Predicting food web structure in future climates is a pressing goal of ecology. These predictions may be impossible without a solid understanding of the factors that structure current food webs. The most fundamental aspect of food web structure—the relationship between the number of links and species—is still poorly understood. Some species interactions may be physically or physiologically ‘forbidden’—like consumption by non-consumer species—with possible consequences for food web structure. We show that accounting for these ‘forbidden interactions’ constrains the feasible link-species space, in tight agreement with empirical data. Rather than following one particular scaling relationship, food webs are distributed throughout this space according to shared biotic and abiotic features. Our study provides new insights into the long-standing question of which factors determine this fundamental aspect of food web structure.

1. Introduction

Anticipating food web responses to environmental change is a pressing issue [1–4]. Yet, doing so is challenging owing to the inherent complexity of food web structure and dynamics [5,6]. Failure to account for the many underlying mechanisms—biotic and abiotic—that constrain food webs makes predicting possible future responses difficult. Gaining the fundamental understanding upon which to ground predictive models should thus be a major focus of food web ecology [7,8], but progress has proceeded slowly owing to a lack of highly resolved food web data that span across environmental conditions [9].

Despite these limitations, ecologists have uncovered many fundamental laws governing food web structure [1,10–22]. For example, larger ecosystems harbour more diverse communities with more trophic levels [23]—although the number of trophic levels is lower than expected owing to the inefficient transfer of energy and matter [24,25]. More fundamentally, the number of feeding interactions, or links ($L$), increases—unsurprisingly—with the number of species ($S$) [26,27]. Competing hypotheses have been proposed to explain this foundational pattern, yet it remains poorly understood [28], suggesting a striking lack of understanding of the most fundamental factors constraining food web structure.

According to Martinez’s constant connectance hypothesis, $L$ increases in proportion to the total possible number of species interactions, $S^2$ [29]. The minimal connectance model represents the extreme scenario where species are connected to at least one other species, resulting in $L$ increasing as $S - 1$ [29]. Last, the recent flexible links model unifies these two extremes, thus making scaling predictions within those boundaries [28]. Tests of these hypotheses suggest that the actual scaling lies somewhere in between the estimated scaling exponent of 1 predicted by minimal connectance, and 2, for the constant connectance hypothesis [28–30]. As a result, food webs should occur on a continuum, somewhere between minimally connected food webs and fully interconnected ones.
[29]. This fact rejects the notion of a one-size-fits-all scaling between L and S, but this point has received little attention.

The number and quality of available food webs makes it possible to revisit this classic debate. Furthermore, rapidly changing environmental conditions are expected to alter multiple aspects of food web structure [3,7,8,31], some of which may influence, or be influenced by, the relationship between L and S, like the proportion of apex and basal species [3] and the arrangement of feeding interactions across trophic levels (e.g. degree of omnivory, maximum trophic level, modularity) [7,8,31]. Food webs in different ecosystem types and latitudes are also likely to respond to these changes differentially [8,32]. It is therefore more important than ever to understand how these biotic and abiotic factors determine the relationship between L and S in food webs.

Additionally, we now know that phenotypic and physiological mismatches between consumers and their resources can lead to ‘forbidden interactions’ [16,33–36]. Perhaps the most fundamental type of forbidden interaction occurs in organisms that do not rely on the consumption of others to meet their energetic demands (i.e. autotrophs, chemotrophs), but it may also be common among organisms for whom predation is physically or physiologically impossible (e.g. biologically mismatched like gape-limitation in aquatic systems [25,37]). Yet, hypotheses that aim to explain the L–S relationship usually do not account for forbidden interactions, despite their potential influence on food web structure [38–40].

Here, we bridge these gaps by exploring how forbidden interactions, in particular those between non-consumers and other species, may constrain the L–S relationship. We use empirical food webs from across the globe to show that (i) rather than following a single log-linear relationship, food webs occupy a well-defined feasible space within the L–S plane with boundaries that can be predicted from theory, and (ii) food webs closer—or farther away—from those boundaries, share common biotic and abiotic features.

2. Methods

(a) L–S scaling and forbidden interactions

Food webs are represented by an adjacency matrix where 1 indicates that a species in the i-th row is consumed by one in the j-th column and 0 indicates the absence of a link (figure 1a).

Following standard practices [10,16,29,41–48] all links are directed. Directed connectance (C)—the classic metric of linkage density—is calculated as the ratio between realized links (the sum of all 1s in the adjacency matrix) and the total possible number of links (the square of the number of species, S²) (figure 1a). The constant connectance hypothesis assumes that C is constant across food webs, thus L = CS², with a maximum of L = S² when C = 1. Minimal connectance requires that L = S − 1 so that every species has at least one connection to another species.

In all food webs, there should be a fraction of species that do not consume any other species for several possible reasons. If we consider a fraction β of non-consumer species, then the total number of non-consumers is βS, while that of consumers is (1 − β)S (figure 1a). By ignoring non-consumers, classic directed connectance overestimates the maximum possible number of links within food webs. To resolve this, we calculate connectance as the fraction of realized links (L) divided by all possible links (S²) from which the number of forbidden links due to non-consumers (βS²) has been subtracted (or S² − βS², figure 1a). This results in a new expression for connectance: C = L/(S²(1 − β)), which is larger than classical connectance for β > 0. Rearranging and taking logarithms yields log(L) = log(C(1 − β)) + 2 log(S), where log(C(1 − β)) is the intercept. Assuming C = 1, we can find an upper boundary for L in L–S space that depends on 1 − β (figure 1a). Because C = 1 sets the upper limit given the constant connectance hypothesis, this new upper limit should be interpreted as the magnitude by which forbidden interactions lower the theoretical limit predicted by constant connectance, which can be estimated by quantifying β from data.

(b) Empirical food webs and data analyses

The 65 food webs used are described elsewhere [16,45] and were taken from the Interaction Web Database (https://www.nceas.ucsb.edu/interactionweb/), the GlobalWeb food web database (https://www.globalwebdb.com/) and the R package ‘cheddar’ [49]. The data span all continents, encompass terrestrial, marine, estuarine and freshwater ecosystems, and average 64 species and 427 interactions (details in electronic supplementary material, appendix S1). Data and code can be found in Dryad [50] or at https://github.com/JPGibert/FWs_LS_feasibility_space.

Because the fraction of non-consumers, β, may be over- or underestimated in empirical data, we determined a series of possible upper boundaries to L–S space of increasing likelihood based on the quantiles of the empirical distribution of β across all food webs. The ‘true’ upper boundary corresponds to the maximum possible β in nature, but owing to sampling error,
our upper boundary estimated based on the 0.95 quantile represents our best estimate of this boundary. We also tested for the existence of statistical boundaries in the L–S plane through a quantile regression with L as response and S as predictor using the R package quantreg [51]. Quantile regressions indicate the existence of boundaries in the distribution of stochastic variables [52].

Last, we assessed whether food webs that occur closer or farther away from these boundaries share common biotic and abiotic features. To do so we fitted a multiple linear regression that includes descriptors of food web structure. Biotic factors included the prevalence of omnivory (i.e. the fraction of species that feed on multiple trophic levels [53]), the maximum number of trophic levels (calculated as in [54]), the fraction of basal and top species and modularity (i.e. the existence of groups of species that more likely interact with other species within their group [55]). Abiotic factors included ecosystem type (aquatic or terrestrial) and absolute latitude. We used the log_{10} of all numerical variables except for absolute latitude and the fraction of basal species. We considered the interaction between ecosystem type and the fraction of top predators (model selection is presented in electronic supplementary material, appendix S2). To control for multicollinearity, we included variance inflation factors (VIFs) for all predictors using R package car [56] (VIF > 5 indicates strong collinearity). Owing to moderate collinearity, maximum trophic level and the fraction of basal species were dropped from the main model (see the full model in electronic supplementary material, appendix S2).

3. Results

Values of β ranged from 0.013 to 0.73 with a mean of 0.32 ± 0.03. The 5th, 25th, 50th, 75th and 95th percentiles of the distribution of β-values set boundaries to L–S space that are lower than the boundary set by constant connectance and above which no food web in our dataset occurs (figure 1b). The lower boundary of this feasible space is still set by the L = S−1 relationship (figure 1b) and all food webs fell within the limits of this space (figure 1b). The existence of a feasible space was corroborated by a significant quantile regression (figure 1c).

A model with S as the sole predictor of L accounted for 50% of observed variance within the feasible space, while a model including other biotic and abiotic variables accounted for 87% (adj. $R^2 = 0.867$; table 1). Biotic and abiotic variables thus account for all but 13% of the variation observed within the feasible space. Indeed, food webs with a higher prevalence of omnivory also were ones with more links ($b = 0.181 ± 0.023$, $p < 10^{-10}$; figure 2a), while modular food webs had fewer links ($b = -0.081 ± 0.023$, $p < 10^{-10}$; figure 2b). Food webs with more top predators had fewer links in aquatic ecosystems ($b = -0.081 ± 0.022$, $p = 0.001$; figure 2d), but more links in terrestrial ecosystems (interaction term = $0.129 ± 0.053$, $p = 0.018$; figure 2e). Among abiotic factors, food webs closer to the upper boundary also were farther away from the equator ($b = 0.053 ± 0.02$, $p = 0.011$; figure 2f), while ecosystem type did not influence the number of links (terrestrial intercept relative to aquatic = $0.042 ± 0.066$, $p = 0.526$; figure 2f). Multicollinearity precluded us from assessing whether food webs closer or farther away from the upper boundary were also similar in terms of the fraction of basal species and the number of trophic levels (VIF/no. trophic levels = 4.67, VIF/basal spp.) = 4.14; electronic supplementary material, appendix S2).

4. Discussion

Understanding how biotic and abiotic factors influence food web structure is a prerequisite to forecasting environmentally induced changes to food webs. Despite a century of research since Elton’s seminal work [57], we still do not fully understand what determines why some food webs have more interactions than others of similar richness. We show theoretically (figure 1a), empirically (figure 1b) and statistically (figure 1c), that accounting for forbidden interactions sets an upper boundary to the feasible link-species space within which all analysed food webs occur. We also show that food webs that are closer or farther away from this upper boundary share biotic and abiotic features (figure 2), suggesting underlying rules governing how food webs are structured in nature. These results amend our understanding of the processes controlling food web structure and emphasize the importance of interactions that cannot occur in nature as a factor structuring food webs [36,39].

A recent study examined several possible L–S scaling relationships, with a ‘flexible links’ model accurately predicting this scaling as well as the distribution of values in L–S space [28]. Despite this important advance, the flexible links model does not predict an upper limit to the feasible L–S plane, nor does it attribute residual variation to underlying ecological factors. We show that food webs do not occur within feasible L–S space at random. Instead, food webs

<table>
<thead>
<tr>
<th>variable</th>
<th>estimate (β)</th>
<th>s.e.</th>
<th>p-value</th>
<th>VIF</th>
</tr>
</thead>
<tbody>
<tr>
<td>$R^2 = 0.882$</td>
<td>2.45</td>
<td>0.019</td>
<td>$&lt;10^{-15}$</td>
<td>n.a.</td>
</tr>
<tr>
<td>$R_{adj}^2 = 0.867$</td>
<td>0.334</td>
<td>0.02</td>
<td>$&lt;10^{-15}$</td>
<td>1.36</td>
</tr>
<tr>
<td>$p &lt; 10^{-15}$</td>
<td>0.181</td>
<td>0.023</td>
<td>$&lt;10^{-10}$</td>
<td>1.711</td>
</tr>
<tr>
<td>modularity</td>
<td>$-0.131$</td>
<td>0.023</td>
<td>$&lt;10^{-6}$</td>
<td>1.853</td>
</tr>
<tr>
<td>latitude</td>
<td>0.053</td>
<td>0.02</td>
<td>0.011</td>
<td>1.403</td>
</tr>
<tr>
<td>fraction top predators</td>
<td>$-0.081$</td>
<td>0.022</td>
<td>0.001</td>
<td>1.703</td>
</tr>
<tr>
<td>habitat</td>
<td>$-0.042$</td>
<td>0.066</td>
<td>0.526</td>
<td>1.655</td>
</tr>
<tr>
<td>top preds × habitat</td>
<td>0.129</td>
<td>0.053</td>
<td>0.018</td>
<td>1.75</td>
</tr>
</tbody>
</table>

Table 1. Regression summary of a multiple linear model estimating the effects or biotic and abiotic variables on species richness in empirical food webs. Increasing values of variance inflation factors (VIFs) indicate increasing levels of collinearity associated with each predictor variable. $R_{adj}^2$, adjusted $R^2$; n.a., not applicable.
that occupy certain portions of this space also share a combination of biotic and abiotic features like the prevalence of omnivory or the proportion of apex predators (figure 2). We notice, however, that our analysis falls short of assigning causality to these variables, as cause and effect can be difficult to disentangle among structural features. Interestingly, however, and contrary to what was observed in terrestrial food webs, aquatic food webs showed (i) a negative correlation between the proportion of apex predators and the total number of links, and (ii) higher overall connectivity. One possible explanation is that, despite being equally productive, aquatic food webs tend to have high consumer biomass relative to terrestrial food webs (i.e. inverted biomass pyramids [58,59]), potentially leading to more omnivory and interactions. In addition to having higher connectivity, aquatic food webs also tended to be less modular, which could reduce stability, thus providing a possible explanation as to why trophic cascades are seemingly more frequent in aquatic systems.

Our findings suggest possible ways in which food webs might change with the climate. Rising temperatures are expected to decrease the proportion of apex predators and increase the proportion of basal species and primary producers [3]. Our results suggest that as the proportion of top predators decreases, aquatic food webs may become more interconnected while terrestrial ecosystems may become less interconnected (figure 2d,e). A simple numerical exploration yields mixed support for this prediction (electronic supplementary material, appendix S3), highlighting an avenue for future research. Latitude had a small but significant effect on the $L$–$S$ relationship, consistent with previous studies [45,60]. If temperature is partially responsible for the decrease in linkage density with latitude, then food webs could become less interconnected in a warmer world, but more work is needed to understand whether that is the case. Our results stress the need to revisit the fundamental factors that determine food web structure and emphasize the role of forbidden interactions in that endeavour.

Data accessibility. Data and code can be found in Dryad Digital Repository: https://doi.org/10.5061/dryad.4b8ghtc3 [50].

Authors’ contributions. J.P.G. conceived the study. J.P.G. and D.J.W. analysed the data. J.P.G. wrote the first draft and both authors contributed substantially to subsequent versions and revisions. Both authors approved the final version of the manuscript and agreed to be held accountable for the content therein.

Competing interests. We declare we have no competing interests.

Funding. J.P.G. and D.J.W. were funded through a US Department of Energy, Office of Science, Office of Biological and Environmental Research, Genomic Science Program under award no. DE-SC0020362 to J.P.G.

Acknowledgements. We thank two anonymous reviewers for their insights.

References


