ECOLOGY LETTERS

Ecology Letters, (2018) doi: 10.1111/ele.12955

LETTER

Feeding environment and other traits shape species' roles in marine food webs

Alyssa R. Cirtwill* and Anna Eklöf

Department of Physics, Chemistry and Biology (IFM) Linköping University, Linköping SE-581 83, Sweden

**Correspondence: E-mail: alyssa.cirtwill@gmail.com*

Abstract

Food webs and meso-scale motifs allow us to understand the structure of ecological communities and define species' roles within them. This species-level perspective on networks permits tests for relationships between species' traits and their patterns of direct and indirect interactions. Such relationships could allow us to predict food-web structure based on more easily obtained trait information. Here, we calculated the roles of species (as vectors of motif position frequencies) in six well-resolved marine food webs and identified the motif positions associated with the greatest variation in species' roles. We then tested whether the frequencies of these positions varied with species' traits. Despite the coarse-grained traits we used, our approach identified several strong associations between traits and motifs. Feeding environment was a key trait in our models and may shape species' roles by affecting encounter probabilities. Incorporating environment into future food-web models may improve predictions of an unknown network structure.

Keywords

Apparent competition, body mass, direct competition, feeding environment, food chain, indirect interactions, trophic level.

Ecology Letters (2018)

INTRODUCTION

Understanding how species interact in ecological communities is a great challenge for ecologists, conservationists and other stakeholders. Food webs – networks describing consumer–resource interactions – place species in their community context and can facilitate such endeavours. The food-web perspective is important because direct as well as indirect interactions between species can lead to responses to perturbations that are unpredictable when considering each species in isolation (Wootton 1994; Schmitz 1997). The occurrence of an interaction is influenced by species traits (Bartomeus *et al*. 2016); traits are therefore likely to be valuable for predicting and understanding direct and indirect interactions between species (Woodward *et al*. 2005; Pearse & Altermatt 2013). If species' traits are consistently related to their patterns of direct and indirect interactions (their food-web roles), then it may be possible to extrapolate the effects of species on their communities from their traits without knowing the full food web.

The first steps in this direction are food-web models based on predator–prey mass ratios (e.g. Woodward *et al*. 2005; Brose *et al*. 2006; Riede *et al*. 2011), but body mass alone does not fully explain the structure of food webs (Zook *et al*. 2011). Nevertheless, it has been shown that combinations of relatively few traits can predict a large part of the interactions (Eklöf *et al*. 2013). In fact, less than 10 (and usually six or fewer) dimensions, where each dimension is a potential trait-axis in the multidimensional niche-space (Chase & Leibold 2003), are needed to fully describe food-web structure (Eklöf *et al*. 2013). This result suggests that a small number of traits is sufficient to predict each species' feeding interactions although it still remains to be seen which traits have the most predictive power.

Here, we are interested in relating species' traits to their roles in networks rather than predicting the overall structure of a network. We define species' roles based on their participation in meso-scale structures called 'motifs' (Milo *et al*. 2002; Alon 2007; Stouffer *et al*. 2012). These motifs are unique arrangements of n interacting species (Alon 2007) and, for any given species, describe direct and indirect interactions with other species in the network. We focus on roles because they provide a species-level summary of a food web, unlike most commonly reported measures of network structure (e.g. connectance) that summarise network structure (Dunne 2006; Otto *et al*. 2007) and therefore obscure most of the species-level details. Each motif contains one or more unique positions (e.g. the top, middle or bottom species in a three-species chain; Fig. 1) specifying precisely which interactions each species contributes to the motif. These positions can be used to provide a higher level of detail when defining species' roles. As it is reasonable to expect that species will be most strongly affected by species with which they interact or by species that directly interact with the focal species' interaction partners (Jordán & Scheuring 2002; Jordán *et al*. 2006), we used three-species motifs as our units of analysis (Stouffer *et al*. 2012; Cirtwill & Stouffer 2015).

These three-species motifs can be considered the 'building-blocks' of food webs (Milo *et al*. 2002). Some motifs have clear ecological meanings (Bascompte & Melián 2005). For example, the omnivory or intraguild predation motif (in which one resource is consumed by two predators, one of which also consumes the other predator) is believed to contribute to network stability by moderating non-equilibrium dynamics (McCann *et al*. 1998). The three-species food chain motif, meanwhile, has been used to better understand trophic cascades (Hastings & Powell 1991; Laws & Joern 2013). Other

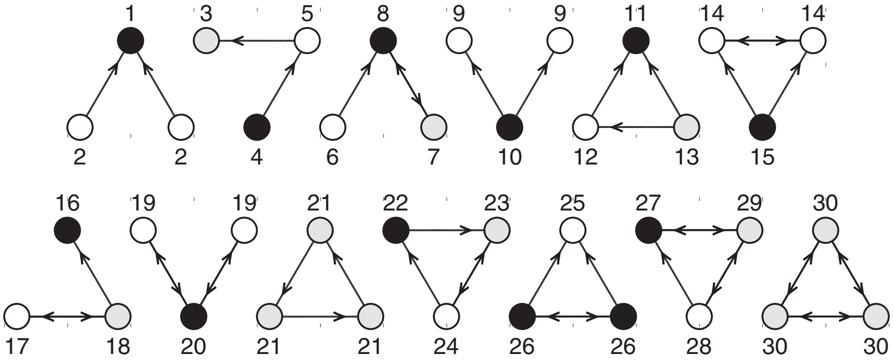

Figure 1 Set of three-species motifs with positions numbered. Motifs may include one, two or three unique positions. Within a motif, species (circles) with the same fill occupy the same position.

motifs have not yet been studied independently, but each unique arrangement of interacting species implies different consequences for the flow of energy and biomass through a food web and therefore provides information about how a focal species fits into its community (Cirtwill & Stouffer 2015). Defining a species' role this way (i.e. as a vector of frequencies with which it appears in each position within each motif) provides a holistic summary of the way in which it is embedded in its food web. We can then explore how these species-level summaries may be related to species' traits. Here, we use species' roles in six well-resolved marine food webs to identify broad traits that are strongly associated with variation in species' roles and may therefore be useful in future efforts to model network structure. We found that most of the traits we consider are related to at least one key position but that feeding environment is a particularly good candidate for inclusion in future food-web models.

METHODS

Dataset

Food webs

To test which traits explain the most variation in species' structural roles, we used a set of 6 well-resolved marine food webs that include trait data for most species. The webs contain 92–488 species and 417–15 880 predator–prey interactions and describe communities ranging from Caribbean reefs to polar communities (Table S1, Supplemental methods: food web details). By covering such a broad range of communities, these webs allowed us to identify traits associated with network structure independent of local environmental variables.

Species traits

As well as predator–prey links, each web contained trait data. The traits we use in our analyses are such that they can be easily collected either from the literature or in the field when new data are assembled. The traits are body mass (g), trophic level, feeding environment, metabolic category, mobility, feeding type and feeding mode (see Supplemental methods: trait details for details). Feeding type refers to the type of resources a species consumes while feeding mode refers to the way in which the species obtains these resources. Body mass and trophic level were measured as continuous variables while all other traits were treated categorically. These traits, which were selected based on ease of collection and applicability to the broadest possible range of species, are unlikely to be the only traits which influence the likelihood of an interaction between two species. A wide variety of traits affect foraging, vulnerability and encounter probabilities (Gravel *et al.* 2016). Due to this variety, the exact set of traits affecting any pair of species is unlikely to be known for most systems (Dormann *et al.* 2017); we therefore focus on broad traits that are likely to affect many species in many systems.

Calculating species' roles

By 'species' role', we refer to the pattern of direct and indirect interactions in which a species participates, within a food web. To describe species' roles, we began by decomposing each food web into its set of three-species motifs (Stouffer *et al.* 2012; Cirtwill & Stouffer 2015). Motifs are unique arrangements of species that describe direct and indirect interactions (Milo *et al.* 2002). Each motif can be subdivided into 1–3 unique positions (e.g. predators and prey in the direct competition motif; Fig. 1) in order to obtain a more detailed picture of the interaction pattern of a species. Over the set of 13 three-species motifs there are 30 unique positions. We can therefore define each species' role as a vector of the number of times the species appears in each of these positions. Note that each set of three interacting species is assigned to exactly one motif: that which includes all of the interactions among the triad. By extension, each species is assigned to only one position per triad of interacting species.

We were not interested in differences between roles due to some species being involved in more feeding links than others. We therefore normalised species' roles by dividing the count of each position by the total number of times the species appeared in any position. In some cases, this would result in frequencies of zero which cannot be properly modelled using logistic regressions. To correct this, we added one to the count of each position for each species before dividing by the new total count across all species. Thus, all possible positions had non-zero frequencies in the roles of each species.

Roles and trophic groups

As well as the traits described above, we expect that roles will differ between species in different trophic or taxonomic groups. The effects of trophic groups (i.e. basal resources, intermediate consumers and top predators) on species roles derive from the fact that basal resources and top predators can only occupy prey or predator positions respectively. We therefore analysed species in different trophic groups separately. If a species had observed predators but no observed prey, it was classed as a basal resource. If a species had observed prey but no observed predators, it was classed as a top predator. Species with both predators and prey were classed as intermediate consumers. Basal resources and top predators can appear in only six positions each while intermediate consumers can occupy any of the 30 positions; we defined the roles of each group based only on the positions that group could occupy.

Roles and taxonomic groups

Species in different taxonomic groups might also have different roles as these species are likely to have similar traits (e.g. physical or chemical defences, dentition, behaviour), which could influence their interactions as predators or prey. Grouping species based on high-level taxonomies provides an alternative approach to grouping species based on trophic groups. Importantly, taxonomy does not depend on the structure of the network and therefore can be used to group species where the network structure is unknown. This avoids the slightly circular reasoning of assigning species to trophic groups based on network structure and then analysing their roles within the same network. To confirm that our approach is relevant for taxonomic, as well as trophic, groups, we also analysed the roles of species in the five classes with more than 50 species in our dataset (Actinopterygii, Malacostraca, Gastropoda, Polychaeta and Aves) separately from those of other species.

Identifying key positions

To assess relationships between species' traits and their roles, we first identified the motif positions that explained the most variation in the roles of each trophic and taxonomic group. We performed a principal components analysis (PCA) to collapse the 30-dimensional (intermediate consumers and taxonomic groups) or six-dimensional (basal resources and top predators) role vectors into orthogonal axes. We then identified the three motif positions that covaried most strongly with the first two axes as 'key positions' accounting for a great deal of the variation in species' roles. These positions are important for distinguishing the roles of different species but may or may not be positions which are important in other contexts (e.g. the most common positions or those which contribute to network stability). We ran separate PCAs for each trophic and taxonomic group as key positions could vary between them.

Association of traits with positions

After identifying the key positions, we then used a model-selection approach to determine which traits were most strongly associated with the frequency of each position. We fitted models separately for the frequency of each key position in the roles of each trophic or taxonomic group. To obtain more normal distributions, we log-transformed body masses and scaled and centred body mass and trophic level, the only continuous traits in our dataset. All other traits were treated as categorical variables.

Because of the large number of variables involved, we selected the best-fitting model by building up from a null model containing only an intercept and a random effect of network. The random effect was included to account for differences in species roles both due to environmental factors and due to the different distributions of traits in different food webs. We created a first set of alternative models by adding predictors associated with each trait to the null model in turn. These alternative models took the form:

$$\psi_{ijk} \approx \beta_0 + \beta_1 trait_{xj} + N_k, \tag{1}$$

where ψ_{ijk} is the count of position i for species j in network k, β_0 is a general intercept, β_1 is the effect of the value of trait x (e.g., feeding environment) for species j and N_k is a random effect of network k. Note that each alternative model included only one trait.

Because ψ_{ijk} are frequencies bounded between 0 and 1, all models used a binomial error distribution.

We then ranked these models according to their AIC to determine which trait most improved upon the fit of the null model. We then extended the best-fitting model to create a second set of alternative models by adding the predictors associated with the traits not included in the best-fit model from the first round. We then ranked these models according to their AIC and again selected the best-fitting model as the base for the next set of alternative models. These second set of alternative models took the form

$$\psi_{ijk} \approx \beta_0 + \beta_1 trait_{xj} + \beta_2 trait_{yj} + N_k, \tag{2}$$

where all symbols are as described above except that trait x is the trait included in the best-fitting model from the first round and trait y may be any trait except for trait x.

We continued this process until adding an additional trait did not improve on the previous best-fitting model (i.e. no model had $\delta_{AIC} > 2$ relative to the previous best-fitting model). Where there were two or more 'best' models with similar AIC scores ($\delta_{AIC} > 2$), we used both models as bases for the next round of models. We fit all models using the function `glmer` (family 'binomial') from the R (R Core Team, 2016) package lmerTest (Kuznetsova *et al.* 2014). As trophic level is defined based on network structure, like trophic groups, we repeated our analysis of taxonomic groups omitting all models including trophic level. This did not substantially change our results (see Supplemental results: within classes for details).

RESULTS

Trophic groups

We first divided the species into basal resources, intermediate consumers and top predators and determined the major axes of variation in the roles of each trophic group. The first two axes explained the majority of the variance in species' roles in all cases (58.3 and 33.1%; 33.3 and 20.2%; and 60.2 and 29.8% for the first and second axes of each group respectively). In each of the trophic groups, positions in the same three motifs were strongly correlated with one or both of the major axes (Figs 2a, 3a, 4a; Fig. S2, Table S6, Supplemental results: vector loadings and mean position frequencies). For basal resources and top predators, these key motif positions denoted the bottom and top positions (respectively) in the apparent competition, three-species chain and direct competition motifs (positions 2, 4 and 10 for basal resources; positions 1, 3 and 9 for top predators). For intermediate consumers, the key positions were the consumer and resource species in the apparent competition motif (positions 1 and 2) and the consumer species in the direct competition motif (position 9). After identifying these key positions, we could search for associations between traits and the frequency of each key motif.

Basal resources

Many of the traits we considered had little or no variation among basal resources: almost all basal resources had the metabolic category, feeding mode and feeding type 'primary producer', a trophic level of 1 and mobility level 1 (sessile/floating). Only body mass and feeding environment had enough variation to permit well-fitting models. This restricted the set of possible models and is reflected in the extremely simple best-fit models for this trophic group. The best-fit models for positions 2 and 4 (resources in the apparent competition motif and three-species chain, respectively; Fig. 1) included only the network random effect.

The best-fit model for the frequency of position 10 (resource in the direct competition motif; Fig. 1) included feeding environment as well as the random effect. Pelagic species had significantly lower frequencies of position 10 than did benthic species ($\beta_{Environment:pelagic} = -4.71$, $P = 0.004$; Fig. 2). Benthopelagic species did not have significantly different frequencies of position 10 (Table S3, Supplemental results: basal resources). The frequency of position 10 also decreased with increasing body mass ($\beta_{BodyMass} = -30.2$, $P = 0.004$).

Intermediate consumers

The best-fit models for the frequencies of key motifs in the roles of intermediate consumers were more complicated than those for basal resources and all included at least three traits. This is unsurprising given the variety of both species and traits included in this trophic group and the fact that the roles of intermediate consumers nearly fill the space defined by the frequencies of positions 1, 2 and 9 (Fig. 3a). The best-fit model for the frequency of position 1 (consumer in the apparent competition motif; Fig. 1) was the simplest, including terms for body mass, feeding environment and feeding type. The frequency of this position increased with the log of body mass ($\beta_{BodyMass} = 9.75$, $P < 0.001$; Fig. 3b). The frequency of position 1 was significantly lower for benthopelagic species than other species ($\beta_{Environment:benthopelagic} = -1.19$, $P = 0.049$), but other feeding environments were not associated with significantly higher or lower frequencies of position 1 (Table S4, Supplemental results: intermediate consumers). Feeding type was also included in the best-fit model, but no levels of this term were associated with significantly different frequencies of position 1.

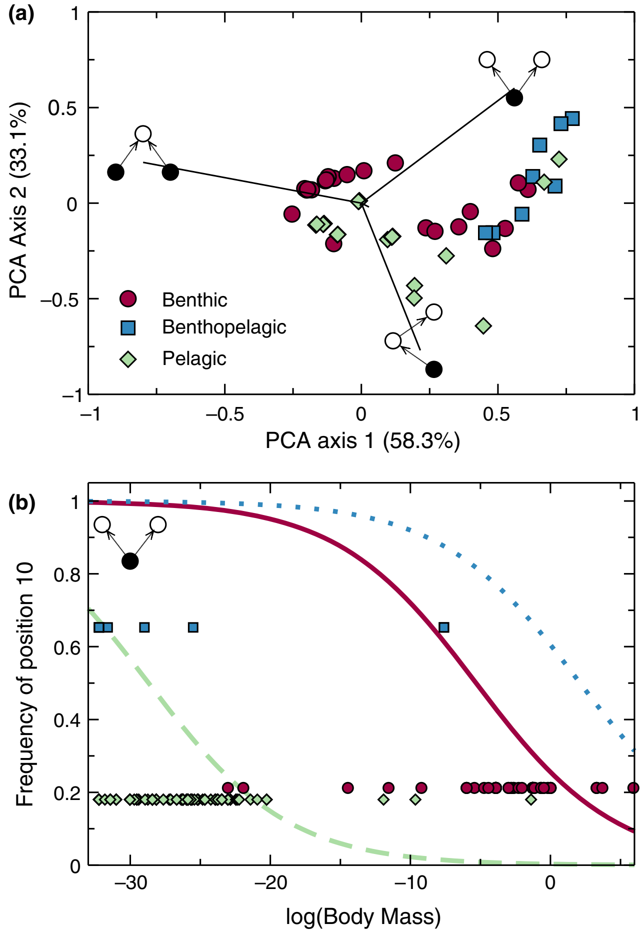

Figure 2 The roles of basal resources varied with feeding environment. (a) Here, we show the roles of each basal resource plotted against the first two principal components axes. Different feeding environments are indicated by symbol shape and size. Feeding environment was included in the best-fit model for the frequency of position 10 (resource in the direct competition motif; top right). The best-fit models for the other positions included only a random effect of network. We also show the loading of positions 2, 4 and 10 against the same axes. Positions are shown in their motif contexts with the focal positions highlighted in black. (b) The frequency of position 10 varied with feeding environment and the log of body mass. Here, we show the predicted frequency of position 10 for resources in each environment (lines). The frequency of position 10 decreases with increasing body mass and was significantly lower for pelagic resources than benthic or benthopelagic resources. We also show the observed body masses of each basal resource in our dataset (red circles for benthic resources, blue squares for benthopelagic resources and green diamonds for pelagic resources). Note that because benthic resources were generally much larger than pelagic resources the two groups ultimately had similar frequencies of position 10. Position 10 is highlighted in black in the motif depicted in the upper left.

The best-fit model for position 2 (resource in the apparent competition motif; Fig. 1) included terms for all traits except

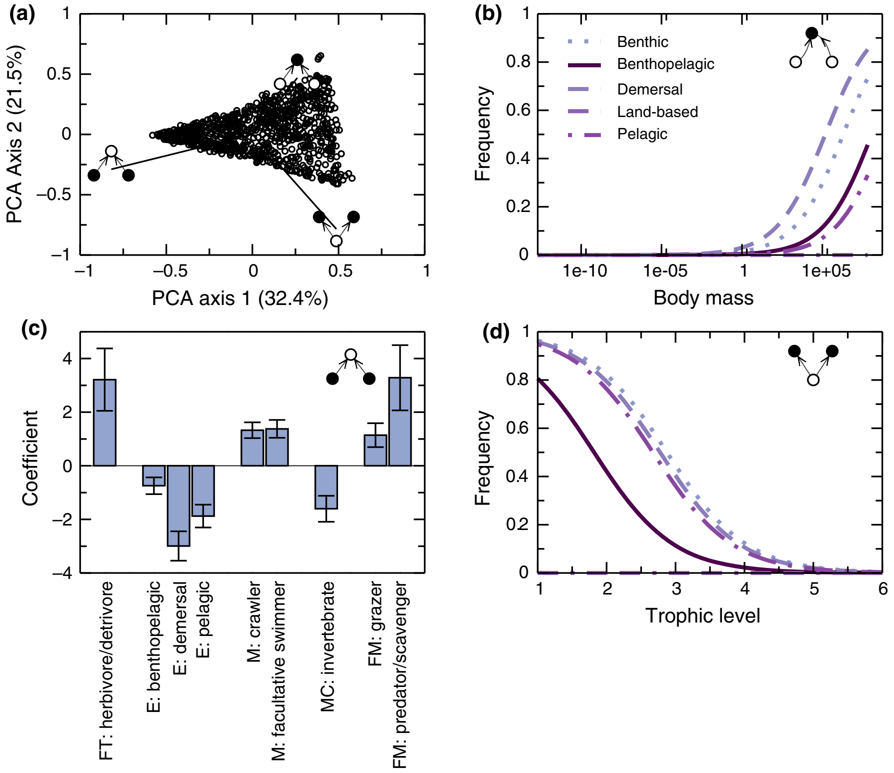

Figure 3 The roles of intermediate consumers filled the space defined by different proportions of positions 1, 2 and 9. (a) Here, we show the roles of each intermediate consumer plotted against the first two principal components axes. We also show the loading of positions 1, 2 and 9 on these same axes. Positions are shown in their motif contexts with the focal positions highlighted in black. These motif diagrams are repeated in panels (b–d). (b) The frequency of position 1 (the consumer in the apparent competition motif) increased with increasing body mass and varied with feeding environment (indicated by line type). (c) The frequency of position 2 (a resource in the apparent competition motif) decreased slightly with increasing trophic level but also varied with all other traits except for body mass. For each categorical trait, we show the coefficient (± SE) for any levels with significantly higher or lower frequencies of position 2 than the baseline. When identifying these levels, 'FT' refers to feeding type, 'E' refers to environment, 'M' refers to mobility, 'MC' refers to metabolic category and 'FM' refers to feeding mode. Baseline levels are given in the caption of Table 1. (d) The frequency of position 9 decreased with increasing trophic level and varied with feeding environment (indicated by line type), feeding mode and feeding type.

body mass. The frequency of this position decreased with increasing trophic level, but this effect was very small ($\beta_{TrophicLevel} = -0.517$, $P < 0.001$; Fig. 3c). Species with the feeding type 'herbivore/detrivore', species with intermediate levels of mobility (i.e, crawlers and facultative swimmers), and species with the feeding modes 'grazer' or 'predator/scavenger' all had higher frequencies of position 2 ($\beta_{FeedingType:herbivore/detrivore}= 3.21$, $P = 0.006$; $\beta_{Mobility:crawler}=1.33$, $P < 0.001$; $\beta_{Mobility:facultativeswimmer} = 1.38$, $P < 0.001$; $\beta_{FeedingMode:grazer} = 1.14$, $P = 0.011$; and $\beta_{FeedingMode:predator/scavenger} = 3.28$, $P = 0.007$ respectively). Conversely, species feeding in benthopelagic, demersal or pelagic environments and invertebrates had lower frequencies of position 2 ($\beta_{Environment:benthopelagic} = -0.744$, $P = 0.017$; $\beta_{Environment:demersal} =-2.99$, $P < 0.001$; $\beta_{Environment:pelagic} = -1.88$, $P < 0.001$; $\beta_{MetabolicCategory:invertebrate} = -1.60$, $P < 0.001$).

The best-fit model for position 9 (consumer in the direct competition motif; Fig. 1) included terms for trophic level, feeding mode, environment and feeding type. Like position 2, the frequency of position 9 decreased with increasing trophic level ($\beta_{TrophicLevel} = -1.37$, $P < 0.001$; Fig. 3d). Species with the feeding mode 'grazer', feeding environment 'benthopelagic' and feeding types 'herbivore/detrivore' or 'omnivore' all had significantly lower frequencies of position 9 ($\beta_{FeedingMode:grazer} = -0.983$, $P = 0.020$; $\beta_{Environment:benthopelagic} = -1.84$, $P < 0.001$; $\beta_{FeedingType:herbivore/detrivore} = -2.07$, $P = 0.005$; and $\beta_{FeedingType:omnivore} = -1.11$, $P = 0.004$ respectively). No other levels of any trait were associated with significantly higher or lower frequencies of position 9 (Table S4, Supplemental results: intermediate consumers).

Top predators

The best-fit models for key positions in the roles of top predators were much simpler than those for intermediate consumers. Trophic level was included in all three models and appears to be strongly associated with the first PCA axis (Fig. 4a). The frequencies of position 1 (consumer in the apparent competition motif; Fig. 1) and position 3 (top predator in the three-species chain motif) both increased with increasing trophic level, although this relationship was only significant for position 3 ($\beta_{TrophicLevel} = 4.44$, $P = 0.289$ and $\beta_{TrophicLevel} = 1.86$, $P < 0.001$ for positions 1 and 3 respectively; Fig. 4b). There was a single species with an unusually high relative frequency of position 1 (the consumer in the apparent competition motif). After removing this potential outlier and repeating our analyses, the best-fit model for

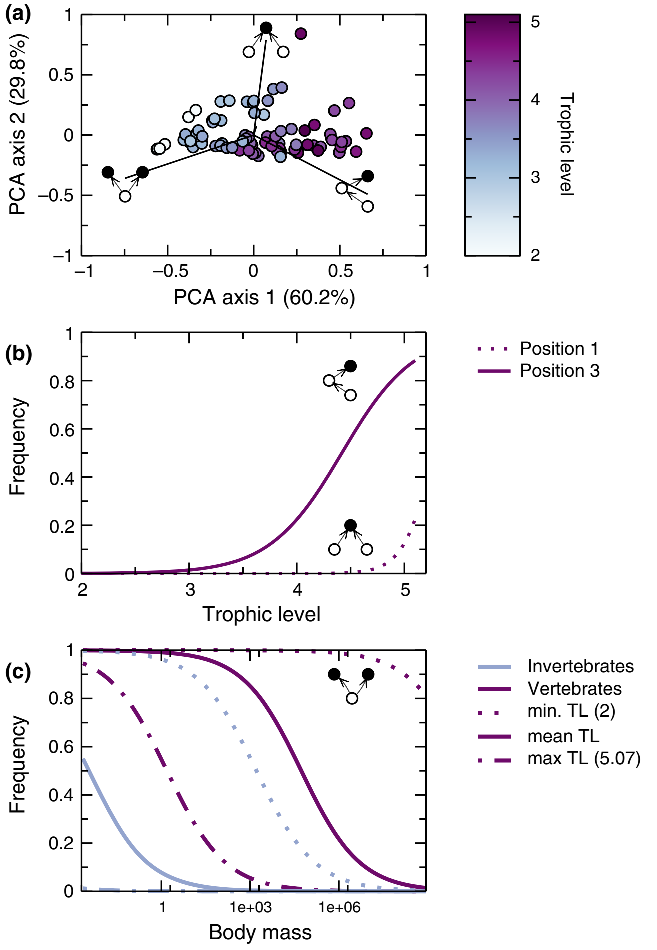

Figure 4 The best-fit models for all three key positions in the roles of top predators included trophic level. (a) Here, we show the roles of each top predator plotted against the first two principal components axes (PCAs). Symbol fill indicates trophic level. Note that trophic level displays a clear gradient along the first PCA axis. We also show the loading of positions 1, 3 and 9 on the same axes. Positions are shown in their motif contexts with the focal positions highlighted in black. These motif diagrams are repeated in panels (b and c). (b) The frequency of positions 1 and 3 (consumer in the apparent competition motif and top predator in the three-species chain) increased with increasing trophic level, although this increase was not significant for position 1. Here, we show predicted frequencies of position 1 (dotted line) and position 3 (solid line) based on the fixed effects in the best-fit models, over the range of trophic levels observed for top predators in our dataset. (c) The frequency of position 9 (consumer in the direct competition motif) was related to body mass and metabolic category as well as trophic level. The frequency of this position decreased with increasing trophic level (indicated by line colour) and increasing body mass. For a given body mass and trophic level, invertebrates (solid lines) had lower frequencies of position 9 than other metabolic categories (dotted lines for ectotherm vertebrates, dashed lines for endotherm vertebrates). Here, we show the predicted frequencies of position 9 for species in each metabolic category with minimum, mean and maximum trophic levels, over the range of body masses observed for top predators in our dataset.

position 1 included only the network-level random effect (Supplemental results: top predators).

The frequency of position 9 (consumer in the direct competition motif; Fig. 1), in contrast, decreased with increasing trophic level ($\beta_{TrophicLevel} = -2.09$, $P < 0.001$). The best-fit model for the frequency of position 9 also included terms for body mass and metabolic category. The frequency of position 9 decreased as the log of body mass increased ($\beta_{BodyMass} = -11.8$, $P = 0.003$; Fig. 4c) and was also lower for invertebrates than for ectotherm vertebrates ($\beta_{MetabolicCategory:invertebrate} = -5.49$, $P = 0.013$; Table S5, Supplemental results: top predators).

Taxonomic groups

We repeated our analyses for the five most species-rich classes (the only classes with > 50 species). One class, Polychaeta, was only represented by intermediate consumers. All other classes included both top predators and intermediate consumers. Metabolic category was constant within each class, all Actinopterygii and all Aves had the same mobility, and all but one Polychaeta shared the same feeding environment. These traits were therefore excluded from our models for the relevant classes. The first two PCA axes explained the majority of variation in each case (Fig. 5). While the key positions varied between classes, all positions were in the same three motifs highlighted by our trophic-group analyses: apparent competition, direct competition, and three-species chains (Fig. 5). The traits retained in the best-fit models also varied between classes (Table 1), but feeding environment and trophic level were commonly included (Supplemental results: within classes).

DISCUSSION

Analyses within trophic groups

Our results show that for all three tropic groups, key positions were within the same motifs: apparent competition, direct competition and the three-species chain. For basal resources, the best-fit models relating traits to these positions were very simple. The most complex model, that for position 10, included body mass and feeding environment. Specifically, the frequency of position 10 decreased with increasing body mass and, all else being equal, was higher for pelagic resources than benthic or benthopelagic resources. As pelagic and benthopelagic resources tended to be smaller than benthic resources such as macroalgae, the net result was that benthopelagic resources had the highest frequencies of position 10. Ecologically, this may reflect the higher vulnerability of benthopelagic resources which are within reach of both pelagic and benthic consumers (Fig. S1, Supplemental results: basal resources). The dearth of traits retained in our other best-fit models for basal resources suggests that their roles are best predicted using different traits than we considered here – possibly traits related to physical or chemical defences against herbivores (Hay & Fenical 1988). Such traits, classified as vulnerability traits by Gravel *et al.* (2016), could affect species' roles if, for example, heavily defended species tend to be consumed by specialised herbivores with few other prey and hence mainly appear in three-species chains.

Our results for intermediate consumers were more complex: the best-fit models each included several traits. This is not surprising as intermediate consumers were the most numerous

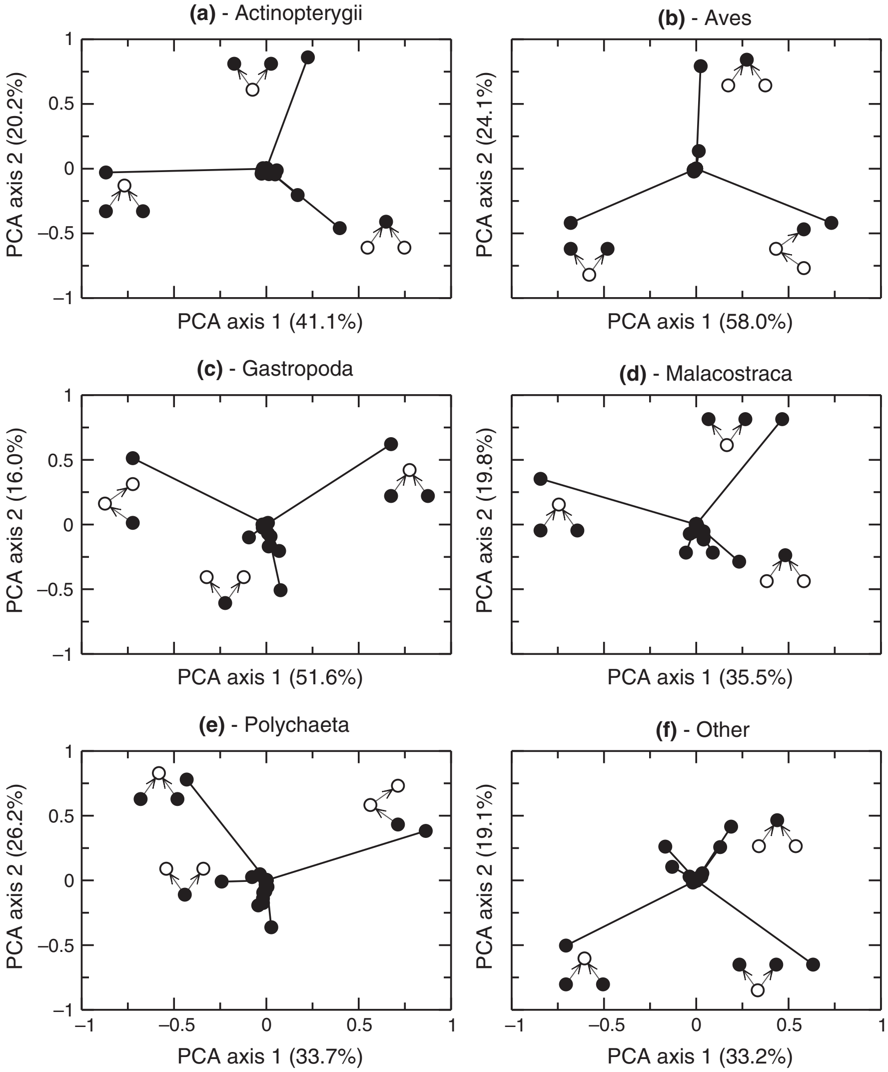

Figure 5 The key motif positions most strongly associated with the major axes in variation between species' roles varied by class, but all positions were part of the apparent competition, direct competition or three-species chain motifs. Here, we show the loadings of each motif position on the first two PCA axes of variation in the roles of the five most species-rich classes in our dataset, as well as for species in all other classes. Each PCA was conducted separately.

and taxonomically diverse trophic group in our dataset. Subdividing this group based on taxonomy or a trait of particular interest might, in future studies, yield clearer trends. Among the large number of traits included in the best-fit models, we note that feeding environment and feeding type were included in the models for all three key positions. Both of these traits, foraging traits in the framework of Gravel *et al.* (2016), may be particularly useful when modelling trophic interactions. As with basal resources, the feeding environments used by intermediate consumers may affect the numbers and sets of predators and prey that they encounter, thereby shaping their roles.

The possibility that species' feeding environments have strong effects on their roles echoes earlier work using fishes' body masses and foraging depths to predict interactions (Gravel *et al.* 2013). In both cases, the probability of two species interacting depends upon their encounter probabilities (Bartomeus *et al.* 2016), with species feeding in similar environments being more likely to encounter each other and hence more likely to interact (Gravel *et al.* 2013, 2016). For example, a benthic consumer is more likely to encounter, and prey upon, another benthic species than it is to prey upon a pelagic species it will rarely or never encounter. Feeding type (e.g. herbivore) was also included in all models for key positions in the roles of intermediate consumers, suggesting that the type of resources a species consumes likewise shapes its role. Herbivores in particular had unusual roles perhaps because they interact with basal resources which themselves have unique roles.

Our results for top predators were simpler, with trophic level included in the best-fit models for all three key positions. Trophic level appears to correspond to the first PCA axis for these species and was therefore related to the frequencies of positions 3 and 9, which were strongly correlated with this axis. The inclusion of trophic level in the best-fit model for the frequency of position 1, which was correlated with the second PCA axis, appears to have been due to the influence of a single outlier and may not reflect a general trend

Table 1 Traits retained in the best-fit models for each trophic group and class analysed

Group or class	Position	Traits included in best-fit model
Basal resources	2	None
	4	None
	10	Body mass, feeding environment
Intermediate consumers	1	Body mass, feeding environment, feeding type
	2	Feeding environment, feeding mode, feeding type, metabolic category, mobility, trophic level
	9	Feeding environment, feeding mode, feeding type, trophic level
Top predators	1	Trophic level
	3	Trophic level
	9	Body mass, metabolic category, trophic level
Actinopterygii	1	Body mass, feeding environment, trophic level
	2	Trophic level
	9	Feeding environment, feeding type, trophic level
Aves	1	None
	3	Trophic level
	9	Feeding type, trophic level
Gastropoda	2	Feeding environment
	9	Feeding environment, feeding type, trophic level
	10	None
Malacostraca	1	None
	2	Body mass, feeding mode
	9	Feeding environment, feeding mode, trophic level
Polychaeta	2	Trophic level
	9	Trophic level
	10	None

Note that each model also included a random effect of network. Benthic species were the baseline level for feeding environment, deposit-feeders were the baseline level for feeding mode, carnivores were the baseline level for feeding type, ectotherm vertebrates were the baseline level for metabolic category, and level 1: sessile/passive floater was the baseline level for mobility. Body mass and trophic level were treated as continuous traits. For a complete list of levels for each trait, see Table S2.

(Supplemental results: top predators). Species with higher trophic levels appeared more frequently at the tops of food chains (position 3) and less frequently as direct competitors (position 9). This is possibly because species with higher trophic levels have more of the food web ‘below’ them and therefore more chains available. Species with higher trophic levels may also be more able to exclude competitors than those at lower trophic levels (which appeared in more competition motifs in our dataset). As body size was also included in the best-fit model for position 3 (smaller top predators appeared more frequently in the direct competition motif), both possibilities seem likely. In our dataset, some of these small, low trophic-level top predators are herbivore/detrivores (i.e. *Abatus* spp.) that likely share prey with intermediate consumer herbivore/detrivores as well as other top predators. These herbivore/detrivore predators also consume basal resources as a large proportion of their prey. This creates many chains of length two and eliminates the possibility of motifs like the three-species chain. These exclusions must increase the frequencies of the competition motifs, which require only two trophic levels, and decrease the frequency of the three-species chain motif because of the zero-sum nature of relative motif frequencies.

Taxonomic groups

We also tested for relationships between the traits and roles of species in the five most species-rich classes in our datasets. High-level taxonomic groups (e.g. class or phylum) have previously been shown to predict species’ interactions (Eklöf *et al.* 2011), suggesting that we might find clearer relationships between species’ traits and roles within a family than across many families. A taxonomic approach also avoids the problems of first using network structure to define trophic groups and then analysing species’ roles as defined using that same structure. Notably, all key positions belonged to the same three motifs as those in our analyses by trophic group – apparent competition, direct competition and the three-species chain – although the set of positions most strongly associated with variation in the roles of each class differed.

Positions within the omnivory motif were not identified as ‘key positions’ in the roles in any class. Despite the fact that the omnivory motif is over-represented in food webs (Stouffer *et al.* 2007) and may contribute to stability (McCann & Hastings 1997; Emmerson & Yearsley 2004), it was not among the most common or most variable positions in our dataset (Table S6, Supplemental results: vector loadings and mean position frequencies). Instead, it appears that three-species chains and the two competition motifs contain the most important positions for distinguishing the roles of species within a group (trophic or taxonomic). This reinforces the instincts of researchers who have previously studied some of these motifs (e.g. Hastings & Powell 1991; McCann *et al.* 1998; Laws & Joern 2013) and suggests that they are worthy of further study. In particular, trait-matching within these motifs may differ such that we can predict which motif a given trio of species will form. Feeding environment and trophic level were retained in many best-fit models for classes, suggesting that these traits may be of interest in such studies. The inclusion of feeding environment reinforces our results for trophic groups in suggesting that the use of different environments strongly shapes species’ roles. Trophic level, as in our models for top predators, may affect the types of predators and prey a focal species encounters, again shaping its role. Although trophic level is defined based on network structure in our dataset, there is good agreement between trophic levels derived from network structure and those obtained using stable isotopes (Williams & Martinez 2004; Carscallen *et al.* 2012), allowing for trophic level to be a useful trait even when network structure is not known.

Our within-class analyses also included many best-fit models which included only the network-level random effect. As with the basal resources, this likely indicates that traits we did not consider here have large effects on species’ roles. Defensive shells, spines or camouflage, for example, might strongly affect species’ interaction partners and hence their roles (Gravel *et al.* 2016). Where the traits that determine interactions are difficult to predict (Dormann *et al.* 2017), restricting analyses to within a class may be helpful in future studies. This is because examining species within a single class should reduce some of the ‘hidden variables’ leading to different roles between classes and yield clearer signals for the traits of interest. Moreover, those with expert knowledge about a given class may be able to suggest plausible traits for testing that are likely to affect the roles of species within the focal class but not necessarily species in other classes. Importantly, these benefits do not come at the expense of simplicity. Although resolving fine-scale taxonomic groups can be very difficult, high-level taxa such as classes are straightforward to identify in most cases, facilitating analyses of novel species.

CONCLUSION

Positions in the same three motifs – three-species chain, apparent competition and direct competition – were associated with the major axes of variation in the roles of each trophic group and class we considered. This suggests that these motifs will reward further study, particularly with regard to how trait-matching may vary between motifs. Feeding environment consistently emerged as a particularly important trait in structuring patterns of interactions. This is likely because, in order to interact, species first must co-occur in the same environment (Bartomeus *et al.* 2016). Our results thus support earlier work in emphasising the importance of including co-occurrence in food-web models and suggest that traits influencing co-occurrence are important in structuring species’ roles.

ACKNOWLEDGEMENTS

This study was supported by a Formas grant (#942-2015-1262) to AE. We thank Ute Jacob for sharing the Lough Hyne and Kongsfjorden datasets, and György Barabás and Jonatan Årevall for providing feedback on the manuscript.

AUTHORSHIP

AE and AC designed the study. AC carried out the analyses and wrote the first draft of the manuscript. Both authors contributed substantially to revisions.

DATA ACCESSIBILITY STATEMENT

Data available from the Dryad Digital Repository: http://doi.org/10.5061/dryad.1mv20r6

REFERENCES

Alon, U. (2007). Network motifs: theory and experimental approaches. *Nat. Rev. Genet.*, 8, 450–461.

Bartomeus, I., Gravel, D., Tylianakis, J.M., Aizen, M.A., Dickie, I.A. & Bernard-Verdier, M. (2016). A common framework for identifying linkage rules across different types of interactions. *Funct. Ecol.*, 30, 1894–1903.

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

Brose, U., Jonsson, T., Berlow, E.L., Warren, P., Banasek-Richter, C., Bersier, L.F. *et al.* (2006). Consumer-resource body-size relationships in natural food webs. *Ecology*, 87, 2411–2417.

Carscallen, W.M.A., Vandenberg, K., Lawson, J.M., Martinez, N.D. & Romanuk, T.N. (2012). Estimating trophic position in marine and estuarine food webs. *Ecosphere*, 3, 1–20.

Chase, J.M. & Leibold, M.A. (2003). *Ecological Niches*. University of Chicago Press, Chicago.

Cirtwill, A.R. & Stouffer, D.B. (2015). Concomitant predation on parasites is highly variable but constrains the ways in which parasites contribute to food web structure. *J. Anim. Ecol.*, 84, 734–744.

Dormann, C.F., Fründ, J. & Schaefer, H.M. (2017). Identifying causes of patterns in ecological networks: opportunities and limitations. *Annu. Rev. Ecol. Evol. Syst.*, 48, 559–584.

Dunne, J.A. (2006). The network structure of food webs. In: *Ecological Networks: Linking Structure to Dynamics in Food Webs* (eds Pascual, M. & Dunne, J.A.). Oxford University Press, New York, chap. 2, pp. 27–86.

Eklöf, A., Helmus, M.R., Moore, M. & Allesina, S. (2011). Relevance of evolutionary history for food web structure. *P Roy Soc. Lond B Bio.*, 279, 1588–1596.

Eklöf, A., Jacob, U., Kopp, J., Bosch, J., Castro-Urgal, R., Chacoff, N.P. *et al.* (2013). The dimensionality of ecological networks. *Ecol. Lett.*, 16, 577–583.

Emmerson, M. & Yearsley, J.M. (2004). Weak interactions, omnivory and emergent food-web properties. *P Roy Soc. Lond B Bio.*, 271, 397–405.

Gravel, D., Poisot, T., Albouy, C., Velez, L. & Mouillot, D. (2013). Inferring food web structure from predator-prey body size relationships. *Methods Ecol. Evol.*, 4, 1083–1090.

Gravel, D., Albouy, C. & Thuiller, W. (2016). The meaning of functional trait composition of food webs for ecosystem functioning. *Phil Trans B*, 371, 20150268.

Hastings, A. & Powell, T. (1991). Chaos in a three-species food chain. *Ecology*, 72, 896–903.

Hay, M.E. & Fenical, W. (1988). Marine plant-herbivore interactions: the ecology of chemical defense. *Annu. Rev. Ecol. Syst.*, 19, 111–145.

Jordán, F. & Scheuring, I. (2002). Searching for keystones in ecological networks. *Oikos*, 99, 607–612.

Jordán, F., Liu, W., Davis, A.J. & Memmott, J. (2006). Topological keystone species: measures of positional importance in food webs. *Oikos*, 112, 535–546.

Kuznetsova, A., Bruun Brockhoff, P. & Haubo Bojesen Christensen, R. (2014). lmerTest: Tests for random and fixed effects for linear mixed effect models (lmer objects of lme4 package). R package version 2.0-11.

Laws, A.N. & Joern, A. (2013). Predator-prey interactions in a grassland food chain vary with temperature and food quality. *Oikos*, 122, 977–986.

McCann, K. & Hastings, A. (1997). Re-evaluating the omnivory-stability relationship in food webs. *P Roy Soc. Lond B Bio.*, 264, 1249–1259.

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

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

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

Pearse, I.S. & Altermatt, F. (2013). Predicting novel trophic interactions in a non-native world. *Ecol. Lett.*, 16, 1088–1094.

R Core Team (2016). R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Riede, J.O., Brose, U., Ebenman, B., Jacob, U., Thompson, R., Townsend, C.R. *et al.* (2011). Stepping in Elton’s footprints: a general scaling model for body masses and trophic levels across ecosystems. *Ecol. Lett.*, 14, 169–178.

Schmitz, O.J. (1997). Press perturbations and the predictability of ecological interactions in a food web. *Ecology*, 78, 55–69.

Stouffer, D.B., Camacho, J., Jiang, W. & Amaral, L.A.N. (2007). Evidence for the existence of a robust pattern of prey selection in food webs. *P Roy Soc B-Biol Sci*, 274, 1931–1940.

Stouffer, D.B., Sales-Pardo, M., Sirer, M.I. & Bascompte, J. (2012). Evolutionary conservation of species' roles in food webs. *Science*, 335, 1489–1492.

Williams, R.J. & Martinez, N.D. (2004). Limits to trophic levels and omnivory in complex food webs: theory and data. *Am. Nat*., 163, 458–468.

Woodward, G., Ebenman, B., Emmerson, M., Montoya, J.M., Olesen, J.M., Valido, A. *et al*. (2005). Body size in ecological networks. *Trends Ecol. Evol*., 20, 402–409.

Wootton, J.T. (1994). The nature and consequences of indirect effects in ecological communities. *Annu. Rev. Ecol. Syst*., 25, 443–466.

Zook, A.E., Eklof, A., Jacob, U. & Allesina, S. (2011). Food webs: ordering species according to body size yields high degree of intervality. *J. Theor. Biol*., 271, 106–113.

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.

Editor, Ulrich Brose
Manuscript received 15 January 2018
First decision made 18 February 2018
Manuscript accepted 4 March 2018