$$

where all parameters are defined in Table 1.

Most of the mathematical terms used in Equation (1) are standard components of metapopulation models and are widely used, and explained, in the existing literature (Tilman, 1994; Li et al., 2020; Liao et al., 2022). We give only a brief intuitive explanation of these terms here to aid understanding of the model. The *colonization* term describes an increase in the population of a basal species by colonizing unoccupied sites. The *mortality* term describes the intrinsic loss of population, while the *predation* term describes population losses due to predation by consumers. The *competitive displacement* term describes population changes due to colonizing a site occupied by a weaker competitor or being displaced from a site by a stronger competitor (Li et al., 2020; Liao et al., 2022). In particular, coefficients of relative competition strength H_{ij} and H_{ji} are the independent probabilities that an individual of species i displaces species j and that an individual of species j displaces species i, respectively. These coefficients can be used to describe various competition structures, for example, a strict hierarchical competition by setting $H_{ij} = 1$ if $i < j$ and 0

otherwise in a matrix $\boldsymbol{H}$ (Tilman, 1994), and intransitive competition by perturbing the hierarchical competition matrix $\boldsymbol{H}$ (Rojas-Echenique & Allesina, 2011).

The *disturbance* term, including a forcing function $f(t,D,T)$, requires some additional explanation. The disturbance regime is characterized by both disturbance extent (D) and frequency $1/T$, i.e., a given fraction D of each basal species is removed within every period T. This can be conceived of as a sudden reduction in species' site occupancies occurring periodically (i.e., pulse disturbance). Other forms of disturbance, e.g., alternative shapes or aperiodicity, are also possible. In principle, all these disturbances are stochastic. As each term in Equation (1) contains a factor of P_i, the *per-capita* growth rate ($\frac{1}{P_i}\cdot\frac{dP_i}{dt}$) of basal species *i* is independent of P_i, meaning that it is a linear and additive model. The time-averaged behavior of such model matches the long-term dynamics of the original (Liao et al., 2022). Consequently, we can replace this stochastic disturbance with its time average for the long-term dynamics (cf. Liao et al., 2022), by setting this average as

$$\overline{f}(t,D,T) = log(1-D)/T. \quad (2)$$

Our results hold for any specific form of $f(t,D,T)$ satisfying this criterion. Liao et al. (2022) have shown that the effects of a disturbance with extent *D* and periodicity *T* are equivalent to the effects of another disturbance with extent $D' = 1 - (1-D)^{1/T}$ and periodicity T'=1. Thus, we vary *D* alone while keeping *T*=1 throughout, which is sufficient for achieving a full understanding of the impact of disturbance.

Then we construct the site-occupancy dynamics for consumers in complex food webs. For simplicity, we assume: (i) these consumers can co-occur in the same habitat

site by ignoring competition among them; (ii) a consumer species has the same colonization rate when feeding on different prey species; and (iii) environmental disturbances (i.e., hydrological or geomorphologic variations) have no direct effect on consumers in aquatic ecosystems (Townsend et al., 1998; Death, 2002; McHugh et al., 2010), as most consumers can react quickly to these disturbances and escape from them via movement/dispersal, sheltering, or similar activities. Thus, we write the site-occupancy dynamics for consumer *i* as

$$\frac{dA_i}{dt} = \underbrace{c_i^A A_i\left(\sum_{j=1}^{n_P}\theta_{ji}P_j + \sum_{k=1}^{n_A}\delta_{ki}A_k\right)(S - A_i)}_{Colonization} \underbrace{- e_i^A A_i}_{Mortality} - \underbrace{A_i\sum_{k=1}^{n_A}\delta_{ik}\varphi_{ik}A_k}_{Predation}, \quad (3)$$

where parameters are defined in Table 1. The interpretation of these terms is similar to that for the basal species, with the additional emphasis on the *colonization* term where species need to feed on other consumers or/and basal species for reproduction. Note that top predators do not suffer from the top-down control, thus the site-occupancy dynamics lack the *predation* term present in Equation (3).

Model analysis of environmental drivers for FCL

Besides *stability analysis* (i.e., whether the feasible equilibrium point is stable; Supporting Information S2), we primarily use numerical methods to simulate the long-term dynamics of the model based on the C-C tradeoff among basal species (Matlab code available at Zenodo: https://doi.org/10.5281/zenodo.8297624). To establish the possibility of C-C tradeoffs, we rank basal species according to their colonization rates, so that species 1 has the lowest colonization rate and species n_P has the highest, i.e., $c_1^P < c_2^P < ... < c_{n_P}^P$. Then we assume a competitive hierarchy by ranking the basal species from the best competitor (species 1) to the poorest (species

n_P), such that colonization rate is negatively correlated with competition ability (Tilman, 1994). With the model, we explore the individual and interactive effects of these drivers on FCL in several complex food webs. Then we offer a mechanistic explanation for the resulting FCL responses by illustrating the changes in basal species diversity and therefore food web structure along environmental gradients. Finally, we demonstrate the robustness of our theoretical outcomes to varying food web complexity.

We use the niche model of Williams and Martinez (2000) to generate various initial food webs (with $n_P \geq 3$), as it can provide an accurate overall fit to the empirical structure of complex food webs. Note that those food webs with loops or cannibalism are excluded. The niche model only requires two input parameters: total number of species N and directed connectance C, which are sampled from truncated normal distributions (cf. Digel et al., 2011). Specifically, the range of initial species richness is $N \in [10,50]$, with a mean $\overline{N} = 30$ and a standard deviation SD=10. The initial connectance is sampled from $C \in [0.05,0.25]$, with a mean $\overline{C}$=0.15 and SD=0.05. These values fall within the range found in natural communities (Digel et al., 2011), so that a variety of plausible food webs can be explored.

To find the steady state, each case is run for a long time. Based on numerous preliminary trials, 50,000 time units are sufficient for all cases to achieve steady state. Accordingly, we run each case for 55,000 time units, using the time-averaged site occupancies during the final 5,000 time units to estimate species abundances at steady state. Species with steady-state abundance less than 10^{-6} are treated as extinct, because

such populations are typically eliminated by environmental fluctuations. We obtain species equilibrium abundances after transient dynamics, and then adopt the commonly used maximum trophic position to estimate FCL (cf. Post & Takimoto, 2007).

Results

Empirical analysis

We use a compiled empirical dataset to investigate the effects of these drivers on FCL, quantified using both *LRR* and linear regressions (Figure 1; Supporting Information S1). Generally, ecosystem size and resource productivity predominantly produce positive responses when measured by *LRR* (positive *versus* negative=30:6 for ecosystem size, and 24:12 for productivity), while responses to disturbance are mixed (7:9). Furthermore, we find some differences in the median effect size of these variables (Figure 1A-C). Specifically, effects of ecosystem size are more positive than effects of resource productivity, whereas negative effects of disturbance occur more frequently than positive ones.

For each empirical case, we perform a simple linear regression to analyze the individual effects of these variables on FCL (Figure 1D-F). We observe that the majority of experiments (29/36 cases) display positive effects of ecosystem size on FCLs (slope>0, with 17 cases being significant $P<0.05$), while only 7 cases are negative but non-significant. For resource productivity, 22 out of 36 cases show positive responses, but more than half (13 cases) are not significant ($P>0.05$). In addition, there are 14 negative cases for productivity, but nearly all of them are not

significant. Similarly, effects of disturbance on FCL are not significant in most cases (14 out of 16 cases: negative *versus* positive = 7:7). Thus, the linear analysis suggests that increasing ecosystem size leads to a significantly monotonic, linear increase in FCL; but that responses of FCL to productivity or disturbance may be quite complex, including non-linear, oscillatory behaviors.

We substantiate this inference with multiple regression analysis of the additive and interactive effects of these drivers (Figure 1G-L), observing that the majority of these effects are not significant (Supporting Information S1). Of the few studies with significant effects, positive effects tend to outnumber negative ones. Thus, there is no strong evidence that positive or negative correspondences are the rule. Instead, we suggest that these relationships are typically non-linear, and thus performing linear analysis produces inconclusive results.

Numerical analysis

We implement a basic numerical simulation for several typical food webs with the C-C tradeoff among basal species (Figures 2-3; Supporting Information S3: Figures S1-S4). In these simulations, we find that food chains are elongated monotonically as ecosystem size increases, regardless of productivity or disturbance (Figures 2-3A & D). However, the effects of productivity and disturbance are more complex, depending on other environmental conditions. In large ecosystems with little disturbance, FCL oscillates as productivity increases (Figures 2-3B & E). Similarly, FCL oscillates with increasing disturbance in large, productive ecosystems (Figures 2-3C & F). In particular, as the range of basal species' colonization rates increases,

more FCL oscillations emerge along the productivity or disturbance gradient (Figure 2). However, these oscillations are reduced significantly in small ecosystems with more disturbance or low productivity (Figure 3B-C & E-F). As environments become harsher, the FCL responses eventually become monotone increasing with productivity (Figure 3B &E) and monotone decreasing with disturbance (Figure 3C & F). One particularly noteworthy observation is that the intact environments ($D = 0$ and $R = 1$) do not always guarantee the longest food chains (Figures 2-3).

To explain these observations, we ignore the effects of top-down predation and consider only how basal species diversity and their relative abundances vary along environmental gradients (Figure 4; Supporting Information S3: Figures S5-S7). Basal species diversity (measured by species richness and the inverse Simpson index) rises and falls several times along environmental gradients. The points on the environmental gradient at which a basal species enters or leaves the system are "turning points". At these points, trends in abundance reverse, with species in decline starting to increase in abundance and vice versa, forming a zig-zag pattern. Thus, it is natural that this zig-zag pattern results in oscillations in basal species diversity (see proof in Supporting Information S2). Due to the bottom-up control, variation in the basal species composition, driven by environmental change, can affect how many consumers are able to survive (Supporting Information S3: Figure S8). Thus, oscillations in basal species diversity can induce oscillations in overall community diversity and hence FCL (illustrated in Figure 3). Although basal species diversity oscillates as ecosystem size changes, this does not result in oscillations in FCL

(Figure 3). This is because ecosystem size has a more direct, constraining effect on the entire food web (Equations 1 & 3). This masks the bottom-up effect of variation in basal species diversity, resulting in a monotonic change in overall diversity and FCL.

So far, we have operated with a set of stringent assumptions: a fully competitive hierarchy, evenly spaced colonization rates, a small number of basal species ($n_P \leq 6$), and several typical food webs. However, our predicted outcomes are robust to relaxing these assumptions. In particular, we obtain qualitatively similar results under a variety of other conditions (Supporting Information S3: Figures S9-S18), including trophic systems: (i) without a strict competitive hierarchy; (ii) with irregularly spaced colonization rates; (iii) with weakened competitive hierarchy; (iv) with more basal species ($n_P = 10$); and (v) with 100 initial food webs generated by the niche model (see *Methods*). To demonstrate the generality of our outcomes, we also consider a completely different competitive structure: intransitive competition (Supporting Information S4). Under this competitive structure, we do not impose a global C-C tradeoff among basal species; instead, local C-C tradeoffs, involving only a subset of the basal species, are created at random. However, we again obtain qualitatively similar results.

Finally, we use the 100 initial food webs above to produce 100 simulation cases in a wide range of parameter settings, to mimic the empirically-available data (Figure 5). When measured by *LRR* (Figure 5A-C), positive FCL responses to ecosystem size predominate (positive *versus* negative=90:8), while positive and negative responses to productivity (56:41) or disturbance (41:56) occur with similar frequency. In addition,

we observe some difference in the median effect size of these variables, i.e., ecosystem size (median≈0.44) has more positive effects than productivity (median≈0.03), while negative effects of disturbance occur more frequently than positive ones (median≈-0.021). These predictions are generally consistent with the empirical pattern in Figure 1A-C.

Using a simple linear regression on FCL against each of these variables (Figure 5D-F), we find that almost all simulation cases (97/100 cases) show positive responses to ecosystem size (slope>0), with nearly half being significant (P<0.05; Figure 5D). For resource productivity, nearly all cases show non-significant responses, with positive *versus* negative=52:47 (Figure 5E). Similarly, effects of disturbance on FCL are not significant in solid majority (88 out of 100 cases: negative *versus* positive=49:39; Figure 5F). Likewise, these FCL patterns in response to each variable accord with the empirical analysis in Figure 1D-F.

We further undertake multiple regression analysis of the additive and interactive effects of these drivers (Figure 5G-L). We again find that the majority of these effects are not significant. Yet, a few significant effects are found, with positive effects generally outnumbering negative ones. Thus, this analysis produces similar patterns to those observed in our empirical analysis (Figure 1G-L), supporting the conclusion that there is no overall tendency towards either positive or negative relationships between these drivers and FCL.

Discussion

It is well-established that environmental drivers can influence the complexity of trophic structures and particularly FCL. Existing hypotheses describing these effects share a common feature: the response to the driver is monotonic (Post, 2002a). Yet, how interactions between drivers modify these responses has been rarely considered.

While our empirical meta-analysis is broadly supportive of the *ecosystem-size* hypothesis, it suggests that responses to productivity and disturbance are oscillatory rather than monotonic. Furthermore, we find that the majority of additive and interactive effects of these drivers are not significant, weakening support for the simple *productive-space* hypothesis. The absence of significant linear interactions between drivers suggests that any such interactions are likely to be non-linear and, in particular, oscillatory. These observations necessitate a paradigmatic shift in the debate about relationships between these drivers and FCL; which until now has focused on linear responses (Post, 2002a; Sabo et al., 2010).

We can draw these conclusions because our dataset includes more aquatic empirical studies than before (Takimoto & Post, 2013), and because we use linear regressions in addition to effect size to evaluate responses. However, detection of oscillatory effects from empirical data ideally requires high resolution data to permit fitting using oscillating functions. Due to this limitation, we develop a unified C-C model to further investigate these responses and the interactions between drivers. In contrast to previous mechanistic models (Kondoh & Ninomiya, 2009; Takimoto et al., 2012; Ward & McCann, 2017), we focus on spatial competition between basal species, and use complex food web structures rather than simple trophic modules. Interestingly, simulations obtained from this model are in close agreement with our empirical

meta-analysis. Furthermore, these modelling outcomes are relatively generic, requiring only the classic assumption of C-C tradeoffs among basal species.

The mechanism by which oscillatory responses emerge in our model is variation in basal species diversity along environmental gradients due to the C-C tradeoff. Specifically, if a strong competitor is present at high abundance in the system, it will suppress the abundance of all weaker competitors. However, the species directly below it in the competitive ranking will be suppressed most, as it gains the least compensation for its competitive inferiority from its advantage in colonization rate. In turn, this benefits the species one step down the competitive ranking. As environmental conditions change to favor this competitor less, its abundance, and that of the other basal species its dominance favors, decline while those species that were suppressed increase in abundance. Eventually, this process results in the extinction of the original strong competitor and the emergence of a new dominant competitor. This shapes an alternating pattern of abundance peaks along this ranking (Figure 4G-L) and oscillations in basal species diversity along the environmental gradient. As each basal species is required by different predators, variation in the basal species composition can reassemble the overall food web. Therefore, the interaction between C-C tradeoffs and environmental changes, which facilitates different subsets of basal species to coexist, creates oscillating patterns in FCL via bottom-up control. Note that oscillations in FCL can only be seen if the potential food web is sufficiently complex, while a community structure consisting of simple trophic modules of similar size would produce little variation in FCL (Pimm & Lawton, 1977; Holt, 1996).

Consequently, we do not see this phenomenon in harsh environments (e.g., with low productivity and/or high disturbance in Figure 3), where the potential food web is relatively simple. Furthermore, it does not occur in response to ecosystem size, as all species in the food web are directly constrained by ecosystem size, and such constrained effects overwhelm the effect of variation in the basal species composition.

Previous empirical (Bengtsson, 1991; Mackey & Currie, 2001; Fraser et al., 2015) and theoretical (Hastings, 1980; Nee & May, 1992; Banitz et al., 2008; Liao et al., 2022) studies have found variation in biodiversity along environmental gradients. There are numerous empirical examples of hump-shaped biodiversity responses to disturbance and productivity (reviews in Mackey & Currie, 2001; Smith, 2007; Fraser et al., 2015). Furthermore, some experiments have actually observed multiple distinguishable peaks in biodiversity along disturbance gradients in aquatic ecosystems (Lenz et al., 2004; Cadotte, 2007; Hall et al., 2012; Gibbons et al., 2016). These observations demonstrate that the underlying variation in basal species diversity on which our mechanism for oscillations in FCL relies do occur in nature. However, this variation has not previously been linked to the effect of environmental drivers on FCL.

Effects of interactions between environmental drivers appear in both empirical and theoretical analyses. In particular, our results suggest that the effects of combining changes in productivity or disturbance with changes in ecosystem size are highly unpredictable without additional knowledge. For example, while increasing resource productivity will intuitively increase FCL, it can limit FCL in small

ecosystems by strengthening competition effects. However, this effect can be ameliorated, in turn, by greater habitat heterogeneity in larger ecosystems (Takimoto et al., 2012). Similarly, we might intuitively expect disturbance processes to limit FCL by reducing population levels, an effect that depends on the spatial scale of disturbance (Takimoto et al., 2008). However, if disturbance disrupts dominance of basal species, and therefore increases biodiversity (Power et al., 1996), it could have a positive effect on FCL. In either case, the scale of spatial disturbance, relative to the ecosystem size, determines how significant its impact will be (Sabo et al., 2010). Previous models on interactions between drivers (Takimoto et al., 2012; Terui & Nishijima, 2019) have considered only relatively simple trophic structures, while our systematic analysis highlights the importance of these interactions and suggests the need for further work.

Predators might mediate such oscillatory responses in FCL via top-down control. For instance, in a complex food web, each basal species is controlled by different predators and therefore displays different predation rates. The variation in predation rates would perturb the C-C tradeoff among basal species and alter their coexistence, thereby modifying the resulting FCL. In a simple food chain, Liao et al. (2017) found asymmetric top-down control mechanism that can cause oscillations in species abundance along habitat loss gradients, similar to our predicted oscillatory responses in basal species. Thus, we could expect that this mechanism might also induce oscillations in FCL along environmental gradients in complex food webs (e.g., addition of omnivory structure). In fact, preliminary models for a simple IGP module

have shown that omnivory responding to environmental gradients predicts context-dependency in drivers of FCL (Takimoto et al., 2012; Ward & McCann, 2017). Therefore, this asymmetric top-down control might provide an alternative mechanism to explain natural variation in FCL.

This study demonstrates how environmental drivers can act on FCL through bottom-up control, i.e., by inducing variation in basal species diversity. This complements the work of Takimoto et al. (2012) and Ward & McCann (2017) that showed how top-down control can drive changes in FCL. Future research could integrate both horizontal competitive and vertical trophic interactions into novel mechanistic models to further develop understanding of FCL responses to multiple drivers. Similarly, further empirical data is needed to confirm the existence of oscillatory responses to productivity and disturbance, and to more completely characterize interactions between these drivers. To investigate these issues, aquatic microcosm experiments are perhaps the most effective experimental system, as they have the key advantage that rapid microbial reproduction allows multigenerational community dynamics to be studied within short time frames (Gibbons et al., 2016). Additional comparative analysis of FCL along natural environmental gradients in aquatic ecosystems is also warranted. For example, the stable isotopes of nitrogen ($\delta^{15}N$) and carbon ($\delta^{13}C$) permit estimation of species' trophic positions in complex food webs (Townsend et al., 1998; Post, 2002b). Overall, our analysis suggests that the determinants of FCL are much more complex than previously thought, enriching

our understanding of the relationships between multiple drivers and FCL in complex trophic systems.

Acknowledgments

Funding: National Natural Science Foundation of China 32271548 & 31901175 (JL)

Competing interests: The author declares no competing interests.

Supplementary Information includes:

Supporting Information S1 – *Statistical analysis*

Supporting Information S2 – *System analysis*

Supporting Information S3 – *Figures S1-S18*

Supporting Information S4 – *Intransitive competition*

References

Arim, M., Borthagaray, A.I. & Giacomini, H.C. (2016) Energetic constraints to food chain length in a metacommunity framework. *Canadian Journal of Fisheries and Aquatic Sciences*, 73, 685–692.

Banitz, T., Huth, A., Grimm, V. & Johst, K. (2008) Clumped versus scattered: How does the spatial correlation of disturbance events affect biodiversity? *Theoretical Ecology*, 1, 231–240.

Bengtsson, J. (1991) Interspecific competition in meta-populations. *Biological Journal of the Linnean Society*, 42, 219–237.

Cadotte, M.W. (2007) Competition–colonization trade-offs and disturbance effects at multiple scales. *Ecology*, 88, 823–829.

Calcagno, V., Massol, F., Mouquet, N., Jarne, P. & David, P. (2011) Constraints on food chain length arising from regional metacommunity dynamics. *Proceedings of the Royal Society B*, 278, 3042–3049.

Callaway, J.C. & Josselyn, M.N. (1992) The introduction and spread of smooth cordgrass (*Spartina alterniflora*) in South San Francisco Bay. *Estuaries*, 15, 218–226.

Cardinale, B.J., Bennett, D.M., Nelson, C.E. & Gross, K. (2009) Does species diversity drive productivity or vice versa? A test of the multivariate productivity−diversity hypothesis in streams. *Ecology*, 90, 1227–1241.

Death, R.G. (2002) Predicting invertebrate diversity from disturbance regimes in forest streams. *Oikos*, 97, 18–30.

Digel, C., Riede, J.O. & Brose, U. (2011) Body sizes, cumulative and allometric degree distributions across natural food webs. *Oikos*, 120, 503–509.

Doi, H., et al. (2009) Resource availability and ecosystem size predict food-chain length in pond ecosystems. *Oikos*, 118, 138–144.

Fraser, L.H., et al. (2015) Worldwide evidence of a unimodal relationship between productivity and plant species richness. *Science*, 349, 302–305.

Gibbons, S.M., Scholz, M., Hutchison, A.L., Dinner, A.R., Gilbert, J.A. & Coleman, M.L. (2016) Disturbance regimes predictably alter diversity in an ecologically complex bacterial system. *mBio*, 7, e01372–16.

Hairston, N.G., Jr & Hairston, N.G., Sr (1993) Cause-effect relationships in energy flow trophic structure and interspecific interactions. *American Naturalist*, 142, 379–411.

Hall, A.R., Miller, A.D., Leggett, H.C., Roxburgh, S.H., Buckling, A. & Shea, K. (2012) Diversity-disturbance relationships: frequency and intensity interact. *Biology Letters*, 8, 768–771.

Hastings, A. (1980) Disturbance, coexistence, history, and competition for space. *Theoretical Population Biology*, 18, 363–373.

Häussler, J., Barabás, G. & Eklöf, A. (2020) A Bayesian network approach to trophic metacommunities shows that habitat loss accelerates top species extinctions. *Ecology Letters*, 23, 1849–1861.

Hedges, L.V., Gurevitch, J. & Curtis, P.S. (1999) The meta-analysis of response ratios in experimental ecology. *Ecology*, 80, 1150–1156.

Holt, R.D. (1996) Food webs in space: An island biogeographic perspective. In: *Food webs: Integration of patterns & dynamics* (eds Polis, G.A. & Winemiller, K.O.) pp. 313–323. Boston, MA: Springer US.

Huisman, J., Jonker, R.R., Zonneveld, C. & Weissing, F.J. (1999) Competition for light between phytoplankton species: experimental tests of mechanistic theory. *Ecology*, 80, 211–222.

Kidd, K.A., Schindler, D.W., Muir, D.C.G., Lockhart, W.L., Hesslein, R.H. (1995)

High concentrations of toxaphene in fishes from a subarctic lake. *Science*, 269, 240–242.

Kondoh, M. (2001) Unifying the relationships of species richness to productivity and disturbance. *Proceedings of the Royal Society B*, 268, 269–271.

Kondoh, M. & Ninomiya, K. (2009) Food-chain length and adaptive foraging. *Proceedings of the Royal Society B*, 276, 3113–3121.

Lenz, M., Molis, M. & Wahl. M. (2004) Testing the intermediate disturbance hypothesis: Response of fouling communities to various levels of emersion intensity. *Marine Ecology-Progress Series*, 278, 53–65.

Li, Y., Bearup, D. & Liao, J. (2020) Habitat loss alters effects of intransitive higher-order competition on biodiversity: a new metapopulation framework. *Proceedings of the Royal Society B*, 287, 20201571.

Liao, J., Bearup, D. & Blasius, B. (2017) Diverse responses of species to landscape fragmentation in a simple food chain. *Journal of Animal Ecology*, 86, 1169–1178.

Liao, J., Barabás, G. & Bearup, D. (2022) Competition–colonization dynamics and multimodality in diversity–disturbance relationships. *Ecology*, 103, e3672.

Mackey, R.L. & Currie, D.J. (2001) The diversity–disturbance relationship: Is it generally strong and peaked? *Ecology*, 82, 3479–3492.

McHugh, P.A., McIntosh, A.R. & Jellyman, P.G. (2010) Dual influences of ecosystem size and disturbance on food-chain length in streams. *Ecology Letters*, 13, 881–890.

McIntyre, P.B., Jones, L.E., Flecker, A.S. & Vanni, M.J. (2007) Fish extinctions alter nutrient recycling in tropical freshwaters. *Proceedings of the National Academy of Sciences USA*, 104, 4461.

Nee, S. & May, R.M. (1992) Dynamics of metapopulations: habitat destruction and competitive coexistence. *Journal of Animal Ecology*, 61, 37–40.

Pace, M.L., Cole, J.J., Carpenter, S.R. & Kitchell, J.F. (1999) Trophic cascades revealed in diverse ecosystems. *Trends in Ecology & Evolution*, 14, 483–488.

Pimm, S.L. & Lawton, J.H. (1977) Number of trophic levels in ecological communities. *Nature*, 268, 329–331.

Post, D.M., Pace, M.L. & Hairston, N.G. (2000) Ecosystem size determines food chain length in lakes. *Nature*, 405, 1047–1049.

Post, D.M. (2002a) The long and short of food-chain length. *Trends in Ecology & Evolution*, 17, 269–277.

Post, D.M. (2002b) Using stable isotopes to estimate trophic position: models, methods, and assumptions. *Ecology*, 83, 703–718.

Post, D.M. (2007) Testing the productive-space hypothesis: rational and power. *Oecologia*, 153, 973–984.

Post, D.M. & Takimoto, G. (2007) Proximate structural mechanisms for variation in food-chain length. *Oikos*, 116, 775–782.

Power, M.E., Parker, M.S. & Wootton, J.T. (1996) Disturbance and food chain length in rivers. In: Polis, G.A. & Winemiller, K.O. (eds.) *Food webs: integration of pattern and dynamics*. Chapman and Hall, New York, pp. 286–297.

Rojas-Echenique, J. & Allesina, S. (2011) Interaction rules affect species coexistence in intransitive networks. *Ecology*, 92, 1174–1180.

Sabo, J.L., Finlay, J.C., Kennedy, T. & Post, D.M. (2010) The role of discharge variation in scaling of drainage area and food-chain length in rivers. *Science*, 330, 965–967.

Schindler, D.E., Carpenter, S.R., Cole, J.J., Kitchell, J.F. & Pace, M.L. (1997) Influence of food web structure on carbon exchange between lakes and the atmosphere. *Science*, 277, 248–251.

Smith, V.H. (2007) Microbial diversity–productivity relationships in aquatic ecosystems. *FEMS Microbiology Ecology*, 62, 181–186.

Spencer, M. & Warren, P.H. (1996) The effects of habitat size and productivity on food web structure in small aquatic microcosms. *Oikos*, 75, 419–430.

Sun, W.H., Yu, Z.W. & Yu, S.W. (1988) Inhibitory effect of *Eichhornia crassipes* (Mart.) Solms on algae. *Acta Phytophysiologica Sinica*, 14, 294–300.

Takimoto, G., Spiller, D.A. & Post, D.M. (2008) Ecosystem size, but not disturbance, determines food-chain length on islands of the Bahamas. *Ecology*, 89, 3001–3007.

Takimoto, G., Post, D.M., Spiller, D.A. & Holt, R.D. (2012) Effects of productivity, disturbance, and ecosystem size on food-chain length: insights from a metacommunity model of intraguild predation. *Ecological Research*, 27, 481–493.

Takimoto, G. & Post, D.M. (2013) Environmental determinants of food-chain length: a meta-analysis. *Ecological Research*, 28, 675–681.

Terui, A. & Nishijima, S. (2019) Spatial disturbance synchrony alters the association of food chain length and ecosystem size. *Ecological Research*, 34, 864–871.

Tilman, D. (1994) Competition and biodiversity in spatially structured habitats. *Ecology*, 75, 2–16.

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. *Ecology Letters*, 1, 200–209.

Vander Zanden, M.J. & Rasmussen, J.B. (1999) Primary consumer δ^{13}C and δ^{15}N and the trophic position of aquatic consumers. *Ecology*, 80, 1395–1404.

Ward, C.L. & McCann, K.S. (2017) A mechanistic theory for aquatic food chain length. *Nature Communications*, 8, 2028.

Warfe, D.M., et al. (2013) Productivity, disturbance and ecosystem size have no influence on food chain length in seasonally connected rivers. *PLoS One*, 8, e66240.

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

Worm, B., Lotze, H., Hillebrand, H. & Sommer, U. (2002) Consumer versus resource control of species diversity and ecosystem functioning. *Nature*, 417, 848–851.

Young, H.S., et al. (2013) The roles of productivity and ecosystem size in determining food chain length in tropical terrestrial ecosystems. *Ecology*, 94, 692–701.

Tables

Table 1. Definitions of variables and parameters

Symbols	Definitions
R	Resource productivity, i.e., resource availability per-unit-ecosystem-size
S	The proportion of habitat sites that are available for species colonization in the entire ecosystem, i.e., ecosystem size
D	The fraction of individuals in each basal species being removed within every period, i.e., disturbance extent
P_i	The fraction of habitat sites that are inhabited by basal species *i*, i.e., its site occupancy
A_i	The fraction of habitat sites that are inhabited by consumer species *i*, i.e., its site occupancy
n_P	The number of basal species
n_A	The number of consumer species, including top predators
c_i^P	The colonization rate of basal species *i* (per unit time)
c_i^A	The colonization rate of consumer species *i* (per unit time)
e_i^P	The mortality rate of basal species *i* (per unit time)
e_i^A	The mortality rate of consumer species *i* (per unit time)
H_{ij}	The competition strength of basal species *i* relative to basal species *j*
θ_{ij}	The elements in a adjacency matrix for the trophic interaction between basal and consumer species, with $\theta_{ij} = 1$ if consumer *j* feeds on basal species *i* (otherwise $\theta_{ij} = 0$)
δ_{ij}	The elements in a adjacency matrix for the trophic interaction between consumers, with $\delta_{ij} = 1$ if consumer *j* feeds on another consumer *i* (otherwise $\delta_{ij} = 0$)
μ_{ij}	The top-down mortality rate of basal species *i* due to predation by consumer *j* (per unit time)
φ_{ij}	The top-down mortality rate of consumer *i* due to predation by another consumer *j* (per unit time)
N	Total number of species in the food web ($N = n_P + n_A$)
C	Food web connectance

Figure legends

Figure 1. Meta-analysis on FCL responses to multiple environmental variables (ecosystem size *S*, resource productivity *R* and disturbance *D*) in an empirical dataset compiled from aquatic ecosystems, using both (A-C) the log response ratio (*LRR*) and (D-L) linear regressions. Panels (A-C): effect sizes (*LRR* – grey circle) of *S*, *R* and *D* in these empirical cases, summarized by box-plots. Panels (D-L): both simple and multiple linear regressions on the log-transformed predictors are used to respectively test the individual (D-F), additive (G-I) and interactive effects (J-L) of these variables, with the regression coefficients being summarized by box-plots (blue circles – significant effects at $P<0.05$; yellow circles – non-significant effects).

Figure 2. Interactive effects of ecosystem size (*S*), resource productivity (*R*) and disturbance extent (*D*) on food chain length (FCL) in a typical food web (with species diversity $N=20$, connectance $C=0.15$ and basal species richness $n_P=4$; red circles – species, black lines – trophic links, and dotted lines – competition between basal species), extracted from a lake ecosystem. The basal species are ranked from the best competitor (species 1) to the poorest (species n_P) in a strict competitive hierarchy, i.e., $H_{ij}=1$ for $i<j$ and 0 otherwise in a matrix $\boldsymbol{H}$. To establish the possibility of competition-colonization (C-C) tradeoffs, basal species' colonization rates are evenly spaced in increasing order at both small (A-C: $c_i^P \epsilon E[0.45, 0.8]$) and large (D-F: $c_i^P \epsilon E[0.25, 1]$) ranges. Other parameters: $D=0$ in panels (A & D), $S=1$ in panels (B & E), and $R=1$ in panels (C & F), all species mortality rates $e_i^P = e_i^A=0.1$, all consumers' colonization rates $c_i^A=0.625$ and all top-down mortality rates due to predation $\mu_{ik} = \varphi_{ik}=0.05$.

Figure 3. Individual effects of ecosystem size (*S*), resource productivity (*R*), and disturbance extent (*D*) on FCL in a given food web as displayed in Figure 2. Panels (A & D) $R=0.2$, 0.6 & 1 with $D=0$; panels (B & E) $S=0.6$, 0.8 & 1 with $D=0$; and panels (C & F) $R=S=0.6$, 0.8 & 1. Meanwhile, four food web structures along different environmental gradients are displayed for each panel. Other parameter settings are the same as in Figure 2.

Figure 4. Individual effects of ecosystem size (*S*), resource productivity (*R*) and disturbance extent (*D*) on basal species diversity (A-F) and their relative abundances (G-L) for initial richness n_P=4, while ignoring the top-down predation. Basal species diversity is characterized by (A-C) species richness and the inverse Simpson index ($1/\sum q_i^2$, with $q_i = P_i/\sum P_j$ being the relative abundance of basal species *i*). Panels (A, D, G & J): *R*=1 and *D*=0; panels (B, E, H & K): *S*=1 and *D*=0; and panels (C, F, I & L): *R*=*S*=1. Other parameters are the same as in Figure 2.

Figure 5. Analysis of FCL responses to multiple environmental variables in 100 simulation cases, using both (A-C) the log response ratio (*LRR*) and (D-L) linear regressions. The niche model is used to generate 100 initial food webs, excluding those with loops and cannibalism (see *Methods*). In each case, we first sample 150 values for *S*ϵ[0, 1], *R*ϵ[0, 1] and *D*ϵ[0, 0.6] respectively, and then randomly combine them into 150 groups as our model input for each initial food web. Finally, these 150 samples are used to estimate *LRR* and perform linear regressions (ignoring a large number of samples with FCL=0 & 1). In each initial food web, basal species' colonization rates are uniformly drawn from $c_i^P \epsilon$[0.25, 1] and sorted in increasing order, but with a strict competitive hierarchy. Panels (A-C): effect sizes (*LRR* – grey circle) of ecosystem size (*S*), resource productivity (*R*) and disturbance extent (*D*) in these cases, summarized by box-plots. Panels (D-L): both simple and multiple linear regressions are used to respectively test the individual (D-F), additive (G-I) and interactive effects (J-L) of these variables, with the regression coefficients being summarized by box-plots (blue circles – significant effects at *P*<0.05; yellow circles – non-significant effects). Others: $e_i^P = e_i^A$=0.1, c_i^A=0.625 and $\mu_{ik} = \varphi_{ik}$=0.05.

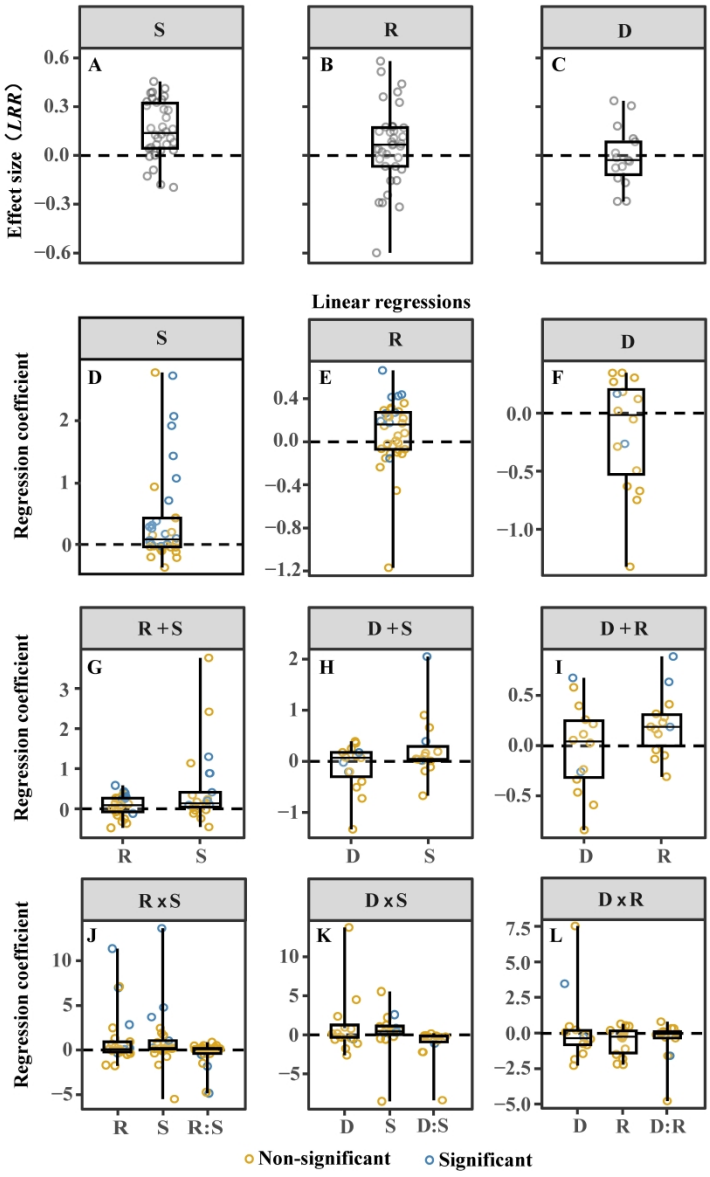


Figure 1. Meta-analysis on FCL responses to multiple environmental variables (ecosystem size S, resource productivity R and disturbance D) in an empirical dataset compiled from aquatic ecosystems, using both (A-C) the log response ratio (LRR) and (D-L) linear regressions. Panels (A-C): effect sizes (LRR – grey circle) of S, R and D in these empirical cases, summarized by box-plots. Panels (D-L): both simple and multiple linear regressions on the log-transformed predictors are used to respectively test the individual (D-F), additive (G-I) and interactive effects (J-L) of these variables, with the regression coefficients being summarized by box-plots (blue circles – significant effects at $P<0.05$; yellow circles – non-significant effects).

105x168mm (1200 x 1200 DPI)

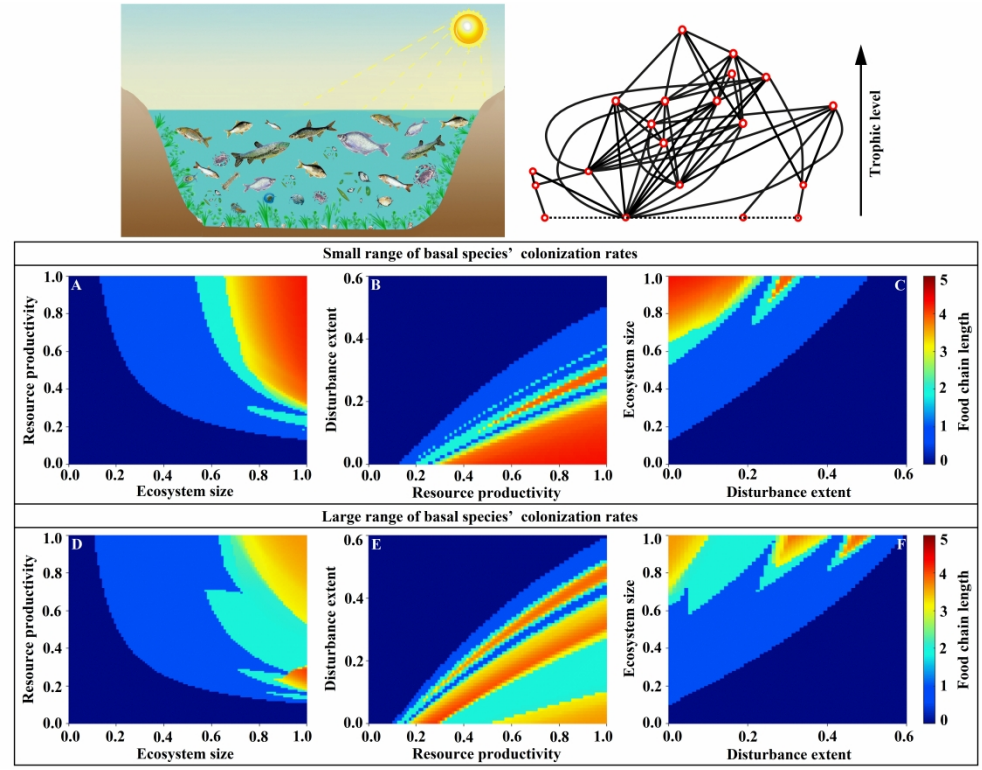


Figure 2. Interactive effects of ecosystem size (S), resource productivity (R) and disturbance extent (D) on food chain length (FCL) in a typical food web (with species diversity N=20, connectance C=0.15 and basal species richness n_P=4; red circles – species, black lines – trophic links, and dotted lines – competition between basal species), extracted from a lake ecosystem. The basal species are ranked from the best competitor (species 1) to the poorest (species n_P) in a strict competitive hierarchy, i.e., H_ij=1 for i<j and 0 otherwise in a matrix H. To establish the possibility of competition-colonization (C-C) tradeoffs, basal species' colonization rates are evenly spaced in increasing order at both small (A-C: c_i^P□E[0.45, 0.8]) and large (D-F: c_i^P□E[0.25, 1]) ranges. Other parameters: D=0 in panels (A & D), S=1 in panels (B & E), and R=1 in panels (C & F), all species mortality rates e_i^P=e_i^A=0.1, all consumers' colonization rates c_i^A=0.625 and all top-down mortality rates due to predation μ_ik=φ_ik=0.05.

272x208mm (600 x 600 DPI)

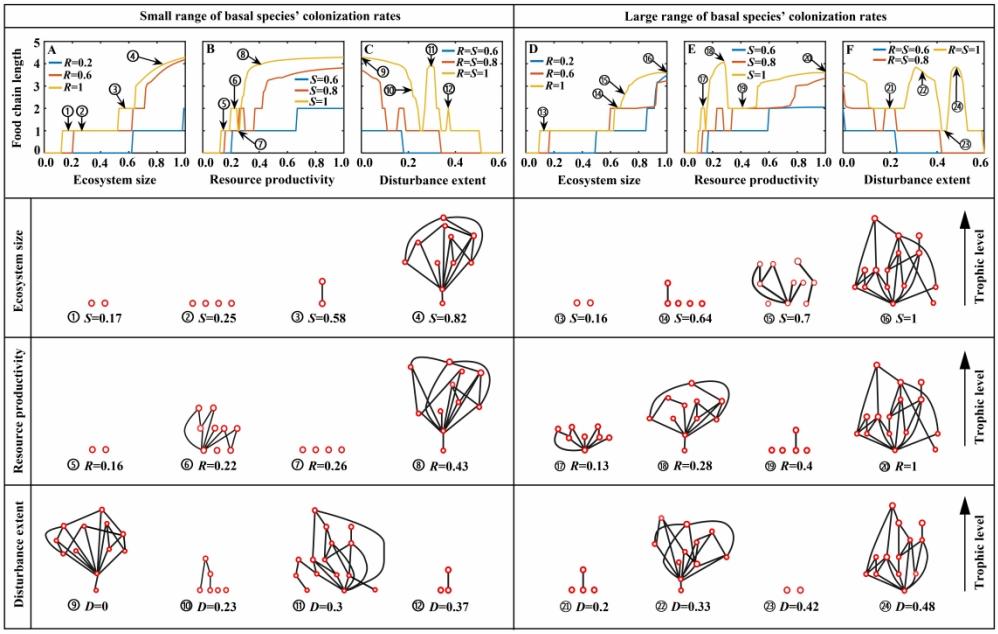


Figure 3. Individual effects of ecosystem size (S), resource productivity (R), and disturbance extent (D) on FCL in a given food web as displayed in Figure 2. Panels (A & D) R=0.2, 0.6 & 1 with D=0; panels (B & E) S=0.6, 0.8 & 1 with D=0; and panels (C & F) R=S=0.6, 0.8 & 1. Meanwhile, four food web structures along different environmental gradients are displayed for each panel. Other parameter settings are the same as in Figure 2.

272x208mm (600 x 600 DPI)

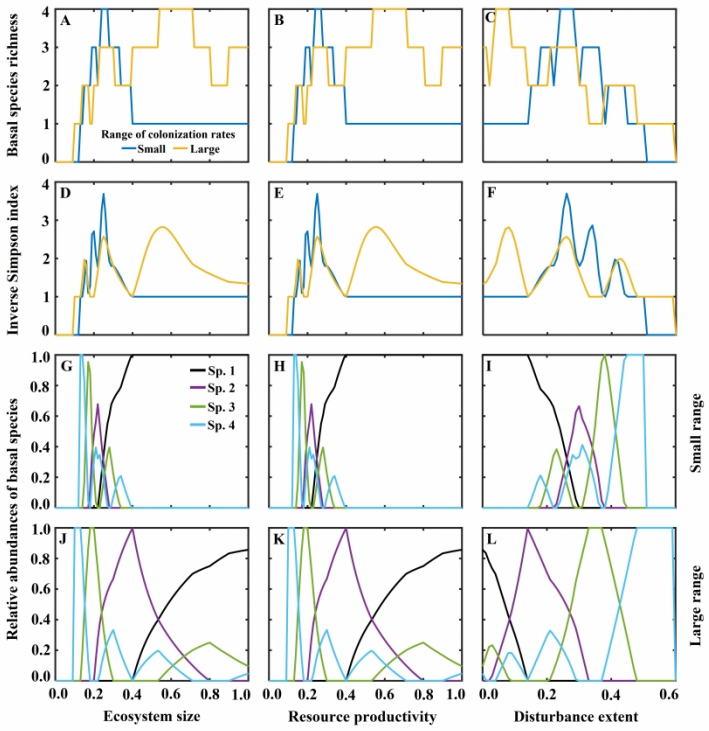


Figure 4. Individual effects of ecosystem size (S), resource productivity (R) and disturbance extent (D) on basal species diversity (A-F) and their relative abundances (G-L) for initial richness n_P=4, while ignoring the top-down predation. Basal species diversity is characterized by (A-C) species richness and the inverse Simpson index ($1/\sum q_i^2$, with $q_i=P_i/\sum P_j$ being the relative abundance of basal species i). Panels (A, D, G & J): R=1 and D=0; panels (B, E, H & K): S=1 and D=0; and panels (C, F, I & L): R=S=1. Other parameters are the same as in Figure 2.

272x208mm (600 x 600 DPI)

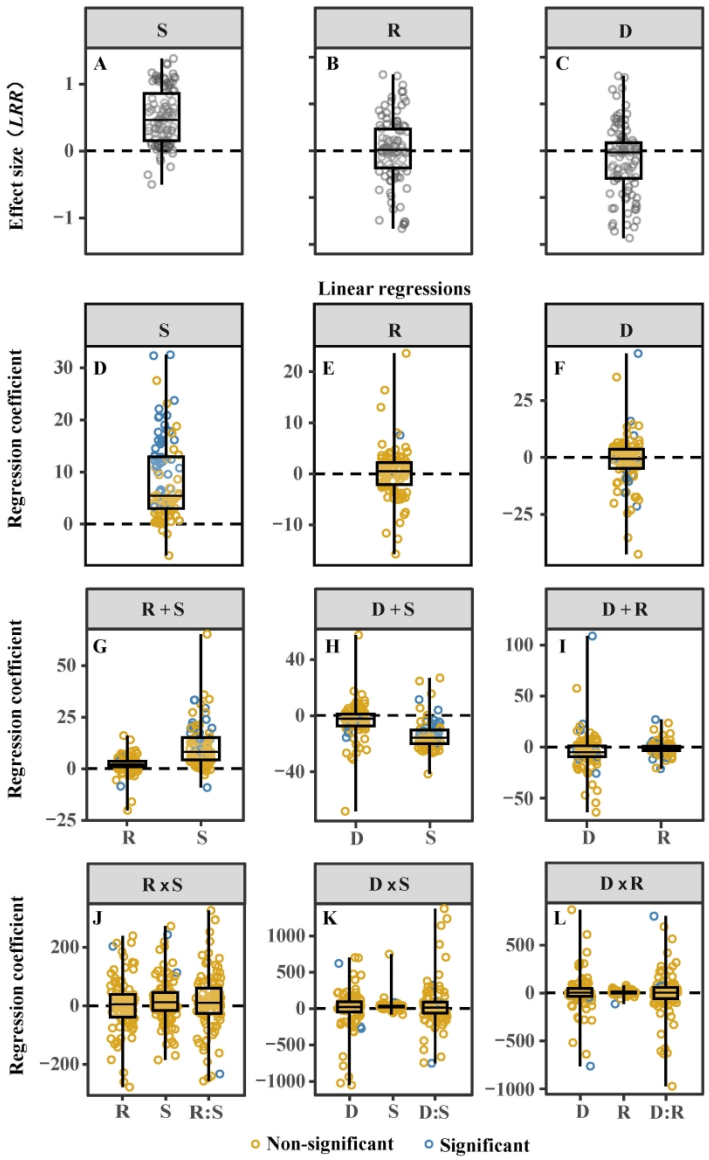
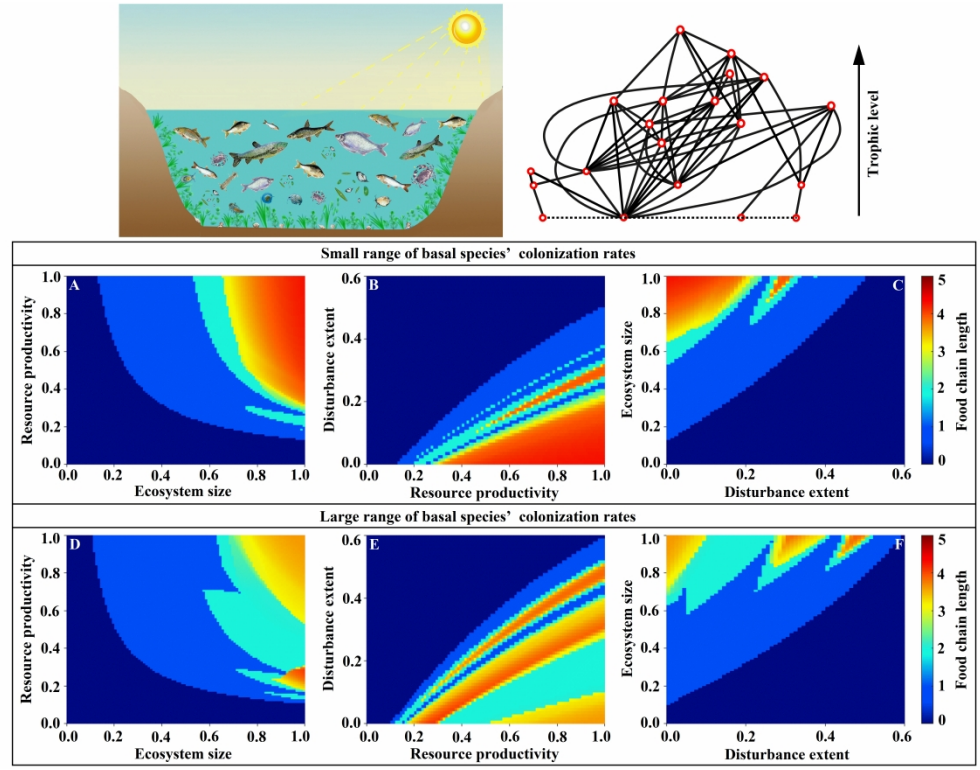


Figure 5. Analysis of FCL responses to multiple environmental variables in 100 simulation cases, using both (A-C) the log response ratio (LRR) and (D-L) linear regressions. The niche model is used to generate 100 initial food webs, excluding those with loops and cannibalism (see Methods). In each case, we first sample 150 values for S□[0, 1], R□[0, 1] and D□[0, 0.6] respectively, and then randomly combine them into 150 groups as our model input for each initial food web. Finally, these 150 samples are used to estimate LRR and perform linear regressions (ignoring a large number of samples with FCL=0 & 1). In each initial food web, basal species' colonization rates are uniformly drawn from c_i^P□[0.25, 1] and sorted in increasing order, but with a strict competitive hierarchy. Panels (A-C): effect sizes (LRR – grey circle) of ecosystem size (S), resource productivity (R) and disturbance extent (D) in these cases, summarized by box-plots. Panels (D-L): both simple and multiple linear regressions are used to respectively test the individual (D-F), additive (G-I) and interactive effects (J-L) of these variables, with the regression coefficients being summarized by box-plots (blue circles – significant effects at P<0.05; yellow circles – non-significant effects). Others: e_i^P=e_i^A=0.1, c_i^A=0.625 and μ_ik=φ_ik=0.05.

109x175mm (600 x 600 DPI)



272x208mm (600 x 600 DPI)

Supplementary Information for

Towards a mechanistic understanding of variation in aquatic food chain length revealed by meta-analysis

Guanming Guo, György Barabás, Gaku Takimoto, Daniel Bearup, William F. Fagan, Dongdong Chen, Jinbao Liao*

* To whom correspondence should be addressed. Email: jinbaoliao@163.com

This PDF file includes:

Supporting Information S1 - *Statistical analysis*

We begin by loading the empirical dataset:

```
library(tidyverse)
dat <- read_csv("Empirical_food_chain_data.csv") # Read data
knitr::kable(head(dat)) # Display first few rows of the table
```

case	reference	size	resource	dist	FCL
1	McHugh et al. 2010	0.3516870	8.6909871	9.750174	3.137930
1	McHugh et al. 2010	1.4036029	2.5107296	11.775951	2.756533
1	McHugh et al. 2010	1.1624540	1.5450644	10.767550	3.311276
1	McHugh et al. 2010	2.5901071	1.1587983	11.670968	3.170593
1	McHugh et al. 2010	0.3946072	0.5793991	9.996132	3.423433
1	McHugh et al. 2010	0.3033092	1.8669528	11.651023	3.561813

Here is a description of the columns of the data:

- `case`: The ID number of the case study from which the data were taken.
- `reference`: The reference for the original source from which the data were taken.
- `size`: Measure of ecosystem size (if available). Its units are dependent of the case study (and can be found in the original Excel files), but they are always consistent within a given study.
- `resource`: Measure of resource availability. (Same comments apply as for `size`.)
- `dist`: Disturbance magnitude. (Again, same comments apply as for `size`.)
- `FCL`: The observed food chain length.

Most studies do not contain measurements of all three predictors (ecosystem size, resource availability, and disturbance magnitude). More commonly, only two predictors are available—and, in some cases, only one. We can see this from the table below where `size`, `resource`, and `dist` now represent the availability of data in those categories:

```
dat %>%
  group_by(case, reference) %>%
  summarise(n = n(), across(c(size, resource, dist), ~!all(is.na(.x)))) 
 %>%
  ungroup() %>%
  mutate(across(c(size, resource, dist), ~if_else(.x, "yes", "no"))) %>%
  knitr::kable()
```

case	reference	n	size	resource	dist

case	reference	n	size	resource	dist
1	McHugh et al. 2010	15	yes	yes	yes
2	Ruhi et al. 2016	9	yes	yes	yes
3	He et al. 2020	8	yes	yes	yes
4	Sabo et al. 2010	35	yes	yes	yes
5	Kautza & Sullivan 2016	12	yes	yes	yes
6	Sullivan et al. 2015	7	yes	yes	yes
7	Sullivan et al. 2015	5	yes	yes	yes
8	Sullivan et al. 2015	7	yes	yes	yes
9	Chanut et al. 2020	24	yes	yes	yes
10	Chanut et al. 2020	20	yes	yes	yes
11	Chanut et al. 2020	12	yes	yes	yes
12	Wang et al. 2016	16	yes	yes	yes
13	Warfe et al. 2013	66	yes	yes	yes
14	Townsend et al. 1998	10	no	yes	yes
15	Jackson & Sullivan 2017	31	yes	no	yes
16	Schriever & Williams 2013	9	yes	no	yes
17	Ward & McCann 2017	62	yes	yes	no
18	Post et al. 2000	25	yes	yes	no
19	Zanden et al. 1999	17	yes	yes	no
20	Reid et al. 2011	10	yes	yes	no
21	Reid et al. 2011	8	yes	yes	no
22	Thompson & Townsend 2005	18	yes	yes	no
23	Doi et al. 2009	15	yes	yes	no
24	Parker & Huryn 2013	5	yes	yes	no
25	Parker & Huryn 2013	14	yes	yes	no
26	Ziegler et al. 2015	20	yes	yes	no
27	Doi et al. 2012	7	yes	yes	no
28	Doi et al. 2012	10	yes	yes	no
29	Doi et al. 2012	3	yes	yes	no
30	Zhang et al. 2013	52	yes	no	no
31	Zhang et al. 2013	68	yes	no	no
32	Zhang et al. 2013	34	yes	no	no
33	Tunney et al. 2012	40	yes	no	no
34	Zanden & Fetzer 2007	66	yes	no	no
35	Zanden & Fetzer 2007	24	yes	no	no

case	reference	n	size	resource	dist
36	Zanden & Fetzer 2007	14	yes	no	no
37	Fraley et al. 2018	28	yes	no	no
38	Doi et al. 2009	14	no	yes	no
39	Doi et al. 2009	13	no	yes	no
40	Saigo et al. 2016	8	no	yes	no
41	Williams & Trexler 2006	17	no	yes	no
42	Anderson & Cabana 2009	23	no	yes	no
43	Kelly & Schallenberg 2019	5	no	yes	no
44	Kelly & Schallenberg 2019	8	no	yes	no
45	Kelly & Schallenberg 2019	5	no	yes	no
46	Hoeinghaus et al. 2008	10	no	yes	no

The column `n` above shows the number of data points in the corresponding study. Thus, out of the 46 case studies, 13 have measurements for all three predictors. 16 have only two measured predictors, and 17 only one. Here is a summary table:

```
dat %>%
  group_by(case, reference) %>%
  summarise(n = n(), across(c(size, resource, dist), ~1*!all(is.na(.
x)))) %>%
  ungroup() %>%
  mutate(predictors = size + resource + dist) %>%
  count(predictors, name = "number of studies") %>%
  knitr::kable()
```

predictors	number of studies
1	17
2	16
3	13

This means that using statistical models which take all three predictors into account is only feasible for about one third of the cases. But a further problem with using three-way models is the scarcity of data. As the above table shows, many of the 17 cases with all three predictors measured have very few observations, rendering such multi-way models inapplicable. In lieu of three-way models, we fit all possible one- and two-way models to the data in each of these cases. The following helper functions aid in doing that:

```
zscore <- function(vec) (vec - mean(vec)) / sd(vec)

transformData <- function(dat, trans = log) {
  dat %>%
    # Transform the predictors:
```

```
    group_by(case) %>%
    mutate(size = trans(size),
           resource = trans(resource),
           dist = trans(dist)) %>%
    ungroup() %>%
    # Remove potential infinities arising from the transformation:
    filter(!is.infinite(size),
           !is.infinite(resource),
           !is.infinite(dist))
}

buildFormulas <- function(dat) {
  dat %>%
    # Combine data with all possible combinations or 1- and 2-factor mo
dels:
    crossing(formula = c("dist", "resource", "size",
                         "dist+resource", "dist+size", "resource+size",
                         "dist*resource", "dist*size", "resource*size")
 %>%
               # Put "FCL ~ " at the front of each formula:
               str_c("FCL ~ ", .)) %>%
    # Discard models for which there isn't appropriate data in `dat`:
    filter(!(is.na(dist) & str_detect(formula, "dist"))) %>%
    filter(!(is.na(resource) & str_detect(formula, "resource"))) %>%
    filter(!(is.na(size) & str_detect(formula, "size")))
}

analyzeData <- function(datWithFormula) {
  datWithFormula %>%
    # One row per each case-model combination:
    nest(data = !case & !formula) %>%
    # Fit the models and put information into tidy tables:
    mutate(fit = map2(formula, data, lm)) %>%
    mutate(model = map(fit, compose(broom::tidy, summary))) %>%
    mutate(quality = map(fit, broom::glance),
           quality = map(quality, ~select(.x, contains("r."), AIC))) %>%
    unnest(c(model, quality))
}

simplifyFormulas <- function(datWithFormula) {
  datWithFormula %>%
    # Refactor & reorder model formulas:
    mutate(formula = str_remove(formula, "FCL ~ ")) %>%
    mutate(formula = str_replace_all(formula, c("dist"="D","resource"="
R",
                                                "size"="S", "\\*"=" x ",
                                                "\\+"=" + "))) %>%
    mutate(formula = fct_relevel(formula, "S", "R", "D", "R + S", "D +
S",
```

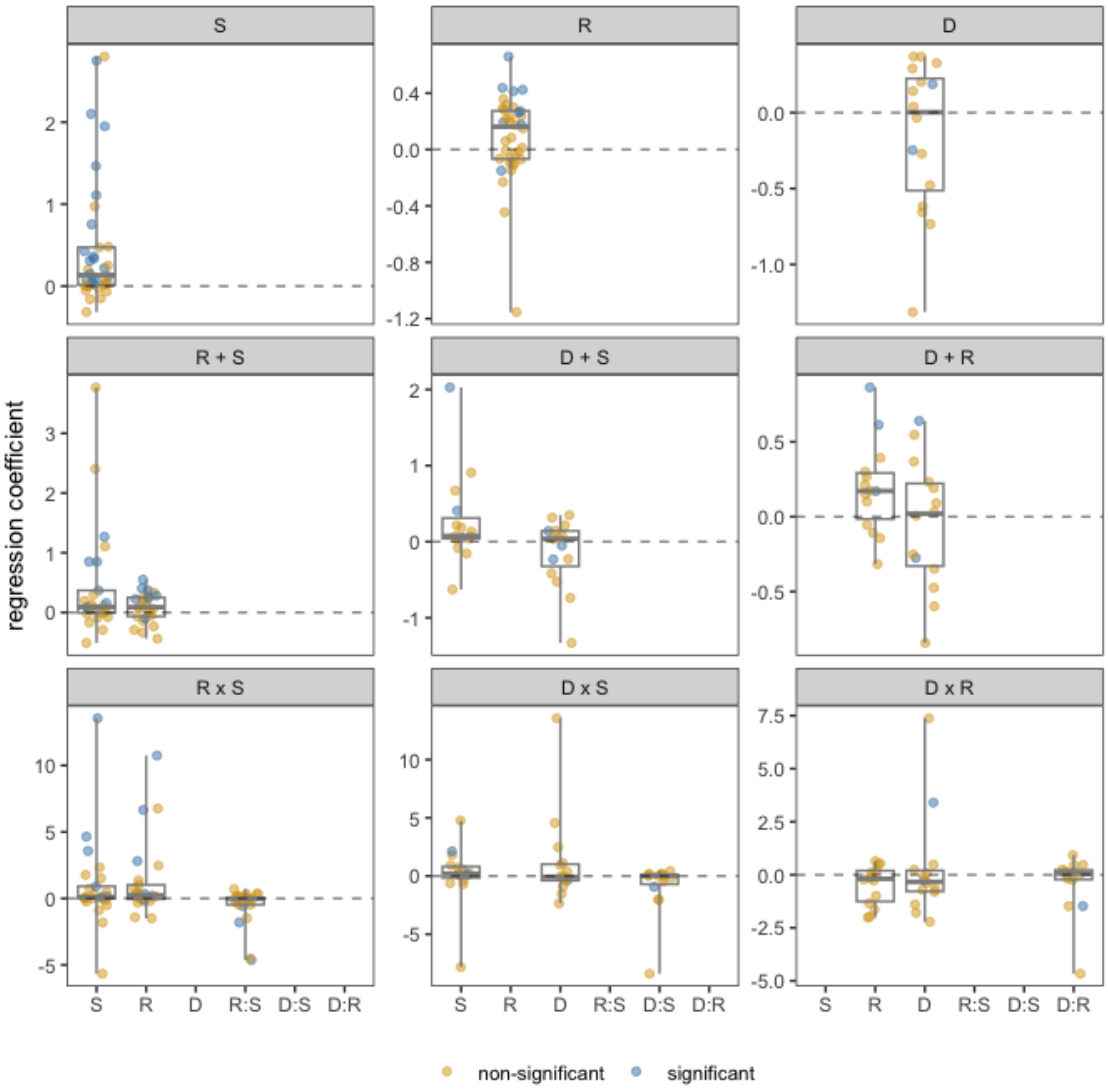
```
                                  "D + R", "R x S", "D x S", "D x R"))
 %>%
    # Refactor & reorder predictors:
    mutate(term = str_replace_all(term, c("dist" = "D","resource" = "R",
                                          "size" = "S"))) %>%
    mutate(term = fct_relevel(term, "S", "R", "D", "R:S", "D:S", "D:R"))
}

plotData <- function(datProcessed, signif = 0.05) {
  datProcessed %>%
    # Are regression coefficients significant?
    mutate(result = ifelse(p.value < signif,
                           "significant", "non-significant")) %>%
    mutate(result = fct_relevel(result,
                                "non-significant", "significant")) %>%
    # Create plot:
    ggplot(aes(x = term, y = estimate, colour = result, alpha = result))
 +
    geom_jitter(width = 0.25, alpha = 0.5) +
    geom_boxplot(aes(x = term, y = estimate), colour = "gray55", alpha
= 0,
                 outlier.shape = NA, coef = 100) +
    geom_hline(yintercept = 0, linetype = "dashed", alpha = 0.5) +
    ylab("regression coefficient") +
    scale_colour_manual(values = c("goldenrod", "steelblue")) +
    scale_alpha_manual(values = c(0.4, 0.7)) +
    facet_wrap(~formula, scales = "free_y") +
    theme_bw() +
    theme(panel.grid = element_blank(), axis.title.x = element_blank(),
          legend.position = "bottom", legend.title = element_blank())
}
```

We now fit all possible models to each case study after log-transforming the predictors, and visualize the results where two-way models were fitted:

```
datProcessed <- dat %>%
  transformData(trans = log) %>%
  buildFormulas() %>%
  analyzeData() %>%
  simplifyFormulas() %>%
  filter(term != "(Intercept)") %>%
  drop_na()

plotData(datProcessed, signif = 0.05)
```



The panels are for the different models. S means size, R resource availability, and D disturbance. When they are by themselves (fist row of panels), we are performing simple regression with a single predictor. When they are added, as in R + S (second row), we are performing two-way regression with main effects only. And when they are multiplied, as in R x S (third row), we are performing two-way regression with interaction effects also included. Within each panel, the x-axis shows which regression coefficient is displayed (the colon means interaction), while the y-axis shows the corresponding coefficient values across the case studies. The points have been jittered sideways to reduce overlap. Statistically significant coefficients at the 0.05 level are shown in blue, all other coefficients are in yellow. Box plots summarize the distribution of the coefficients, with the whiskers encompassing the full range of the data.

Let us summarize, in a table, how many of the results are significant versus non-significant, and whether significant coefficients are positive or negative:

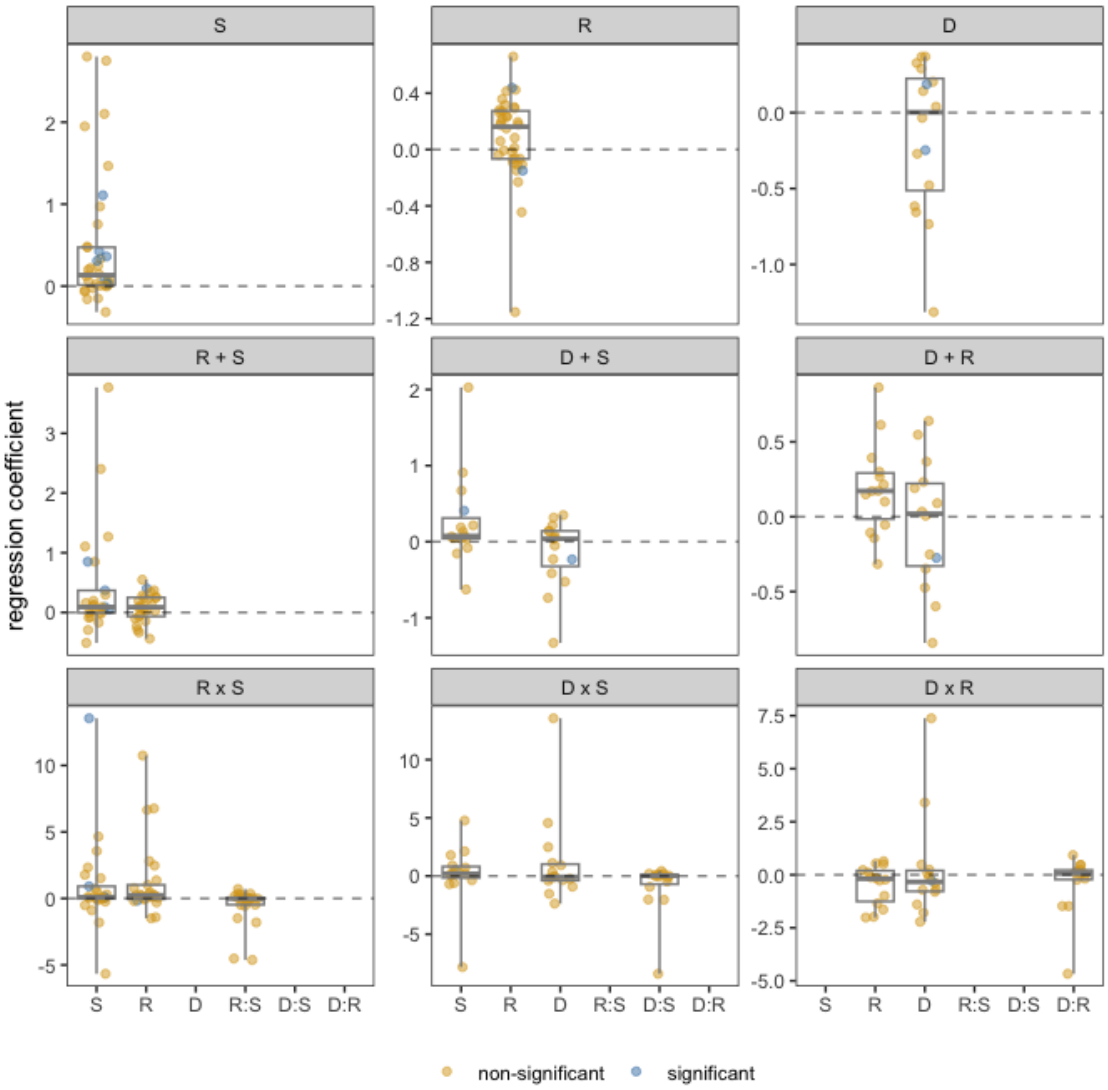
```r
datProcessed %>%
  mutate(result = case_when(
    p.value < 0.05 & estimate >  0 ~ "significant positive",
    p.value < 0.05 & estimate <= 0 ~ "significant negative",
    TRUE ~ "non-significant"
  )) %>%
  group_by(formula) %>%
  count(result) %>%
  mutate(total = sum(n)) %>%
  ungroup() %>%
  pivot_wider(names_from = result, values_from = n, values_fill = 0) %>%
  knitr::kable()
```

formula	total	non-significant	significant positive	significant negative
S	36	19	17	0
R	36	26	9	1
D	16	14	1	1
R + S	50	35	14	1
D + S	30	24	4	2
D + R	28	23	4	1
R x S	75	58	12	5
D x S	45	39	4	2
D x R	42	39	1	2

Looking at the above results, we see that the majority of effects are non-significant regardless of the model. Furthermore, while significant positive effects tend to outnumber significant negative ones, negative effects are not exceptionally uncommon either. In conclusion, any tendency towards positive relationships between the three predictors and food chain length appears to be sporadic at best, and there isn't a strong case for such positive correspondences to be the rule—with the possible exception of the isolated effect of ecosystem size (first row).

Additionally, one could claim that using a significance threshold of 0.05 is overly lenient, due to the problem of multiple testing. Lowering the threshold from 0.05 to 0.005 leads to a much lower prevalence of significant results, further reinforcing the picture that there is no strong general tendency in the data for either positive or negative relationships:

```r
plotData(datProcessed, signif = 0.005)
```




```
datProcessed %>%
  mutate(result = case_when(
    p.value < 0.005 & estimate >  0 ~ "significant positive",
    p.value < 0.005 & estimate <= 0 ~ "significant negative",
    TRUE ~ "non-significant"
  )) %>%
  group_by(formula) %>%
  count(result) %>%
  mutate(total = sum(n)) %>%
  ungroup() %>%
  pivot_wider(names_from = result, values_from = n, values_fill = 0) %>%
  knitr::kable()
```

formula	total	non-significant	significant positive	significant negative
S	36	29	7	0
R	36	35	0	1
D	16	14	1	1
R + S	50	45	5	0
D + S	30	28	1	1
D + R	28	27	0	1
R x S	75	70	3	2
D x S	45	44	1	0
D x R	42	41	0	1

Session information

The analyses were done using R version 4.2.2.[1] Full session information:

```
sessionInfo()

R version 4.2.2 (2022-10-31)
Platform: x86_64-apple-darwin17.0 (64-bit)
Running under: macOS Big Sur ... 10.16

Matrix products: default
BLAS:   /Library/Frameworks/R.framework/Versions/4.2/Resources/lib/libR
blas.0.dylib
LAPACK: /Library/Frameworks/R.framework/Versions/4.2/Resources/lib/libR
lapack.dylib

locale:
[1] en_US.UTF-8/en_US.UTF-8/en_US.UTF-8/C/en_US.UTF-8/en_US.UTF-8

attached base packages:
[1] stats     graphics  grDevices utils     datasets  methods   base

other attached packages:
 [1] lubridate_1.9.2 forcats_1.0.0   stringr_1.5.0   dplyr_1.1.2
 [5] purrr_1.0.1     readr_2.1.4     tidyr_1.3.0     tibble_3.2.1
 [9] ggplot2_3.4.2   tidyverse_2.0.0

loaded via a namespace (and not attached):
 [1] highr_0.9        pillar_1.9.0     compiler_4.2.2   tools_4.2.2
 [5] bit_4.0.5        digest_0.6.30    timechange_0.1.1 jsonlite_1.8.4
 [9] evaluate_0.18    lifecycle_1.0.3  gtable_0.3.1     pkgconfig_2.0.3
 [13] rlang_1.1.1      cli_3.6.1        rstudioapi_0.14  parallel_4.2.2
 [17] yaml_2.3.6       xfun_0.34        fastmap_1.1.0    withr_2.5.0
 [21] knitr_1.40       hms_1.1.2        generics_0.1.3   vctrs_0.6.2
 [25] bit64_4.0.5      grid_4.2.2       tidyselect_1.2.0 glue_1.6.2
 [29] R6_2.5.1         fansi_1.0.3      vroom_1.6.0      rmarkdown_2.18
 [33] farver_2.1.1     tzdb_0.3.0       magrittr_2.0.3   backports_1.4.1
 [37] ellipsis_0.3.2   scales_1.2.1   htmltools_0.5.3  colorspace_2.0-3
 [41] labeling_0.4.2   utf8_1.2.2       stringi_1.7.8    munsell_0.5.0
 [45] broom_1.0.4      crayon_1.5.2
```

[1] R Core Team (2022). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Supporting Information S2 - *System analysis*

Stability analysis

For mathematical tractability, we assume donor-controlled dynamics by ignoring the top-down predation, i.e., $P_i \sum_{k=1}^{n_A} \theta_{ik}\mu_{ik}A_k = 0$ in Equation (1) and $A_i \sum_{k=1}^{n_A} \varphi_{ik}\delta_{ik}A_k = 0$ in Equation (3) (cf. Holt, 1997; Gravel et al., 2011; Häussler et al., 2020). However, we relax this assumption with numerical simulations and find that our results are robust to recipient-controlled dynamics. Thus, Equation (1) in *Methods* can be rearranged as

$$\frac{dP_i}{dt} = P_i\left[\underbrace{c_i^P RS - e_i^P + f(t,D,T)}_{g_i} + R\sum_{j=1}^{n_P}\underbrace{\left(c_i^P H_{ij} - c_j^P H_{ji} - c_i^P\right)}_{M_{ij}} P_j\right]. \quad \text{(S1)}$$

In this formulation, g_i is the effective intrinsic growth rate of basal species *i*, while M_{ij} is the effective interaction coefficient in a matrix $\boldsymbol{M}$ (i.e., the effects of intra- and inter-specific competition). The net effect of these two terms in the square bracket is the *per-capita* growth rate $r_i = \frac{1}{P_i}\frac{dP_i}{dt}$ of basal species *i*, which is linear with respect to the population size P_j. In particular, the *per-capita* growth rate has the Lotka-Volterra form $r_i = g_i + R\sum_{j=1}^{n_P} M_{ij}P_j$. This linearity allows one to take the time average of the *per-capita* growth rate directly:

$$\bar{r}_i = \bar{g}_i + R\sum_{j=1}^{n_P} M_{ij}\bar{P}_j, \quad \text{(S2)}$$

where the over-bar represents time averaging. Here P_i declines to $(1-D)P_i$ during every period *T*, thus we set $f(t,D,T) = log(1-D)/T$, which gives the same long-term average result as the periodically disturbed model (cf. Liao et al. 2022). Since the effects of a disturbance with extent *D* and periodicity *T* are equivalent to the effects of another disturbance with extent $D' = 1-(1-D)^{1/T}$ and periodicity T'=1 in competition-colonization (C-C) tradeoff communities (Liao et al., 2022). Thus, we only vary *D* alone while keeping *T*=1 throughout, which is sufficient for achieving a full understanding of the impact of disturbance. Thus, Equation (S2) has at most one

fixed point where all species populations P_i^* are positive (i.e., a coexistence steady state). At this steady state ($\bar{r}_i$=0), we can express the long-term average site occupancy of basal species *i* explicitly by inverting the matrix $\boldsymbol{M}$:

$$\bar{P}_i^* = -\sum_{j=1}^{n_P}(\boldsymbol{M}^{-1})_{ij}\left[c_j^P S - e_j^P/R + log(1-D)/R\right], \quad \text{(S3)}$$

where $(\boldsymbol{M}^{-1})_{ij}$ is the $(i,j)th$ entry of the inverse of the effective interaction matrix $\boldsymbol{M}$. Furthermore, if the tournament matrix $\boldsymbol{H}$ is fully hierarchical ($H_{ij} = 1$ if $i < j$ and 0 otherwise), the feasible equilibrium point in which the most species survive is stable (cf. Liao et al., 2022).

For consumers, if we ignore the top-down predation in Equation (3), we have

$$\frac{dA_i}{dt} = A_i\left[c_i^A\left(\sum_{j=1}^{n_P}\theta_{ji}P_j + \sum_{k=1}^{n_A}\delta_{ki}A_k\right)(S - A_i) - e_i^A\right]. \quad \text{(S4)}$$

Given that the equilibrium point is feasible, we can express the long-term average occupancies for consumer i at steady state as

$$\bar{A}_i^* = S - \frac{e_i^A}{c_i^A\left(\sum_{j=1}^{n_P}\theta_{ji}\bar{P}_j^* + \sum_{k=1}^{n_A}\delta_{ki}\bar{A}_k^*\right)}, \quad \text{(S5)}$$

in which $\bar{P}_j^*$ is already determined from Equation (S3), independent of the site-occupancy dynamics of consumers. If $\delta_{ki} = 1$, i.e., the consumer *i* can feed on consumer *k*, then $\bar{A}_i^*$ is related to $\bar{A}_k^*$, but $\bar{A}_k^*$ is irrelevant to $\bar{A}_i^*$. When $\delta_{ki} = 0$, the equilibrium site occupancies of both consumers *i* and *k* are mutually independent, as we neglect all top-down effects in the whole trophic system. As such, the survival of consumer *i* depends on the abundances of its prey species at lower trophic levels. This is actually a recursion relation, with the initial condition for basal species given by Equation (S3). Therefore, using these equations, we can express the equilibrium occupancies of the whole community, provided that the equilibrium is feasible.

The checkerboard pattern of the inverse community matrix

Similar to Liao et al. (2022), here we show that, along the gradients of multiple environmental variables, the observed oscillating patterns in basal species diversity (ignoring the top-down predation) can be understood from

$$\bar{P}_i^* = -\sum_{j=1}^{n_P} (\boldsymbol{M}^{-1})_{ij} \left[c_j^P S - e_j^P / R + log(1-D)/R \right], \quad \text{(S6)}$$

where the effective interaction coefficient $M_{ij} = c_i^P H_{ij} - c_j^P H_{ji} - c_i^P$ in a matrix $\boldsymbol{M}$. This yields the long-term site occupancies as a function of ecosystem size (S), resource productivity (R) and disturbance extent (D). Note that decreasing S, R or increasing D (i.e., in harsher environments) will decrease $\left[c_j^P S - e_j^P / R + log(1-D)/R \right]$, the term multiplied by the inverse of M_{ij}. Since the equilibrium occupancies are functions of S, R and D, we can write $\bar{P}_i^* = \bar{P}_i^*(S, R, D)$.

Here we assume that the matrix H is fully hierarchical: $H_{ij} = 1$ if $i < j$ and 0 otherwise. In that case, $c_i^P H_{ij} - c_i^P$ cancels each other for upper triangular ($i < j$) entries, and $\boldsymbol{M}$ reduces to $-(c_j^P + c_i^P)$ in the lower triangular entries and to $-c_i^P$ along its diagonal. Introducing the matrices C and L, where C is diagonal with its i-th diagonal entry equal to $-c_i^P$, and L is lower triangular with entries $L_{ij} = -(c_i^P + c_j^P)\Theta_{ij}$ (where $\Theta_{ij}=1$ for $i > j$ and 0 otherwise), we can then write $\boldsymbol{M}$ as the sum of the two: $M=C+L$.

Since all $c_i^P > 0$, the diagonal matrix C is invertible. Its inverse C^{-1} is itself a diagonal matrix with the $-1/c_i^P$ along its diagonal. One can then equivalently write $M=C+L$ as

$$M = C\,(I + C^{-1}L). \quad \text{(S7)}$$

The inverse of M as a whole can thus be written as

$$M^{-1} = (I + C^{-1}L)^{-1}C^{-1}. \quad \text{(S8)}$$

We now use the known identity $(I - B)^{-1}=\sum_{k=0}^{\infty} B^k$ that holds for any matrix B with eigenvalues falling inside the unit circle (the Neumann series expansion). In our case, $B = -C^{-1}L$, a strictly lower triangular matrix. The eigenvalues of strictly lower triangular matrices are all equal to 0 (these matrices are nilpotent), which do of course fall in the unit circle. The Neumann series expansion therefore holds, and we can write

$$M^{-1} = \sum_{k=0}^{\infty}(-1)^k\,(C^{-1}L)^k C^{-1}. \tag{S9}$$

Even more is true: since the n_P-th power of a strictly lower triangular matrix is guaranteed to vanish, we can terminate the above infinite sum at n_P-1:

$$M^{-1} = \sum_{k=0}^{n_P-1}(-1)^k\,(C^{-1}L)^k C^{-1}. \tag{S10}$$

For the following, it will be easier if we multiply both sides by C from the right, and work with $M^{-1}C$:

$$M^{-1}C =\sum_{k=0}^{n_P-1}(-1)^k\,(C^{-1}L)^k. \tag{S11}$$

Let us examine the powers of $C^{-1}L$ in more detail. Its 0th power is simply the identity matrix: $(C^{-1}L)^0 = I$, or $(C^{-1}L)^0_{ij} = \delta_{ij}$ for its (i, j)th entry (the Kronecker symbol δ_{ij} is 1 if $i = j$ and 0 otherwise). The (i, j)th entry of the first power $(C^{-1}L)^1 = C^{-1}L$ reads, using $L_{ij} = -(c_i^P + c_j^P)\Theta_{ij}$, as

$$(C^{-1}L)_{ij} = \sum_{k=1}^{n_P}\frac{1}{c_i^P}\delta_{ik}\left(c_k^P + c_j^P\right)\Theta_{kj} = \left(1 + \frac{c_j^P}{c_i^P}\right)\Theta_{ij}, \tag{S12}$$

with Θ_{ij} restricting its nonzero entries below the main diagonal. The (i, j)th entry of the second power is

$$(C^{-1}L)^2_{ij} = \sum_{k=1}^{n_P}\left(1 + \frac{c_k^P}{c_i^P}\right)\left(1 + \frac{c_j^P}{c_k^P}\right)\Theta_{ik}\Theta_{kj} = \sum_{k=j+1}^{i-1}\left(1 + \frac{c_k^P}{c_i^P}\right)\left(1 + \frac{c_j^P}{c_k^P}\right), \tag{S13}$$

where the summation is understood to yield zero if $j + 1 > i - 1$. Clearly, as long as

$j+1 \leq i-1$, the contribution of $(C^{-1}L)^2$ to the $(i, j)$the entry always exceeds the contribution of $C^{-1}L$ (since the c_i^P are all positive). The condition $j + 1 \leq i - 1$ restricts the entries of $(C^{-1}L)^2$ below the first subdiagonal. A similar argument establishes that the nonzero entries of $(C^{-1}L)^3$ exceed the corresponding ones in $(C^{-1}L)^2$ (and are restricted to below the second subdiagonal), and so on: $(C^{-1}L)^k > (C^{-1}L)^{k-1}$ for entries below the $(k-1)$th subdiagonal.

As seen from Equation (S11), $(C^{-1}L)^k$ is multiplied by $(-1)^k$ in the summation. When summing over k, the main diagonal is $(-1)^0(C^{-1}L)^0 = I$. The first subdiagonal is given by the corresponding entries of $-C^{-1}L$, which are all negative. The second subdiagonal is determined by the corresponding entries of $-C^{-1}L+(C^{-1}L)^2$; however, since we established that the nonzero entries of $(C^{-1}L)^k$ exceed those of $(C^{-1}L)^{k-1}$, these entries will be positive. Continuing the same argument, the entries in the second subdiagonal $[-C^{-1}L+(C^{-1}L)^2-(C^{-1}L)^3]$ will again be negative; the ones in the 3rd subdiagonal positive, and so on: the subdiagonals keep alternating signs.

All this is true for $M^{-1}C$ (Equation S11). To obtain M^{-1} itself, to be used in Equation (S6), we multiply from the right with the diagonal matrix C^{-1}. Its effect is to multiply each column of $M^{-1}C$ by $-1/c_i^P$. This flips the sign of each entry and adjusts the magnitudes of the nonzero entries, without affecting the alternating sign-pattern in $\boldsymbol{M}$, which therefore looks like this:

$$M^{-1} \sim \begin{pmatrix} - & 0 & 0 & 0 & \cdots & 0 \\ + & - & 0 & 0 & \cdots & 0 \\ - & + & - & 0 & \cdots & 0 \\ + & - & + & - & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \end{pmatrix}. \qquad \text{(S14)}$$

Let us now see what happens when the top species in the hierarchy (species 1) goes extinct. As the top species also has the lowest c_i^P, we can assume that this species goes extinct firstly with environmental deterioration (decreasing S and R, or increasing D). Thus, the reduced growth due to environmental deterioration will bring the density of the top species to zero before other species. The effect of species 1 on

the other species is summarized by the first column of $\boldsymbol{M}$: species 2 is positively affected by 1, species 3 negatively, species 4 positively again, and so on. Thus, the removal of species 1 hurts 2, helps 3, hurts 4 again etc., resulting in a sharp change in the trajectories of all $\bar{P}_i^*$ as a function of decreasing *S*, *R* or increasing *D* (Equation S6). If the effect is strong enough to not just change the trajectory but turn increasing ones into decreasing ones and vice versa, then a oscillating biodiversity pattern obtains.

In fact, while we did not manage to find a formal proof, even more is true: any entry $(M^{-1})_{ij}$ with $i > j$ is such that

$$\left|(M^{-1})_{ij}\right| \geq \sum_{k=j+1}^{i}(M^{-1})_{ik} \quad (i > j). \tag{S15}$$

The consequence is that the extinction of the current top species will indeed change increasing $\bar{P}_i^*(S,R,D)$ curves to decreasing ones, and vice versa. This conjecture held in every case we checked, and we suspect it is in fact a theorem. However, even if one treats it as just a well-supported conjecture, it helps explain the observed oscillating diversity patterns in response to multiple environmental drivers.

Finally, these results are maintained even if the matrix $\boldsymbol{H}$ is not fully hierarchical. By the general continuity of the $\boldsymbol{M \mapsto M^{-1}}$ mapping that holds for any invertible matrix, a sufficiently small change in $\boldsymbol{H}$ can only cause a small change in $\boldsymbol{M}^{-1}$. If $\boldsymbol{H}$ is still upper triangular, then a small enough change cannot alter the sign pattern of Equation (S14). If $\boldsymbol{H}$ is no longer upper triangular, then the upper triangular entries of $\boldsymbol{M}^{-1}$ will no longer be exactly zero – however, as long as the deviation of $\boldsymbol{H}$ from upper triangularity is small, this will not override the overall $\bar{P}_i^*(S,R,D)$ patterns.

References

Gravel, D., Canard, E., Guichard, F. & Mouquet, N. (2011) Persistence increases with diversity and connectance in trophic metacommunities. *PLoS One*, 6, e19374.

Häussler, J., Barabás, G. & Eklöf, A. (2020) A Bayesian network approach to trophic metacommunities shows that habitat loss accelerates top species extinctions. *Ecology Letters*, 23, 1849–1861.

Holt, R.D. (1997) From metapopulation dynamics to community structure: Some consequences of spatial heterogeneity. *Metapopulation Biology*. Academic Press, pp. 149–164.

Liao, J., Barabás G. & Bearup, D. (2022) Competition–colonization dynamics and multimodality in diversity–disturbance relationships. *Ecology*, 103, e3672.

Supporting Information S3 – *Figures S1-S18*

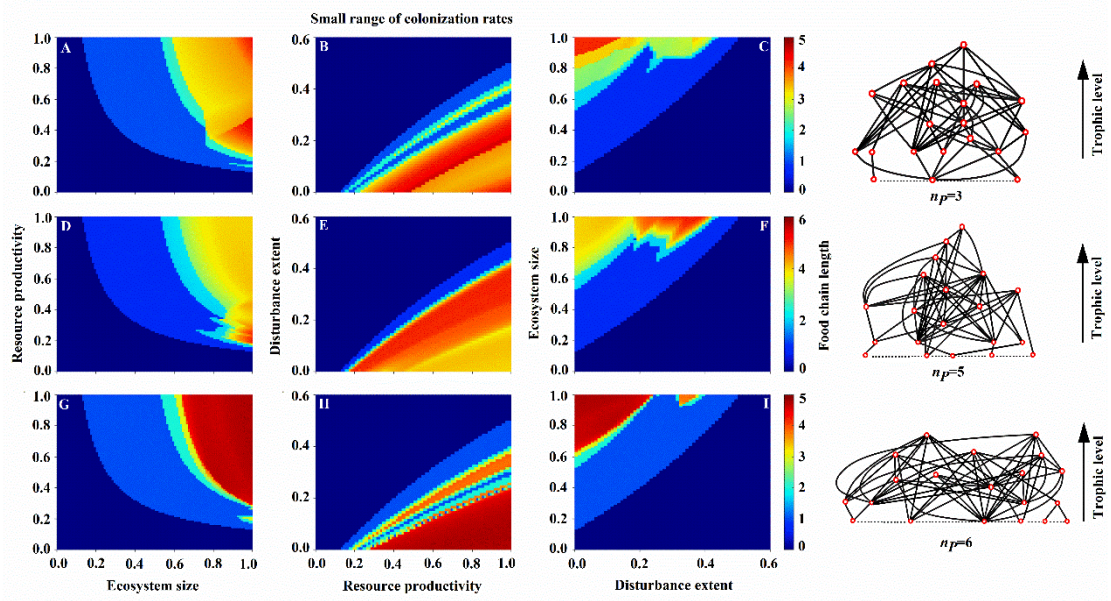


Figure S1. Interactive effects of ecosystem size (*S*), resource productivity (*R*), and disturbance extent (*D*) on food chain length (FCL) in given typical food webs on the right (total species number *N*=20 and connectance *C*=0.15), with basal species n_P=3, 5 & 6 (red circles – species, black lines – trophic links, and dotted lines – basal species competition). The basal species are ranked from the best competitor (species 1) to the poorest (species n_P) in a strict competitive hierarchy, i.e., $H_{ij} = 1$ for $i<j$ and 0 otherwise in a matrix $\boldsymbol{H}$. To establish the possibility of competition-colonization (C-C) tradeoffs, basal species' colonization rates are evenly spaced in increasing order at a small range ($c_i^P \epsilon E$[0.45, 0.8]). Other parameters: all species mortality rates $e_i^P = e_i^A$=0.1, all consumers' colonization rates c_i^A=0.625 and all top-down mortality rates due to predation $\mu_{ik} = \varphi_{ik}$=0.05.

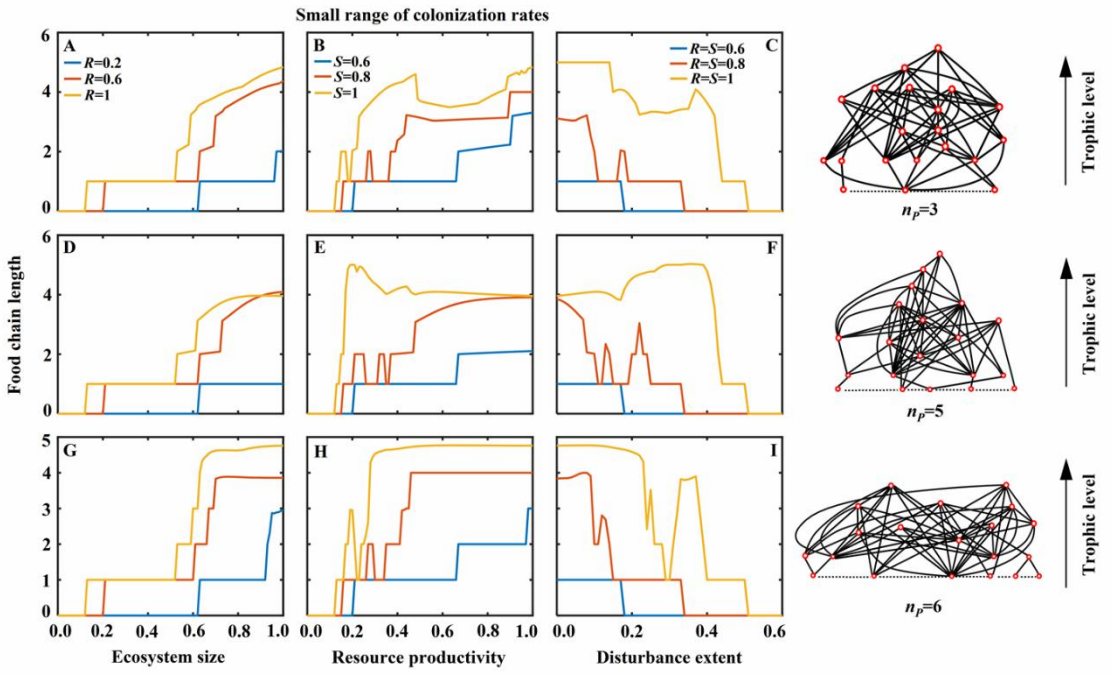


Figure S2. Individual effects of ecosystem size (*S*), resource productivity (*R*), and disturbance extent (*D*) on FCL in given food webs as displayed on the right. Panels (A, D & G) *R*=0.2, 0.6 & 1 with *D*=0; panels (B, E & H) *S*=0.6, 0.8 & 1 with *D*=0; and panels (C, F & I) *R*=*S*=0.6, 0.8 & 1. Other parameter settings are the same as in Figure S1 above.

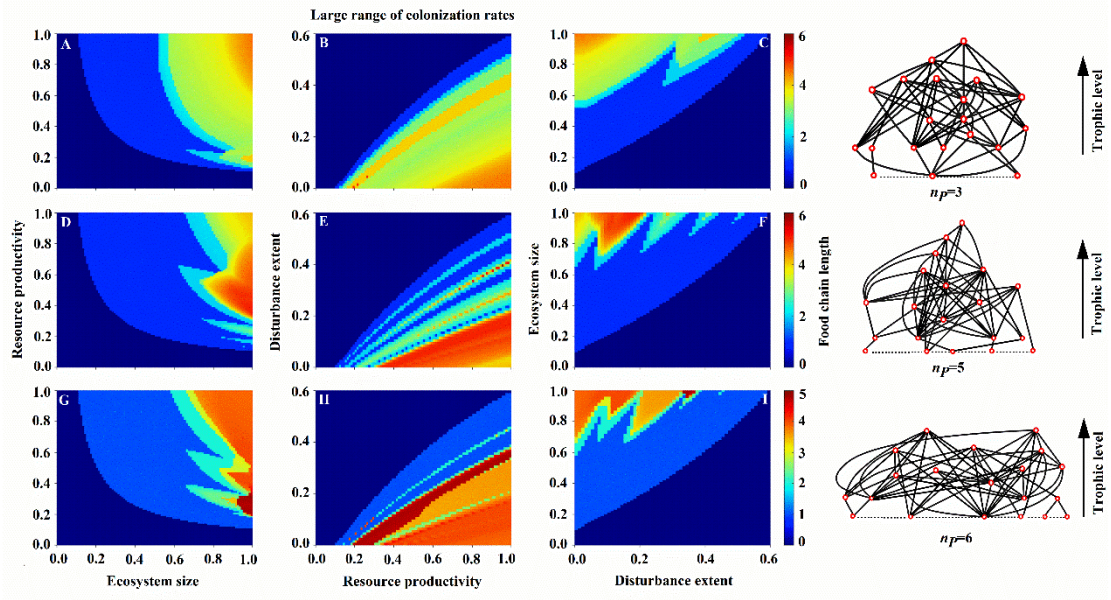


Figure S3. Interactive effects of ecosystem size (*S*), resource productivity (*R*), and disturbance extent (*D*) on FCL in given typical food webs on the right. Basal species' colonization rates are evenly spaced in increasing order at a large range ($c_i^P \epsilon E$[0.25, 1]). Other parameter settings are the same as in Figure S1 above.

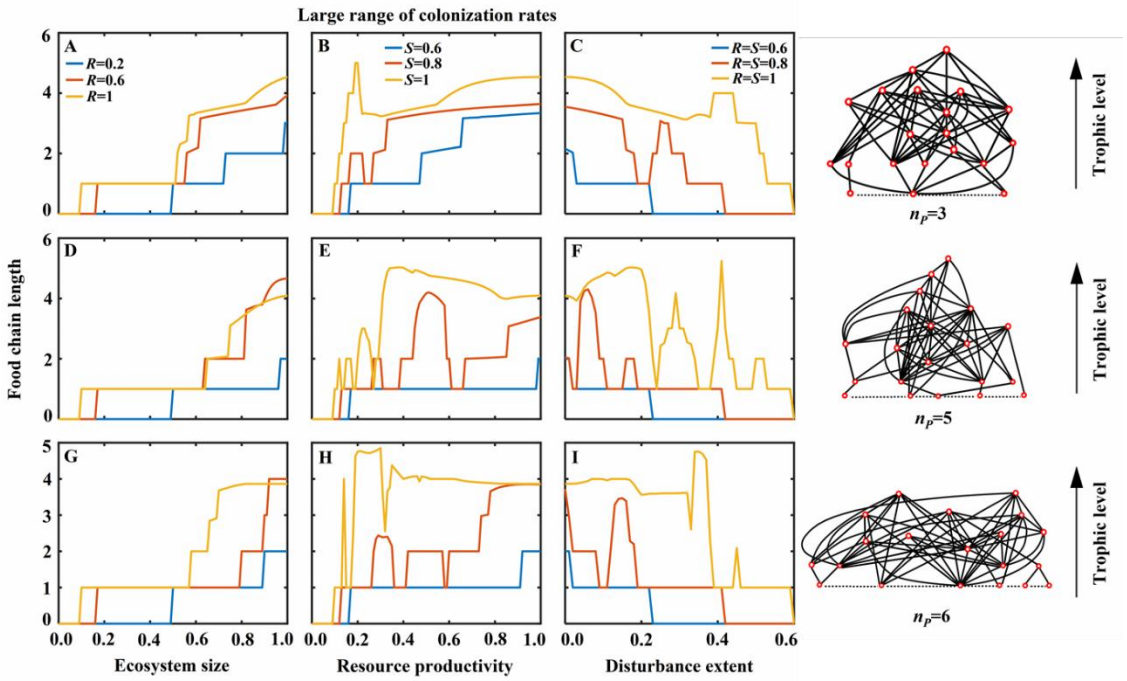


Figure S4. Individual effects of ecosystem size (*S*), resource availability (*R*), and disturbance extent (*D*) on FCL in given food webs as displayed on the right. Panels (A, D & G) *R*=0.2, 0.6 & 1 with *D*=0; panels (B, E & H) *S*=0.6, 0.8 & 1 with *D*=0; and panels (C, F & I) *R*=*S*=0.6, 0.8 & 1. Other parameter settings: see Figure S3 above.

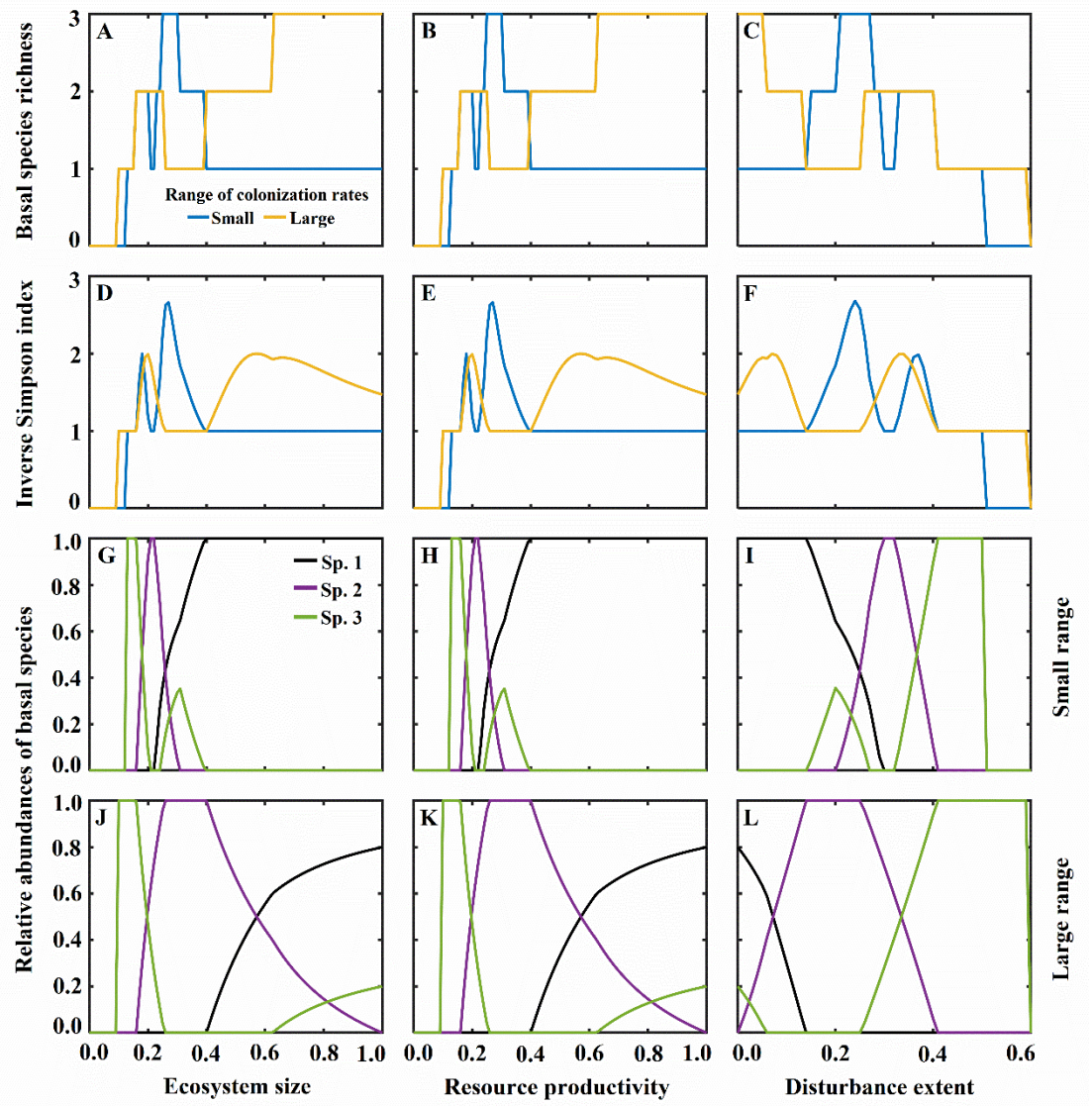


Figure S5. Individual effects of ecosystem size (*S*), resource productivity (*R*) and disturbance extent (*D*) on basal species diversity (A-F) and their relative abundances (G-L) at steady state for initial richness n_P=3, while ignoring the top-down predation. Basal species diversity is characterized by (A-C) species richness and (D-F) the inverse Simpson index ($1/\sum q_i^2$, with $q_i = P_i/\sum P_j$ being the relative abundance of basal species *i*). Panels (A, D, G & J): *R*=1 & *D*=0; panels (B, E, H & K): *S*=1 & *D*=0; and panels (C, F, I & L): *R*=*S*=1. To establish the C-C tradeoff, basal species are ranked from the best competitor (species 1) to the poorest (species n_P) in a strict competitive hierarchy ($H_{ij} = 1$ for *i*<*j* and 0 otherwise in a matrix ***H***), while their colonization rates are evenly spaced in increasing order at both small ($c_i^P \epsilon E[0.45, 0.8]$) and large ($c_i^P \epsilon E[0.25, 1]$) ranges. Other parameters: e_i^P=0.1 for all basal species.

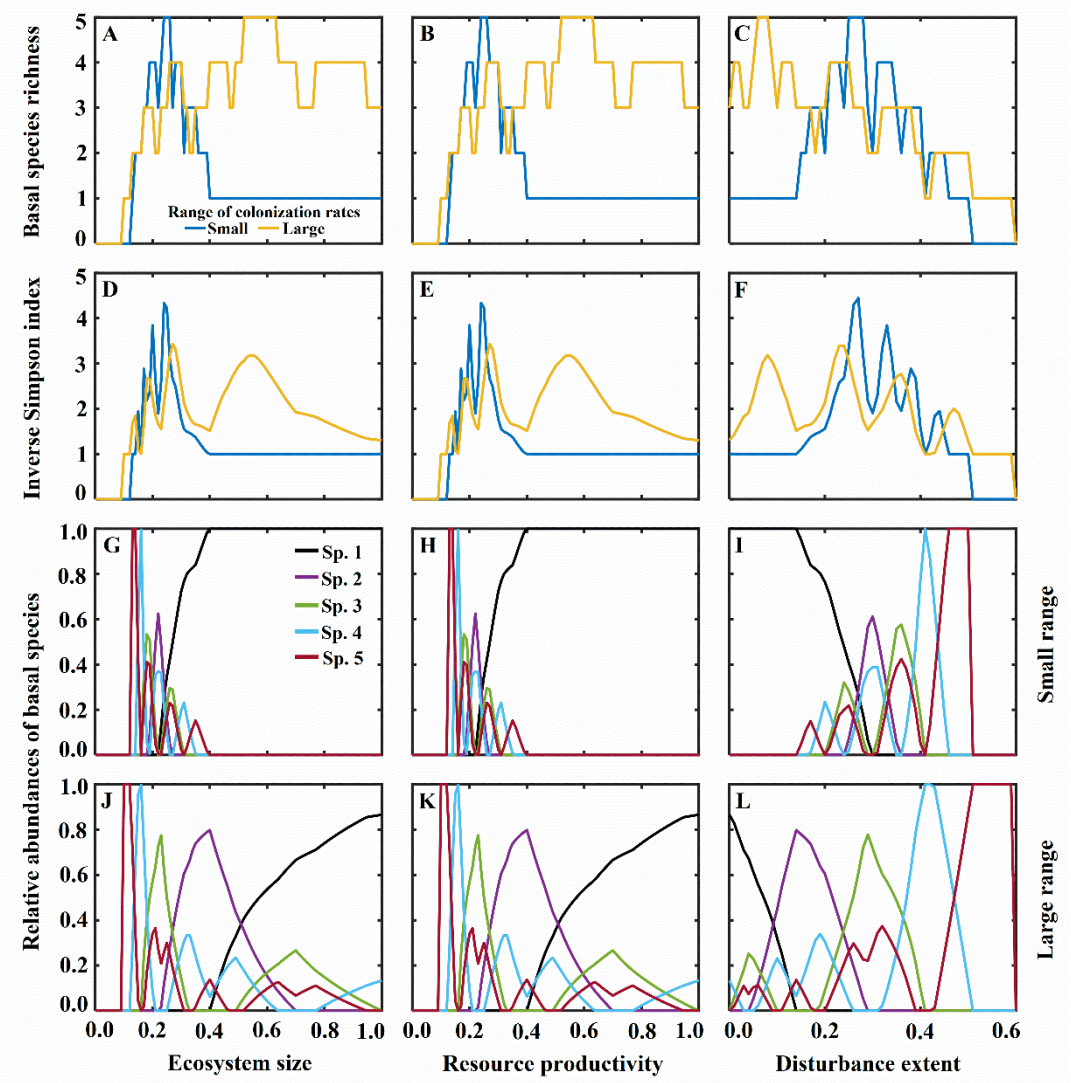


Figure S6. Individual effects of ecosystem size (*S*), resource productivity (*R*) and disturbance extent (*D*) on basal species diversity (A-F) and their relative abundances (G-L) at steady state for initial richness n_P=5, while ignoring the top-down predation. Others are the same as in Figure S5 above.

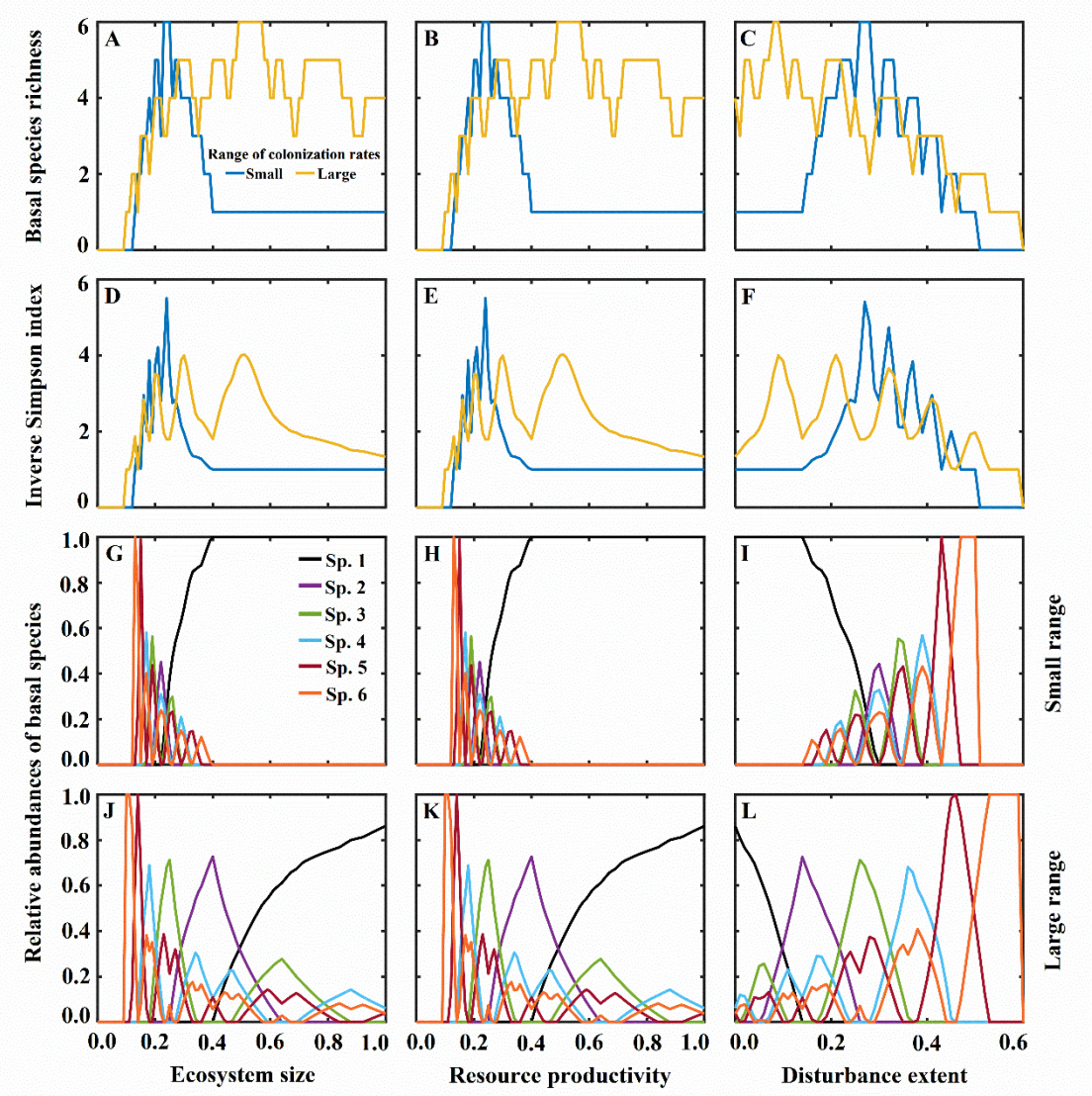


Figure S7. Individual effects of ecosystem size (*S*), resource productivity (*R*) and disturbance extent (*D*) on basal species diversity (A-F) and their relative abundances (G-L) at steady state for initial richness n_P=6, while ignoring the top-down predation. Other parameter settings: see Figure S5 above.

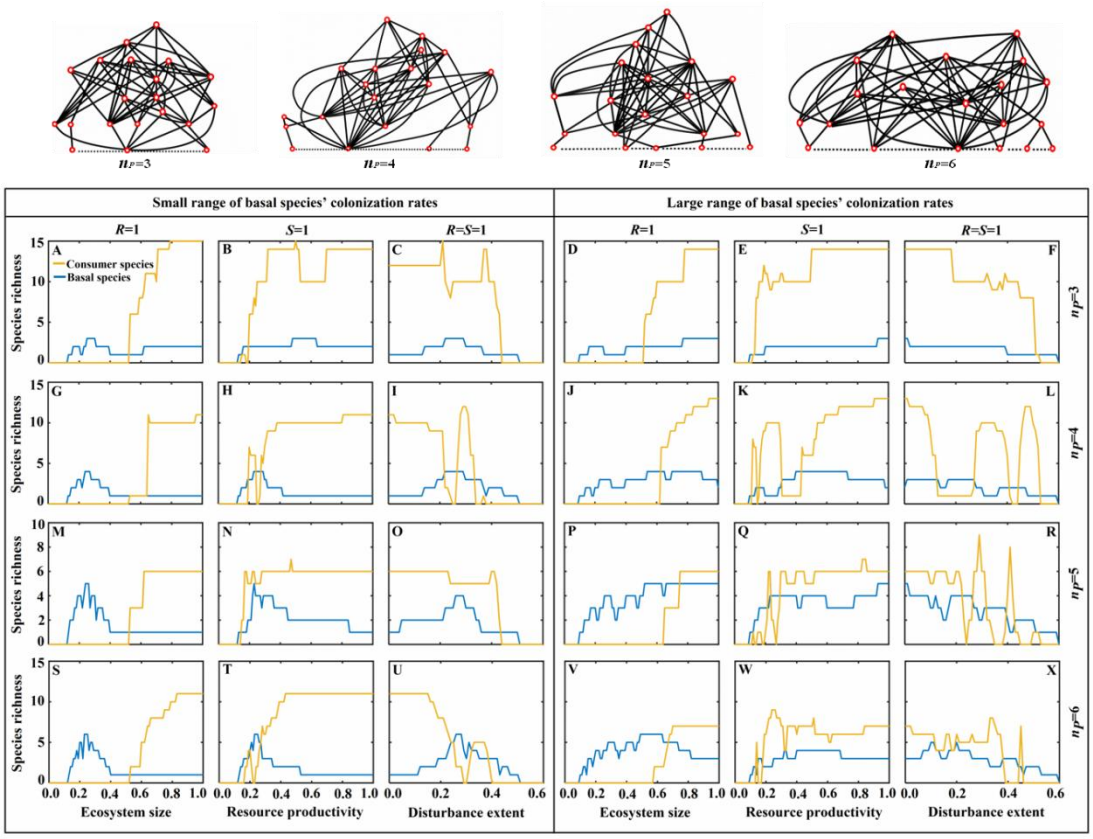


Figure S8. Individual effects of ecosystem size (*S*), resource productivity (*R*) and disturbance extent (*D*) on basal species richness and consumer species richness in given food webs on the top, with different initial basal species richness $n_P = 3, 4, 5$ & 6. Panels (A, D, G, J, M, P, S & V): *R*=1 and *D*=0; panels (B, E, H, K, N, Q, T & W): *S*=1 and *D*=0; and panels (C, F, I, L, O, R, U & X): *R*=*S*=1. Basal species' colonization rates are evenly spaced in increasing order at both small (on the left half: $c_i^P \epsilon E[0.45, 0.8]$) and large (on the right half: $c_i^P \epsilon E[0.25, 1]$) ranges. The basal species are ranked from the best competitor (species 1) to the poorest (species n_P) in a strict competitive hierarchy, i.e., $H_{ij} = 1$ for $i<j$ and 0 otherwise in a matrix $\boldsymbol{H}$. Other parameters: $e_i^P = e_i^A$=0.1, c_i^A=0.625 and $\mu_{ik} = \varphi_{ik}$=0.05.

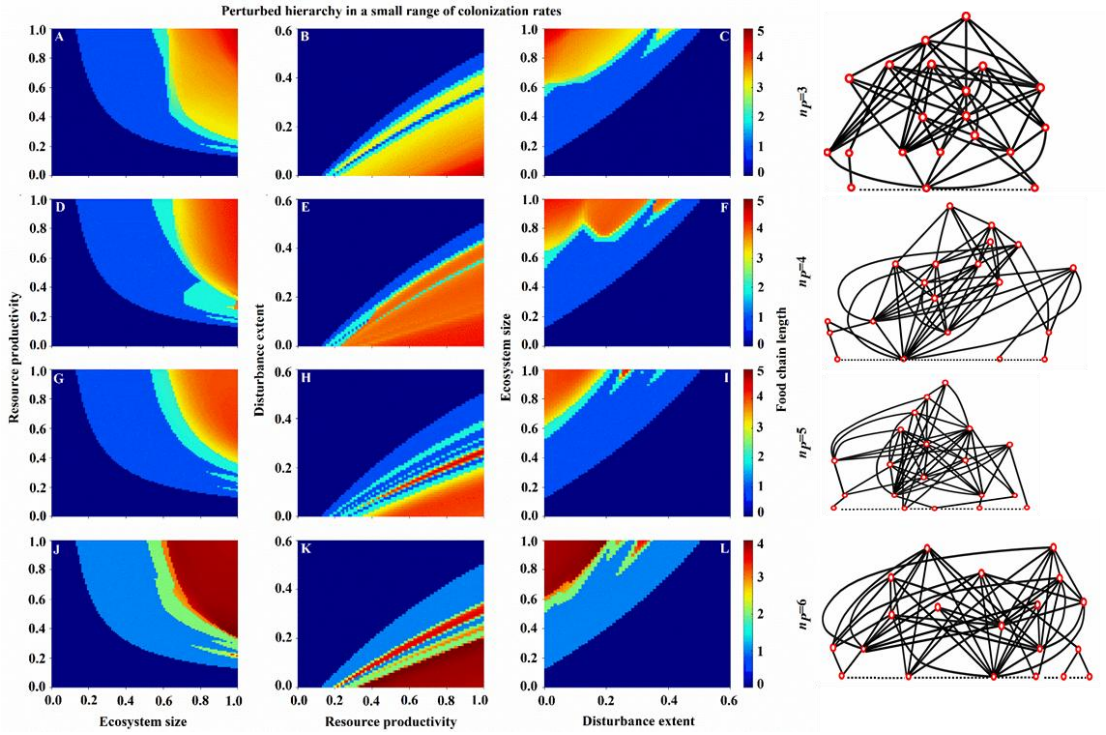


Figure S9. Interactive effects of ecosystem size (*S*), resource productivity (*R*), and disturbance extent (*D*) on FCL in given food webs on the right (with basal species richness $n_P = 3, 4, 5$ & 6). Basal species' colonization rates are evenly spaced in increasing order at a small range ($c_i^P \epsilon E[0.45, 0.8]$), while weakening their competitive hierarchy $\boldsymbol{H}$: the lower and upper triangular entries (H_{ij}) are uniformly sampled from $U[0, 0.25]$ and $U[0.75, 1]$, respectively. Other parameters: $e_i^P = e_i^A$=0.1, c_i^A=0.625 and $\mu_{ik} = \varphi_{ik}$=0.05.

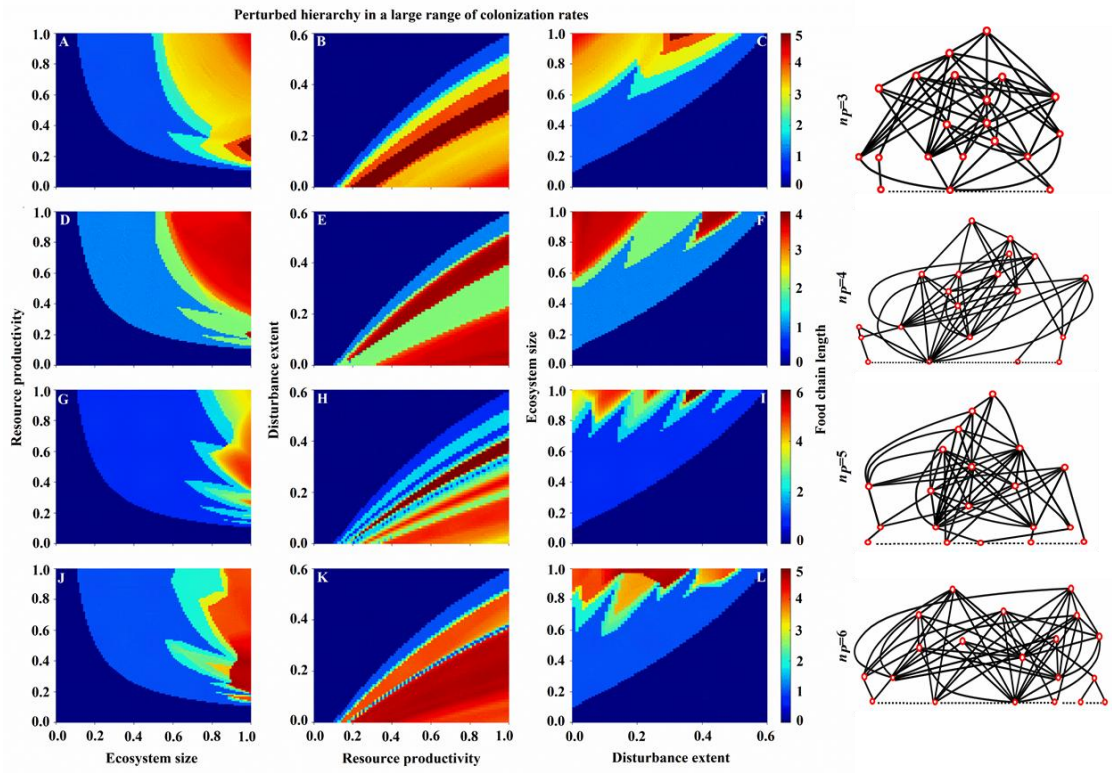


Figure S10. Interactive effects of ecosystem size (*S*), resource productivity (*R*), and disturbance extent (*D*) on FCL in given food webs on the right (with basal species richness $n_P = 3, 4, 5$ & 6). Basal species' colonization rates are evenly spaced in increasing order at a large range ($c_i^P \epsilon E[0.25, 1]$), while weakening their competitive hierarchy ***H***: the lower and upper triangular entries (H_{ij}) are uniformly sampled from $U[0, 0.25]$ and $U[0.75, 1]$, respectively. Other parameter settings are the same as in Figure S9 above.

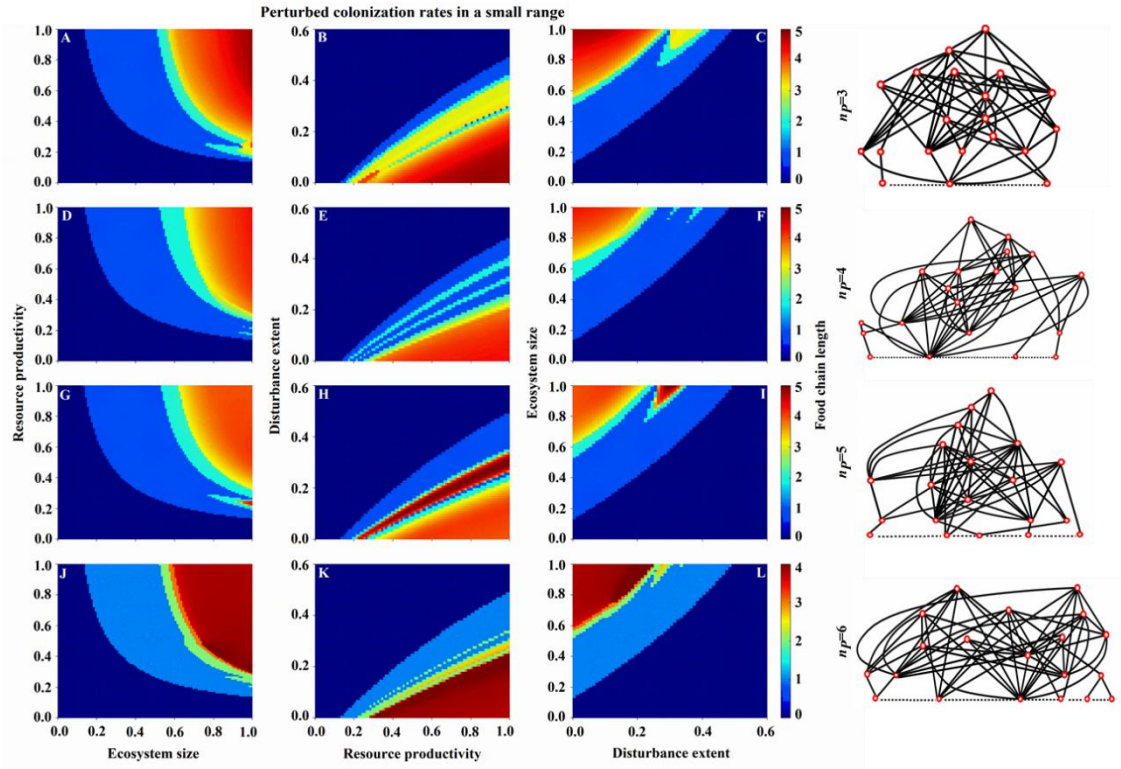


Figure S11. Interactive effects of ecosystem size (*S*), resource productivity (*R*), and disturbance extent (*D*) on FCL in given food webs on the right (with basal species richness $n_P = 3, 4, 5$ & 6). Basal species' colonization rates are uniformly drawn from a small range ($c_i^P \epsilon U[0.45, 0.8]$) and sorted in increasing order, but with a strict competitive hierarchy $\boldsymbol{H}$ ($H_{ij} = 1$ for $i<j$ and 0 otherwise). Other parameters: $e_i^P = e_i^A$=0.1, c_i^A=0.625 and $\mu_{ik} = \varphi_{ik}$=0.05.

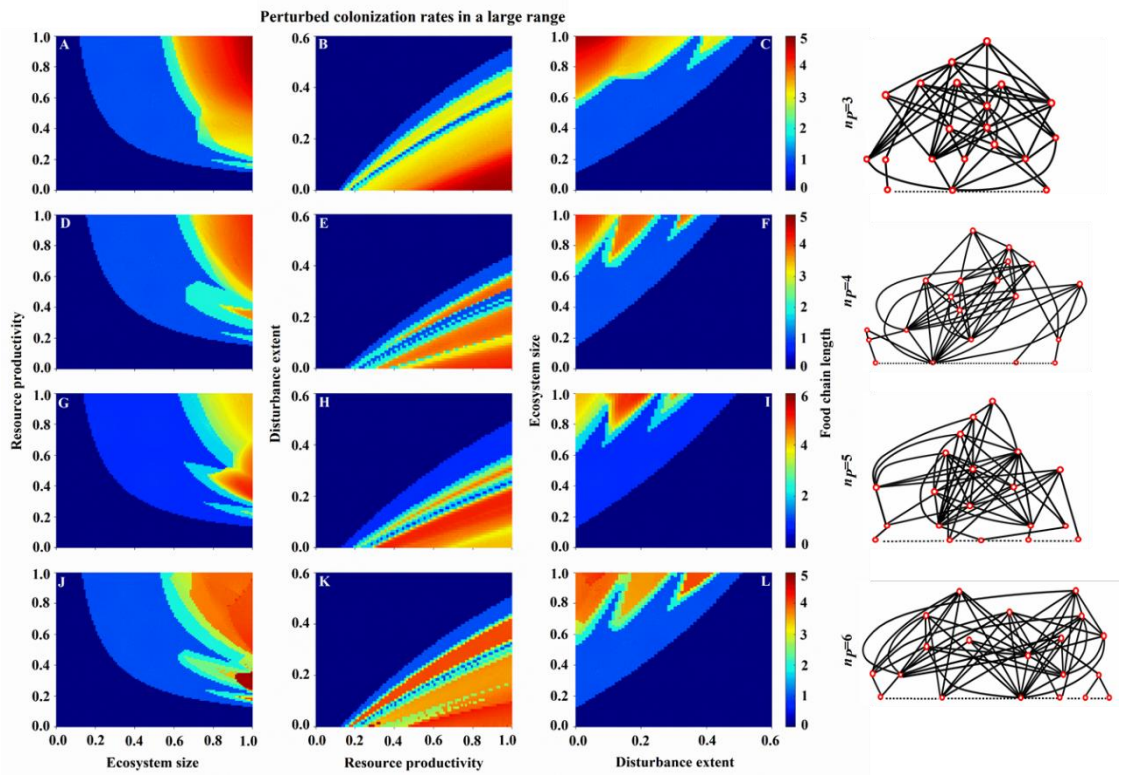


Figure S12. Interactive effects of ecosystem size (*S*), resource productivity (*R*), and disturbance extent (*D*) on FCL in given food webs on the right (with basal species richness $n_P = 3, 4, 5$ & 6). Basal species' colonization rates are uniformly drawn from a large range ($c_i^P \epsilon U[0.25, 1]$) and sorted in increasing order, but with a strict competitive hierarchy $\boldsymbol{H}$ ($H_{ij} = 1$ for $i<j$ and 0 otherwise). Other parameters: $e_i^P = e_i^A$=0.1, c_i^A=0.625 and $\mu_{ik} = \varphi_{ik}$=0.05.

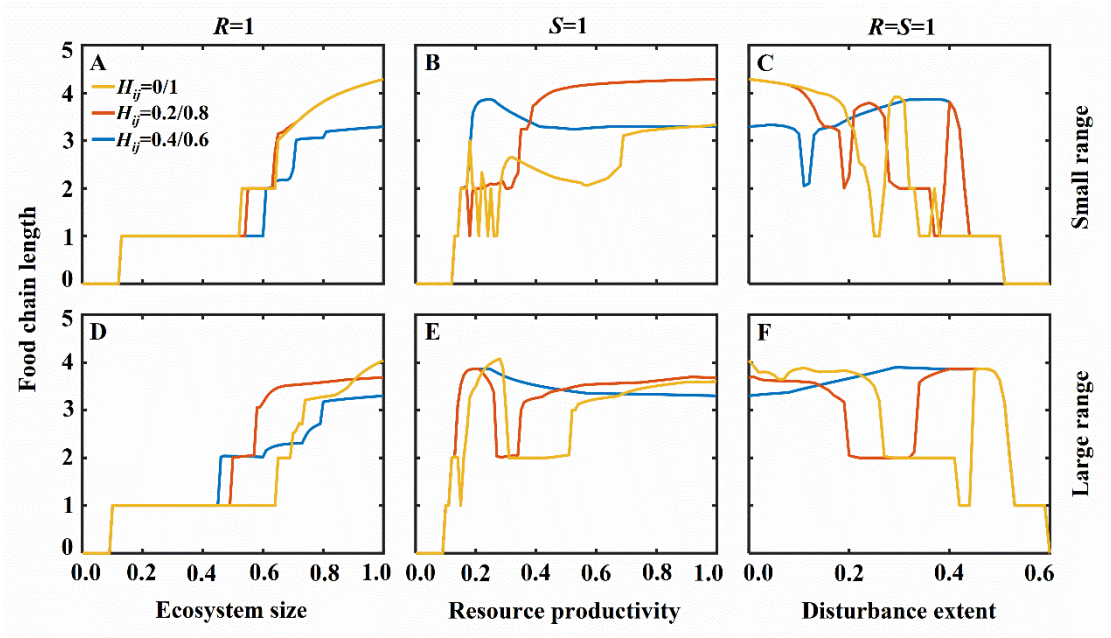


Figure S13. Individual effects of ecosystem size (S), resource productivity (R) and disturbance extent (D) on FCL in a given food web with $n_P = 4$, as displayed in Figure 2. Panels (A & D): *R*=1 & *D*=0; panels (B & E): *S*=1 & *D*=0; and panels (C & F): *R*=*S*=1. Basal species' colonization rates are evenly spaced in increasing order at both small (A-C: $c_i^P \epsilon E[0.45, 0.8]$) and large (D-F: $c_i^P \epsilon E[0.25, 1]$) ranges, while gradually weakening a strict competitive hierarchy ***H***: the upper triangular entries H_{ij}=1, 0.8 or 0.6, corresponding to the lower triangular entries H_{ji}=0, 0.2 or 0.4. Other parameters: $e_i^P = e_i^A$=0.1, c_i^A=0.625 and $\mu_{ik} = \varphi_{ik}$=0.05.

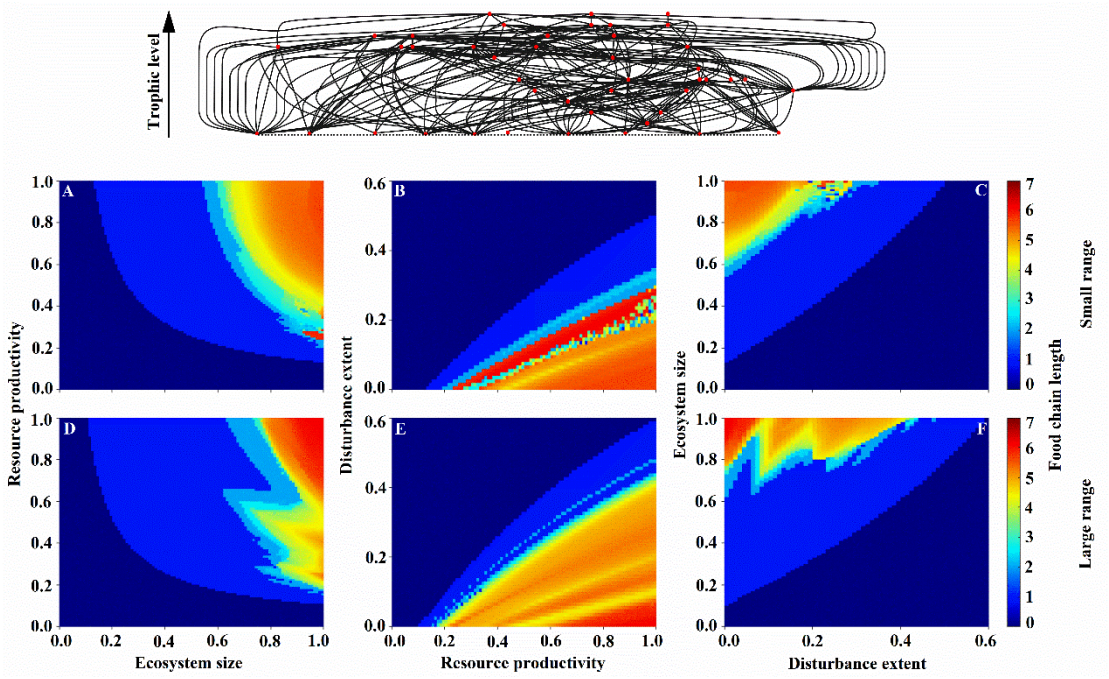


Figure S14. Interactive effects of ecosystem size (*S*), resource productivity (*R*) and disturbance extent (*D*) on FCL in a large food web on the top (with species diversity *N*=44, connectance *C*=0.1322 and basal species richness n_P=10). The basal species are ranked from the best competitor (species 1) to the poorest (species n_P) in a strict competitive hierarchy, i.e., $H_{ij} = 1$ for $i<j$ and 0 otherwise in a matrix $\boldsymbol{H}$. To establish the possibility of C-C tradeoffs, basal species' colonization rates are evenly spaced in increasing order at both small (A-C: $c_i^P \epsilon E[0.45, 0.8]$) and large (D-F: $c_i^P \epsilon E[0.25, 1]$) ranges. Others: $e_i^P = e_i^A$=0.1, c_i^A=0.625 and $\mu_{ik} = \varphi_{ik}$=0.05.

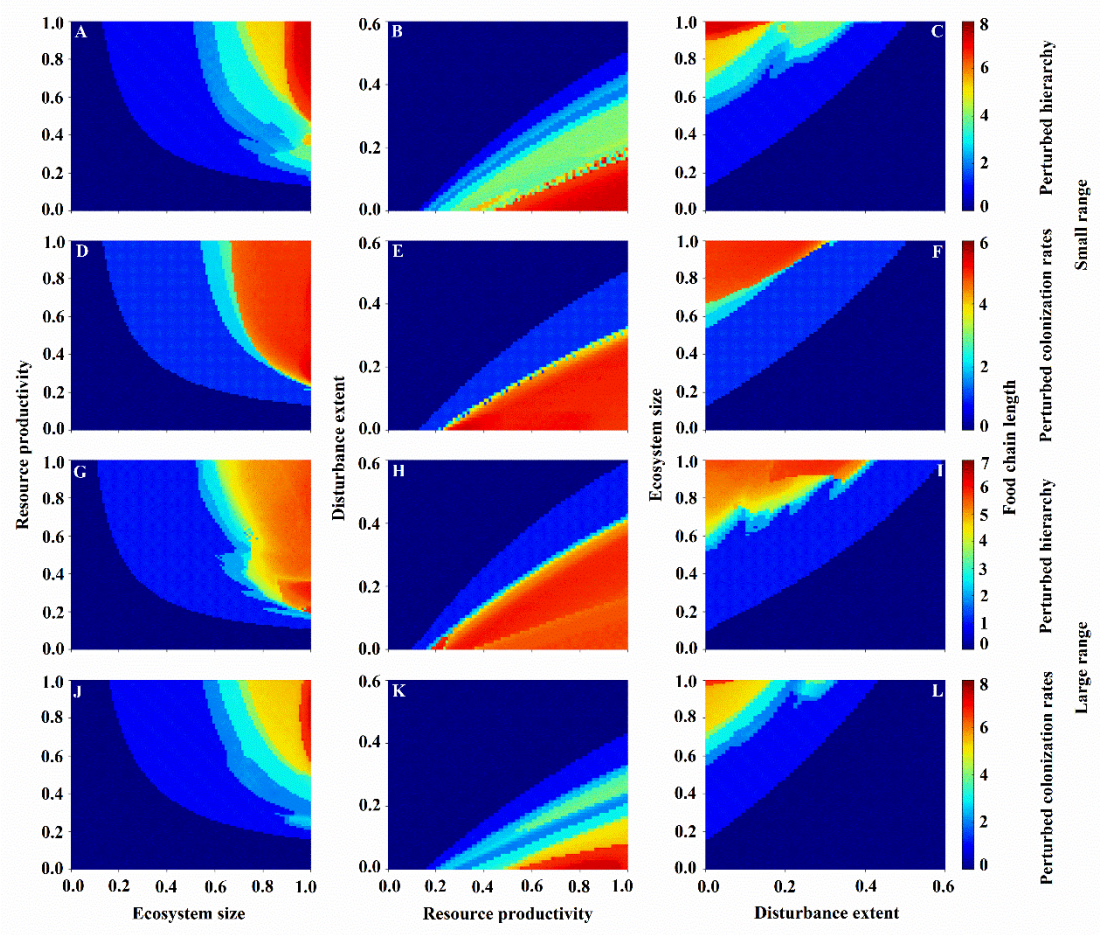


Figure S15. Interactive effects of ecosystem size (*S*), resource productivity (*R*) and disturbance extent (*D*) on FCL in a large food web as displayed in Figure S14 above. (i) Perturbing competitive hierarchy: basal species' colonization rates are evenly spaced at both small ($c_i^P \epsilon E[0.45, 0.8]$) and large ($c_i^P \epsilon E[0.25, 1]$) ranges while weakening their competitive hierarchy $\boldsymbol{H}$: the upper and lower triangular entries (H_{ij}) are uniformly sampled from $U[0.75, 1]$ and $U[0, 0.25]$, respectively. (ii) Perturbing colonization rates: basal species' colonization rates are uniformly drawn from both small ($c_i^P \epsilon U[0.45, 0.8]$) and large ($c_i^P \epsilon U[0.25, 1]$) ranges and sorted in increasing order, but with a strict competitive hierarchy $\boldsymbol{H}$ ($H_{ij} = 1$ for $i<j$ and 0 otherwise; yellow lines). Other parameters: $e_i^P = e_i^A$=0.1, c_i^A=0.625 and $\mu_{ik} = \varphi_{ik}$=0.05.

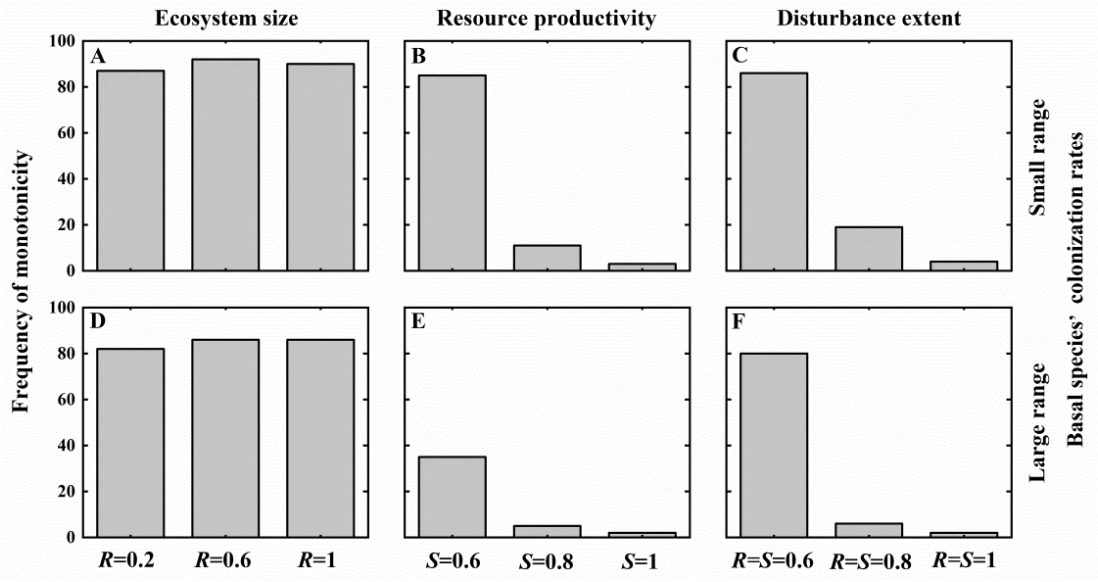


Figure S16. Frequency of the monotonic relationship between FCL and each environmental driver (ecosystem size *S*, resource productivity *R* and disturbance extent *D*) in 100 initial complex food webs simulated by the niche model (excluding those food webs with loops and cannibalism; see details in *Methods*), with varying species diversity ($10 \leq N \leq 50$), connectance ($0.05 \leq C \leq 0.25$) and basal species richness ($n_P \geq 3$). Other parameter settings: see Figure 3 in the main text.

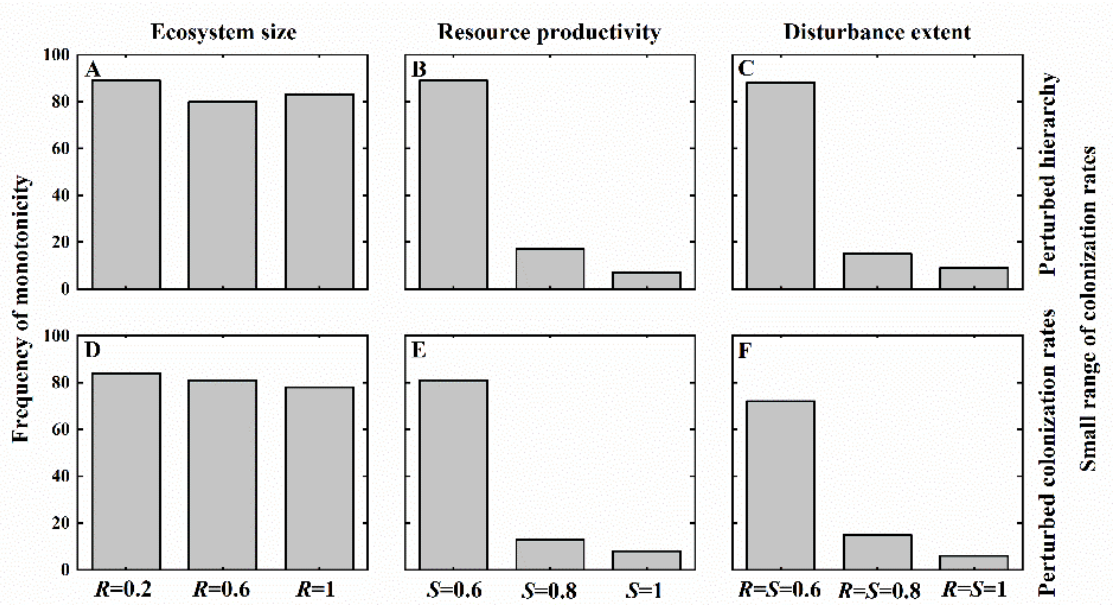


Figure S17. Frequency of the monotonic relationship between FCL and each environmental variable (including ecosystem size *S*, resource productivity *R* and disturbance extent *D*) in 100 initial food webs simulated by the niche model (excluding those food webs with loops and cannibalism), with varying species richness ($10 \leq N \leq 50$), connectance ($0.05 \leq C \leq 0.25$) and basal species richness ($n_P \geq 3$). Panels (A-C): basal species' colonization rates are evenly spaced at a small range ($c_i^P \in E[0.45, 0.8]$) while perturbing the competitive hierarchy $\boldsymbol{H}$: the upper and lower triangular entries (H_{ij}) are uniformly sampled from $U[0.75, 1]$ and $U[0, 0.25]$, respectively. Panels (D-F): basal species' colonization rates are uniformly drawn from a small range ($c_i^P \in U[0.45, 0.8]$) and sorted in increasing order, but with a strict competitive hierarchy $\boldsymbol{H}$ ($H_{ij} = 1$ for $i<j$ and 0 otherwise). Panels (A & D) R=0.2, 0.6 & 1 with D=0; panels (B & E) S=0.6, 0.8 & 1 with D=0; and panels (C & F) R=S=0.6, 0.8 & 1. Other parameters: $e_i^P = e_i^A$=0.1, c_i^A=0.625 and $\mu_{ik} = \varphi_{ik}$=0.05.

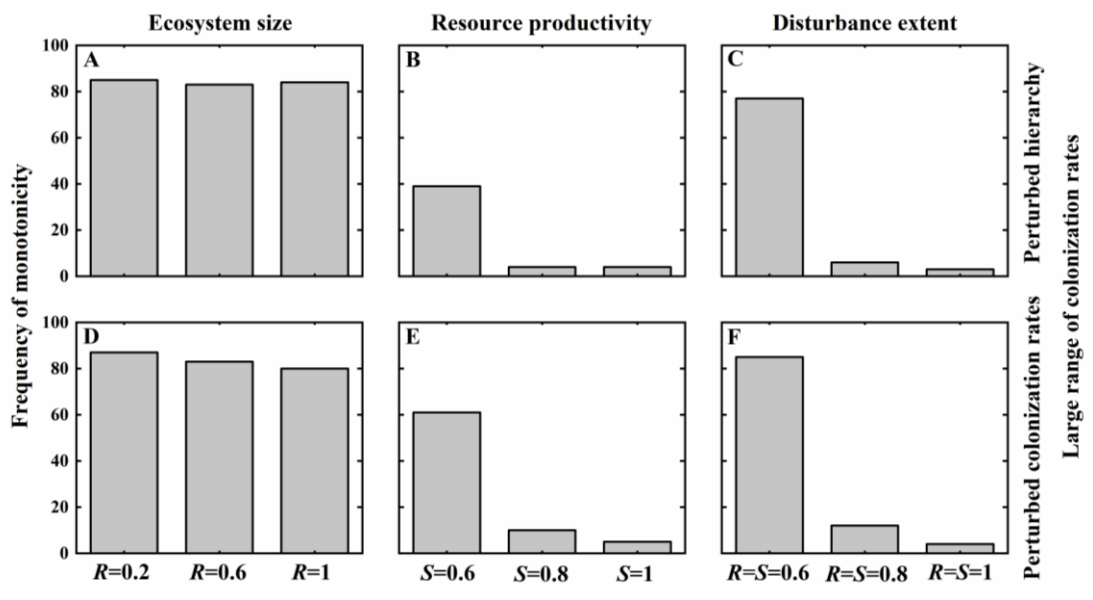


Figure S18. Frequency of the monotonic relationship between FCL and each environmental variable (including ecosystem size *S*, resource availability *R* and disturbance extent *D*) in 100 initial food webs simulated by the niche model (excluding those food webs with loops and cannibalism), with varying species richness ($10 \leq N \leq 50$), connectance ($0.05 \leq C \leq 0.25$) and basal species richness ($n_P \geq 3$). Other parameter settings are the same as in Figure S17 above, except that basal species' colonization rates are set in a large range ($c_i^P \in E[0.25, 1]$ or $c_i^P \in U[0.25, 1]$).

Supporting Information S4 – *Intransitive competition*

Generation of intransitive competition among basal species

In an n_P-species competitive community ($n_P \geq 3$) with a strict zero-sum game ($H_{ij} + H_{ji} = 1$), any pairwise competition event has a certain winner and loser ($H_{ij} = 0\ or\ 1$). This can be summarized by the coefficient of variation $C.V.(\boldsymbol{H}) = \frac{\sigma_H}{\bar{H}}$ with σ_H being the standard deviation for all elements H_{ij} and $\bar{H}$ the mean of these elements. Following Laird & Schamp (2006, 2008) and Rojas-Echenique & Allesina (2011), the degree of intransitivity can be quantified using the relative intransitivity (*RI*) index of the tournament matrix $\boldsymbol{H}$, with $RI = 1 - \frac{Var_{obs} - Var_{min}}{Var_{max} - Var_{min}}$. Here Var_{obs} denotes the variance of the row sums, $h_i = \sum_{j=1}^{n} H_{ij}$ or score sequence, of the tournament matrix $\boldsymbol{H}$. Var_{max} and Var_{min} are the maximum and minimum possible variances for the row sums of a competitive tournament matrix with the same number of species as the observed tournament matrix respectively. The minimum variance Var_{min} for the score sequence is obtained when the row sums are as uniform as possible. High row sum variance means that a few species win the majority of competitions, and hence corresponds to transitive competition. Low row sum variance means all species have similar numbers of species that they can outcompete, i.e., intransitive competition. When Var_{obs} is close to Var_{max}, a low *RI* index is obtained, indicating that transitive competition is prevalent in the community. When Var_{obs} is close to Var_{min}, a high *RI* index is obtained. Note that when $C.V.(H)$ is low, all rows are similar and thus a high *RI* index is always obtained. To obtain different values of *RI*, it is necessary to choose a zero-sum tournament matrix *H* (H_{ij} = 0 or 1) and permute it (Rojas-Echenique & Allesina, 2011). First a matrix with purely hierarchical competition is constructed (1>2>3>⋯>n_P), containing all ones above the diagonal and all zeros below the diagonal. Then, the interaction between each pair of species (*i, j*) is reversed with probability *f* (0<*f*<1), a random perturbation of the tournament matrix. By varying the probability *f*, we can yield a

broad range of *RI* values. For example, in a three-species system, we thus have *RI*=0 for a strict competitive hierarchy, while *RI*=1 for a rock-paper-scissors game.

Using this approach, we primarily generate two levels of *RI* (=0.5 and 1) for basal species competition in complex food webs, to investigate how food chain length changes along these environmental gradients (including ecosystem size, resource productivity and disturbance; see Figures S1-S7 below), while retaining the ranking of basal species by colonization rate ($c_1^P < c_2^P < \cdots < c_{n_P}^P$). As such, a global competition-colonization (C-C) tradeoff does not occur, but local C-C tradeoffs (encompassing a subset of the species) are possible.

References

Laird, R.A. & Schamp, B.S. (2006). Competitive intransitivity promotes species coexistence. *American Naturalist*, 168, 182–193.

Laird, R.A. & Schamp, B.S. (2008) Does local competition increase the coexistence of species in intransitive networks? *Ecology*, 89, 237–247.

Rojas-Echenique, J. & Allesina, S. (2011) Interaction rules affect species coexistence in intransitive networks. *Ecology*, 92, 1174–1180.

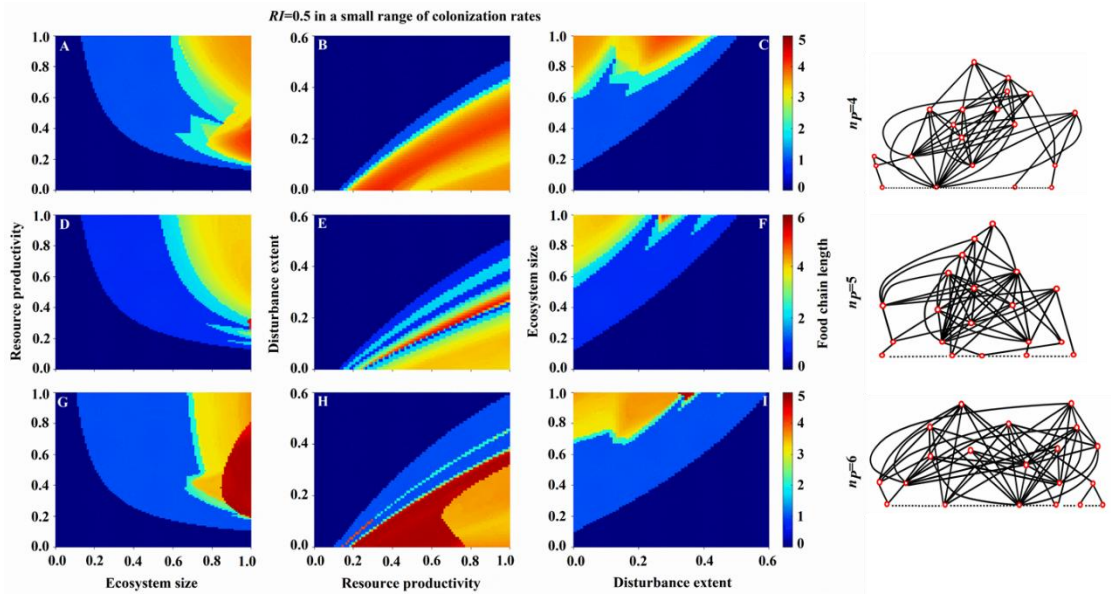


Figure S1. Interactive effects of ecosystem size (*S*), resource productivity (*R*) and disturbance extent (*D*) on FCL in given food webs on the right (with basal species richness $n_P = 4, 5$ & 6). Basal species' colonization rates are evenly spaced in increasing order at a small range ($c_i^P \epsilon E$[0.45, 0.8]), with their competitive intransitivity *RI*=0.5 ($H_{ij} = 0\ or\ 1$). Other parameters: *D*=0 in panels (A & D), *S*=1 in panels (B & E), *R*=1 in panels (C & F), all species mortality rates $e_i^P = e_i^A$=0.1, all consumers' colonization rates c_i^A=0.625 and all top-down mortality rates due to predation $\mu_{ik} = \varphi_{ik}$=0.05.

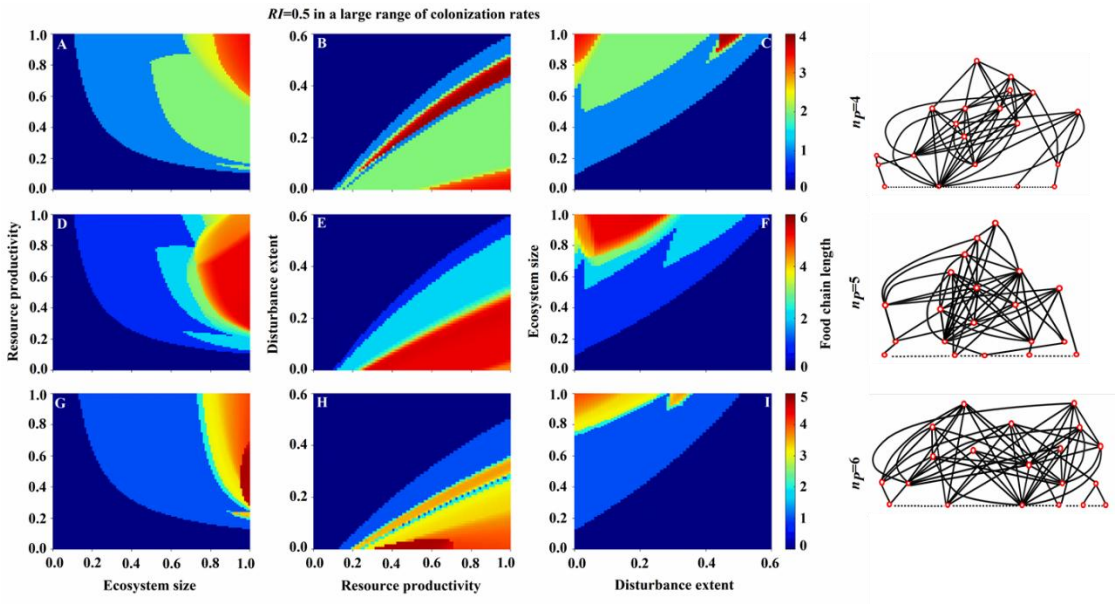


Figure S2. Interactive effects of ecosystem size (*S*), resource productivity (*R*) and disturbance extent (*D*) on FCL in given food webs on the right (with basal species richness $n_P = 4, 5$ & 6). Basal species' colonization rates are evenly spaced in increasing order at a large range ($c_i^P \epsilon E[0.25, 1]$), with their competitive intransitivity *RI*=0.5 ($H_{ij} = 0$ *or* 1). Other parameters are the same as in Figure S1 above.

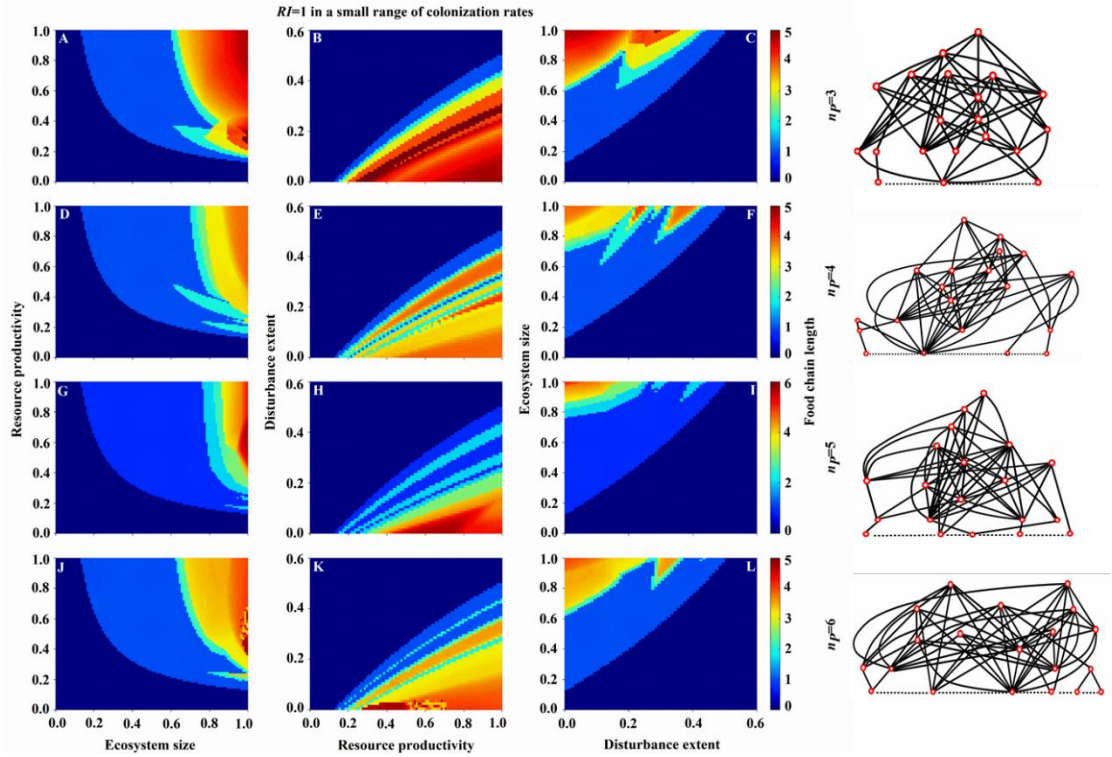


Figure S3. Interactive effects of ecosystem size (*S*), resource productivity (*R*) and disturbance extent (*D*) on FCL in given food webs on the right (with basal species richness $n_P = 3, 4, 5$ & 6). Basal species' colonization rates are evenly spaced in increasing order at a small range ($c_i^P \epsilon E[0.45, 0.8]$), with their competitive intransitivity *RI*=1 ($H_{ij} = 0\ or\ 1$). Other parameters are the same as in Figure S1 above.

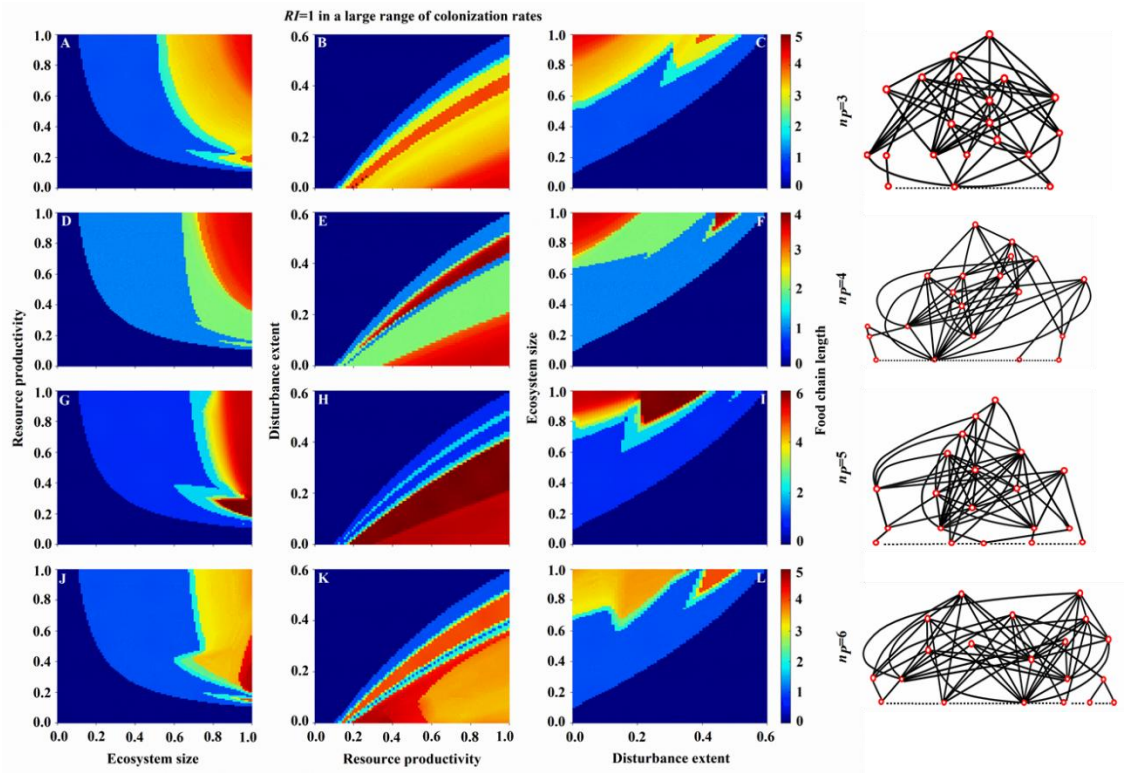


Figure S4. Interactive effects of ecosystem size (*S*), resource productivity (*R*) and disturbance extent (*D*) on FCL in given food webs on the right (with basal species richness $n_P = 3, 4, 5$ & 6). Basal species' colonization rates are evenly spaced in increasing order at a large range ($c_i^P \epsilon E[0.25, 1]$), with *RI*=1 ($H_{ij} = 0 \; or \; 1$). Other parameters are the same as in Figure S1 above.

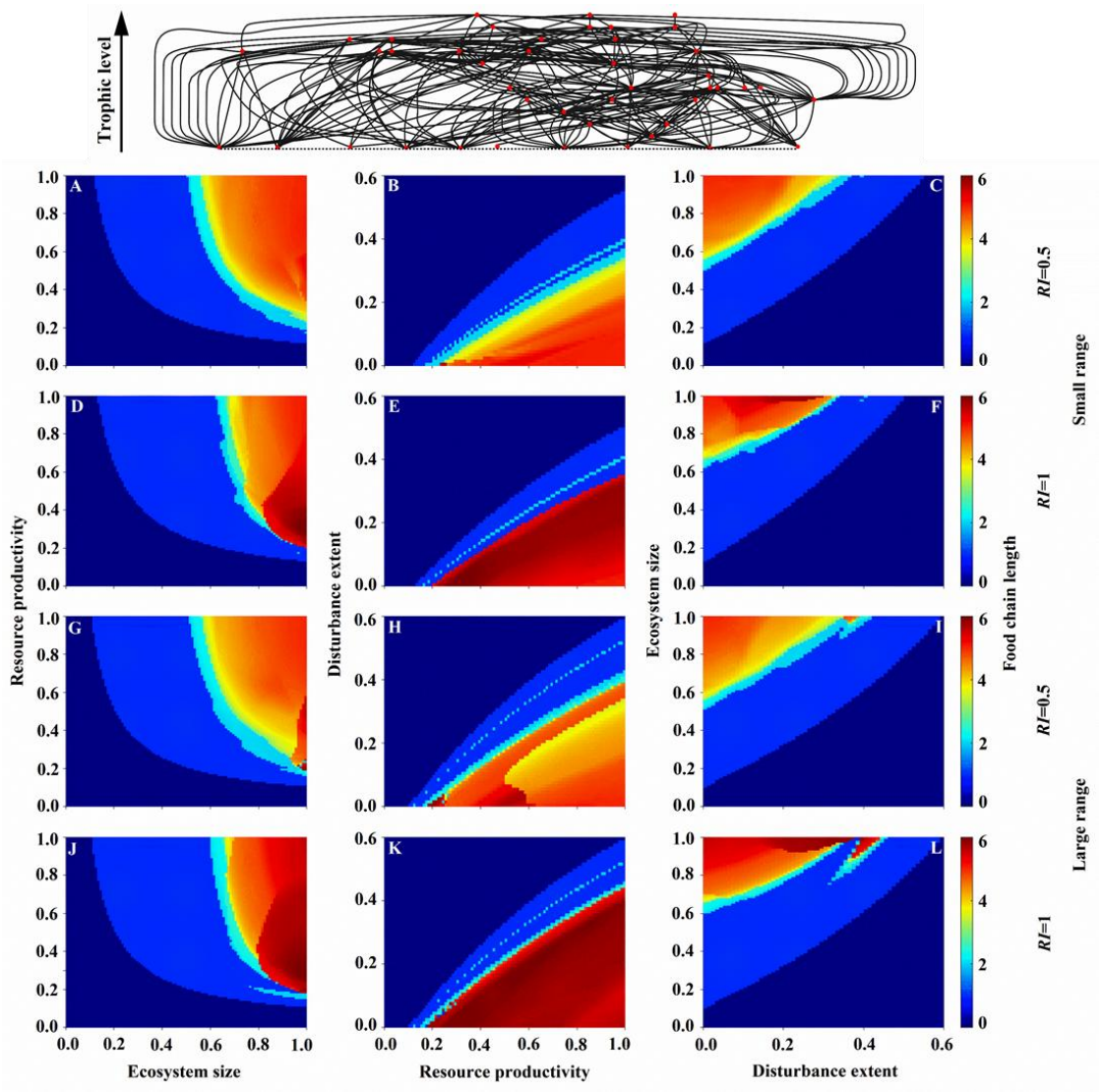


Figure S5. Interactive effects of ecosystem size (*S*), resource productivity (*R*) and disturbance extent (*D*) on FCL in a larger complex food web on the top (with species diversity *N*=44 and basal species $n_P = 10$). Basal species' colonization rates are evenly spaced in increasing order at both small (A-F: $c_i^P \epsilon E$[0.45, 0.8]) and large (G-L: $c_i^P \epsilon E$[0.25, 1]) ranges, with *RI*=0.5 or 1 ($H_{ij} = 0 \; or \; 1$). Other parameters: see Figure S1.

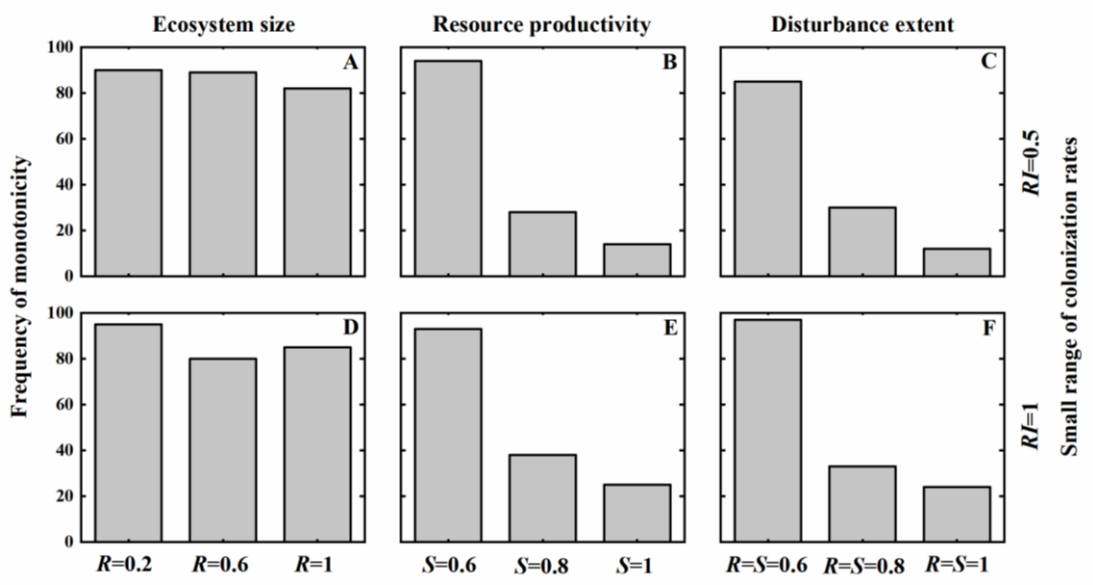


Figure S6. Frequency of the monotonic relationship between FCL and each environmental variable (including ecosystem size *S*, resource productivity *R* and disturbance extent *D*) in 100 initial food webs simulated by the niche model (excluding those food webs with loops and cannibalism), with varying species richness ($10 \leq N \leq 50$), connectance ($0.05 \leq C \leq 0.25$) and basal species richness ($n_P \geq 3$). Basal species' colonization rates are evenly spaced at a small range ($c_i^P \epsilon E[0.45, 0.8]$), with their competitive intransitivity *RI*=0.5 or 1 ($H_{ij} = 0 \; or \; 1$). Panels (A & D): *R*=0.2, 0.6 & 1 with *D*=0; panels (B & E): *S*=0.6, 0.8 & 1 with *D*=0; and panels (C & F): *R*=*S*=0.6, 0.8 & 1. Others: $e_i^P = e_i^A$=0.1, c_i^A=0.625 and $\mu_{ik} = \varphi_{ik}$=0.05.

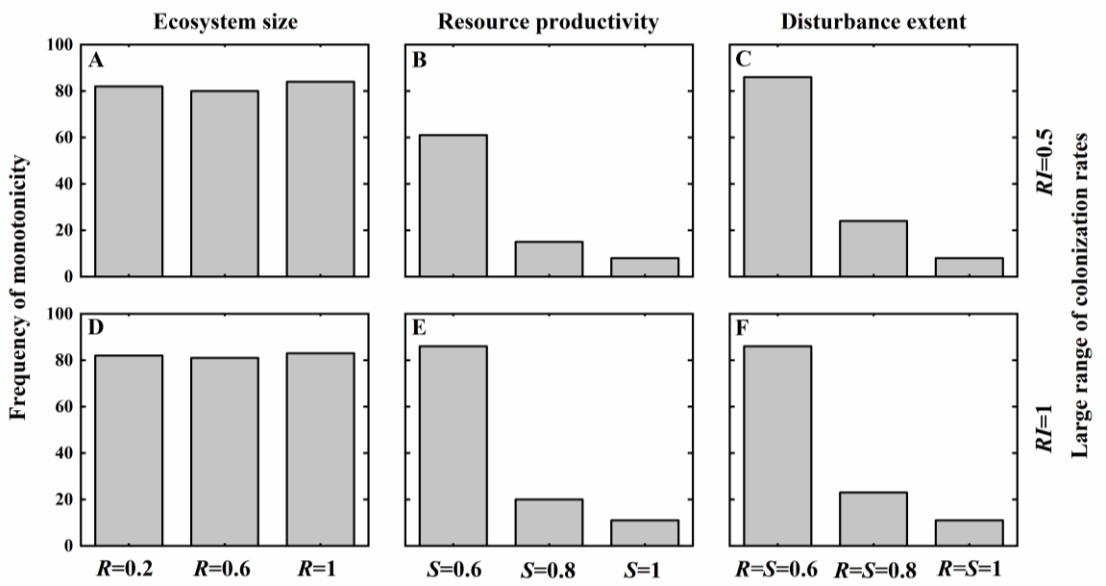


Figure S7. Frequency of the monotonic relationship between FCL and each environmental driver (ecosystem size *S*, resource productivity *R* and disturbance extent *D*) in 100 initial complex food webs simulated by the niche model (excluding those food webs with loops and cannibalism), with varying species diversity ($10 \leq N \leq 50$), connectance ($0.05 \leq C \leq 0.25$) and basal species richness ($n_P \geq 3$). Basal species' colonization rates are evenly spaced at a large range ($c_i^P \epsilon E[0.25, 1]$), with their competitive intransitivity *RI*=0.5 or 1 ($H_{ij} = 0 \; or \; 1$). Others: see Figure S6 above.