Simplifying species interaction matrices by grouping: optimal groupings differ between effects and responses

Christopher R. P. Brown[1,2] and Daniel B. Stouffer[1,3]

[1]Centre for Integrative Ecology, School of Biological Sciences, University of Canterbury, Christchurch, New Zealand

[2]School of the Environment, University of Queensland, Brisbane, Queensland, Australia

[3]Department of Evolutionary and Integrative Ecology, Leibniz Institute of Freshwater Ecology and Inland Fisheries (IGB), Berlin, Germany

Abstract

Most dynamic models of ecological communities require many parameters, making them expensive to fit to experimental or observational data. To reduce the number of parameters, species are often divided into groups *a priori*, using phylogenetic or functional groups, and species within a group are assumed to behave identically. Here, we assess the validity of such grouping by deriving the optimal groupings *a posteriori* based on the fit of the Beverton–Holt model to experimental data from six species of *Drosophila* competing in laboratory conditions. We find evidence that grouping can indeed improve the parsimony of species interaction models, but we also observe that there is no single optimal grouping. Further, those groupings that did prove beneficial could not have been easily predicted from available species attributes, calling into question the prevailing method of *a priori* species grouping. Lastly, we find that species did not group equivalently in their competitive effect and in their response to competition, invalidating a key assumption often made implicitly when defining species groupings. Our study suggests the need for a different, more data-informed, approach to grouping in species interaction models.

Key words: competition; ecological equivalence; grouping; model fitting; parsimony; Drosophila

1 Introduction

Models have long been used in community ecology (Lotka 1925). They allow us to make predictions about things we cannot directly observe, be it the response of a community to novel pressures (Novak et al. 2016) or the stability of an entirely novel assemblage (Maynard et al. 2020). The ability to make these predictions is becoming increasingly important as global change accelerates, introducing alien species and applying climatic or other anthropogenic pressures to communities worldwide (Sutherst et al. 2007; Gilman et al. 2010).

Making practical use of a model requires experimentally or observationally determining the values of each of the model's parameters. This poses a particular problem for models that include species interactions due to the sheer number of parameters that must be determined. Even the

simplest models with no environmental dependence require one parameter to characterise the interaction of each pair of species (Rocha Filho et al. 2005; Novak et al. 2016). Hence the number of parameters scales quadratically with the number of species; at least 100 parameters are required for the interactions in a community of "just" 10 species. Determining these parameters experimentally requires at least one experiment per parameter, and preferably many more (Hart et al. 2018). This is understandably expensive and requires substantial effort. As such, model-based predictions of real communities are typically restricted to a small number of focal species (Ovaskainen et al. 2017; Maynard et al. 2020). However, real communities can contain hundreds of species (Cornell 1999), or even thousands for microbial communities assessed by genetic metabarcoding (Warton et al. 2015). To bridge this gap between the size of real communities and our ability to parameterise models for them, we need an ecologically-defensible way to reduce the number of parameters that predictive models require.

One simplifying approach is to assume that most pairs of species do not interact or interact weakly, and hence to set many of the parameters to zero (Mutshinda et al. 2009; Weiss-Lehman et al. 2022). However, weak interactions have been shown to be very important to the stability of ecological communities (McCann et al. 1998; Neutel et al. 2002; Wootton and Stouffer 2016). Setting any individual parameter to zero thus risks losing important information; moreover, we often do not know with confidence which parameters are likely to be zeros and which are not. Another approach assumes that species fall into groups, such that all species in the same group interact with species in other groups in an identical fashion, thereby collapsing pairwise interaction parameters for all the species in that group into sets of common parameters (Ovaskainen et al. 2017).

In ecological contexts, such groups have been defined *a priori* by phylogeny at some taxonomic level or by functional (i.e., trait) similarity (Uriarte et al. 2004; Ovaskainen et al. 2017; Martyn et al. 2021). The argument for grouping based on phylogeny relies on the principle that closely related species will fill similar ecological niches, and hence will interact with other species in a similar fashion (Darwin 1859; Webb et al. 2002; Godoy et al. 2014). However, empirical tests of

this principle have found remarkably little support for it (Cahill et al. 2008; Mayfield and Levine 2010; Godoy et al. 2014). On the other hand, functional groupings have also proven to be poor predictors of species interactions (Bret-Harte et al. 2008; Funk et al. 2017). While attempts have been made to define functional groupings that are more useful in a modelling context (Boulangeat et al. 2012), these rely upon measurements of biotic interactions to define such groups. As such, integrating these into species interaction models leaves us with a circular argument: a functional group is a group of species that interact with others in a similar way, and we therefore assume that all the species in a functional group interact with others in a similar way. In all, we are left with the question of how best to group species in community models, and whether we can do so in a meaningful way using fewer data than required for a fully pairwise parameterisation.

To the best of our knowledge, all studies that use grouping to simplify models of species interactions have also made one additional, untested assumption: that the same grouping is equally valid for how species affect others as for how species respond to others. There is good reason to question this assumption, as it has been recognised for some time that different traits can govern species' competitive effects than govern their responses to competition (Goldberg and Landa 1991). For resource competition in particular, competitive effects are determined by how the organism affects local resource availability while competitive responses are determined by how the organism is inhibited by depleted resources (Violle et al. 2009). As such, the best groupings of species may be different for each of the effects and responses.

We propose that species might be best partitioned into groups without regard to any *a priori* definition but instead based purely upon which grouping gives the most parsimonious model. Such data-informed grouping approaches have previously proved successful in other contexts, such as finding modules in networks (Girvan and Newman 2002; Guimerà and Nunes Amaral 2005; Newman 2006). In this study, we investigate the usefulness and accuracy of such an *a posteriori* approach using experimental data from competition between different *Drosophila* species provided by Terry et al. (2021). We fit a species competition model, using an underlying Beverton–Holt model for the growth rate, to these data to find the groupings that give the most parsimonious

model, and we then compare these groupings to similarity based upon a species trait (body size). While doing so, we group species independently with regard to their effects upon other species and their responses to other species to find whether the data-informed groupings differ between these two cases.

2 Methods

2.1 Data

The data used in this study were drawn from the work of Terry et al. (2021). They consist of fecundity measurements—number of surviving offspring—for six species of *Drosophila*, grown in single-species or two-species communities with varying densities of each species. This experimental design corresponds to a response–surface experimental design commonly used to parameterise species interactions in population dynamics models (Hart et al. 2018; Cameron et al. 2019; Zhang and Kleunen 2019).

2.1.1 Summary of the experimental methodology

The experiment used six species of *Drosophila* from an Australian rainforest community: *D. birchii*, *D. pallidifrons*, *D. pandora*, *D. pseudoananassae*, *D. simulans*, and *D. sulfurigaster*. The experiment took place in vials with a limited amount of medium for egg-laying and larval development; competition was for space on this medium. A fixed number of adult pairs of one or two species was added to each vial and allowed to mate and lay eggs. The number of adults of each species emerging from these eggs was counted.

Experiments occurred for every single-species and two-species combination of the six species, at a variety of densities. The single-species densities were 6, 18, or 24 adults per vial, and each of these was replicated six times per species. The two-species densities were (6,6), (6,12), (6,18), (6,24), (12,6), (18,6), or (24,6) adults per vial (where the paired values give the densities of each of the two species), and each of these was replicated three times per pair of species. Vials with

a single founding pair of founding adults were excluded from this study and the original authors' due to an Allee effect that the model cannot accommodate. Further details of the experimental methodology can be found in Terry et al. (2021).

2.2 Model

The data provided the change in population size from one generation to the next, so, following the original authors, we used a version of the Beverton–Holt model given by:

$$N_{i,t+1} = \frac{r_i N_{i,t}}{1 + \sum_j^n \alpha_{ij} N_{j,t}}\,, \tag{1}$$

where $N_{i,t}$ is the population of species i at time t (the size of the parent generation), $N_{i,t+1}$ is the population at time $t+1$ (the size of the offspring generation), r_i is the intrinsic growth rate of species i, and α_{ij} is the per capita effect of species j on species i. When fitting the data to the model, the competition coefficients were not constrained to be positive; this allows for the possibility of mutualistic effects between some species. With n species, this model has $n + n^2$ parameters: n growth rates and n^2 competition coefficients.

2.2.1 Species grouping

In order to reduce the number of free parameters in the model, we grouped species together, and constrained the appropriate competition coefficients to be equal for species within the same group. Within the model defined above, each species is effectively in its own group, such that all pairs of species interact in a unique fashion. At the other extreme, all species could be placed in the same group, such that all pairs of species interact equivalently; they are ecologically equivalent *sensu* McPeek and Siepielski (2019). Given n species, there may be any integer number of groups from 1 to n. Groupings are known formally as partitions in mathematics, and the total number of ways to group n species is given by the Bell numbers, denoted B_n (Bell 1934).

It is commonly assumed that species which are similar in their responses to competition will

automatically be similar in their competition effects. However, competitive responses may be governed by a different set of traits than competitive effects, so this assumption may not always hold. We therefore relaxed this assumption by allowing species which were similar in one regard to be dissimilar in the other, and vice versa. As such, each grouped model was defined by two groupings of species: one grouping of the competitive responses and one of the competitive effects.

To better understand the meaning of species grouping of competition coefficients, it is useful to recognise that this component of the model can be represented as an interaction matrix (Novak et al. 2016), where a row contains all the competitive effects on a single species and a column contains all the competitive effects of a single species. If two species are assumed to have the same responses to competition, then two rows of the interaction matrix are constrained to be equal. Similarly, if two species are assumed to have the same competitive effects, then two columns of the interaction matrix are constrained to be equal. This allows us to refer below to the grouping of the competitive responses as the "row grouping" and the grouping of the competitive effects as the "column grouping". Formally, if species a and b are in the same row group, $\alpha_{aj} \equiv \alpha_{bj}\, \forall j$, and if a and b are in the same column group, $\alpha_{ia} \equiv \alpha_{ib}\, \forall i$.

With six species and independently selected row and column groupings, the total number of possible models is $(B_6)^2 = 41\,209$. The species' intrinsic growth rates, r_i, were not affected by grouping; each species always had its own growth rate fitted independently of other species. The total number of parameters in the model is reduced to n growth rates, and a number of competition coefficients equal to the number of row groups multiplied by the number of column groups.

2.2.2 Model fitting

The model in eq. (1) is not linear in the parameters. With the rows of the interaction matrix ungrouped, the regression could be reparameterised to be linear in the parameters by treating $r_i\alpha_{ij}$ as a single regression parameter for each i, j pair. With the rows grouped, however, this becomes impossible since each species still has an independent growth rate. As such, we instead fit the models with the SbPlx algorithm of the NLopt package C++ (Rowan 1990; Johnson 2007).

Following the original authors, we assumed a negative binomial error distribution with a single overdispersion parameter for all species.

2.2.3 Assessing parsimony

To assess the parsimony of the different species groupings, we exhaustively fit all $41\,209$ possible models and computed the resulting AICc for those fits. As a measure of parsimony, we selected AIC over BIC on the basis that we want to find the best model for making ecological predictions rather than seeking the "true" model (Aho et al. 2014). Furthermore, AIC is more conservative in promoting grouping as BIC imposes a greater penalty for the number of parameters for any reasonably sized data set. Some of the models had a relatively high number of parameters compared to the number of observations (727 observations compared to up to 42 parameters), so we used AICc to correct for this "small sample size" effect. AICc values can be compared to determine which model is the most parsimonious, but they cannot directly be interpreted to find the significance of the difference in parsimony. As such, we also converted the set of AICc values to Akaike weights (Wagenmakers and Farrell 2004). These give the probability that each model is the best among the models assessed.

Using these Akaike weights, we calculated two important statistics. First, the total Akaike weight of models with a lower AICc than the entirely ungrouped model; this is the probability that some model is more parsimonious than the model that assumes all species are statistically unique. Second, the total Akaike weight of models that used the same grouping for the rows as for the columns; this is the probability that the models grouping species equivalently in their competitive effects and competitive responses were more parsimonious than all other models. To evaluate the significance of the latter, we performed a Monte Carlo permutation test, comparing the total Akaike weight of those models that grouped rows and columns equivalently to the total Akaike weight of an equally sized random subset of the models.

2.2.4 Coclassification matrices

To summarise the results, we also used the Akaike weights to generate two coclassification matrices, one for the row groupings and one for the column groupings. Each cell of a coclassification matrix represents the coclassification of a pair of species, and contains the total Akaike weight of models that place those two species in the same group. As such, each cell represents the probability that the most parsimonious model is one that coclassifies those two species. For reference, note that the diagonal of a coclassification matrix is tautologically equal to 1.

There are several possible "null" expectations for the off-diagonal elements of these matrices. A neutral model approach would expect all species to be in the same group: every off-diagonal cell would therefore be 1. If each species were expected to be unique, every off-diagonal cell would instead be 0. Alternatively, if every possible grouping were considered to be equally likely, every off-diagonal cell would be $\frac{B_{n-1}}{B_n}$: 0.256 for six species. Unfortunately, without a null *distribution* of coclassification matrices, it was impossible to test the data-informed coclassification matrices against a null expectation.

However, we could test the row-grouping and column-grouping coclassification matrices against one another. By testing for correlation between the coclassification matrix for the rows and that for the columns, we could perform a second test of the assumption that it is appropriate to use the same grouping for both. We did this using a Mantel test, as appropriate for similarity matrices. We could also test the coclassification matrices for correlation with an expected dissimilarity matrix. We calculated one such dissimilarity matrix based on a species trait: body size (as mean wet mass of female adults, obtained from Terry et al. 2021). The trait difference matrix was calculated as absolute difference between trait values while the coclassification matrices were converted to dissimilarity matrices by subtracting them from 1. We then performed Mantel tests for correlation between the trait distance matrix and each of the row-grouping and column-grouping coclassification dissimilarity matrices.

3 Results

Plots of predicted against observed values are shown in figs. S1 and S2. The AICc values ranged from 6166.79 to 6278.59. The ungrouped model gave an AICc value of 6176.14. The null model, with all competition coefficients equal to zero, gave an AICc value of 6298.02. Figure 1 shows the distribution of AICc values from the models, with some models of interest marked.

The "best" model, as indicated by AICc, was not the best with any surety; its Akaike weight indicated a probability of only 0.096 that it was the best. Figure 2 shows the groupings and alpha values of this "best" model. The ungrouped model had an AICc just 9.35 units from the most parsimonious model, but the total Akaike weight assigned to models more parsimonious than the ungrouped model was 0.893; this is the probability that some grouped model was more parsimonious than the ungrouped model.

Considering only models in which the rows and columns were grouped using the same grouping, the AICc values ranged from 6172.26 to 6264.85; neither the best nor worst overall models had equivalently grouped rows and columns. The total Akaike weight assigned to these models was 0.007; this was more than expected by chance (0.005), but not significantly so (Monte Carlo permutation test, $p = 0.183$). This indicates that the most parsimonious model was not especially likely to be one that grouped the rows and columns equivalently.

3.1 Coclassification of species

Overall, no pair of species could confidently be said to be coclassified together in the row or column groupings; the highest probability of coclassification was $64.4\,\%$ (fig. 3). However, some pairs of species were far more likely to be coclassified than others, suggesting some structure underpinning the best groupings.

The structure of the coclassification matrices was very different in the row groups than in the column groups. That is, most pairs of species which had a high probability of being coclassified in one grouping had a very low probability of being coclassified in the other grouping. In particular, in

the column groupings, *D. birchii* was frequently coclassified with *D. pandora* but was almost never coclassified with *D. pallidifrons*, while in the row groupings this was completely reversed. Overall, more coclassification occurred in the row groupings than in the column groupings. A Mantel test of the coclassification matrix for the rows against that for the columns showed that there is no significant correlation between the two ($r_{\mathrm{M}} = -0.243$, $p = 0.827$). The column coclassification matrix was significantly correlated with a dissimilarity matrix based on species body size ($r_{\mathrm{M}} = 0.395$, $p = 0.019$), but the row coclassification matrix was not ($r_{\mathrm{M}} = -0.367$, $p = 0.895$). Testing the correlation of average competitive effects and competitive responses (column- and row-averaged alpha values from the ungrouped model) with species body size, competitive effects are positively correlated with body size ($r = 0.767$, $p = 0.038$), but competitive responses are not ($r = -0.056$, $p = 0.542$), matching the results from the Mantel tests.

Some intransitivity was observed in the coclassification matrices. For example, in the column groupings, *D. pallidifrons* had a relatively high probability of being coclassified with *D. birchii*, and a substantial probability of being coclassified *D. pseudoananassae*, yet *D. birchii* and *D. pseudoananassae* had almost no probability of being coclassified themselves.

4 Discussion

We have shown that it is possible to fit a more parsimonious model of species interactions by grouping the interaction coefficients of species *a posteriori*, and hence reducing the number of free parameters in the model. Indeed, the results show that grouping species was likely to give a more parsimonious model (probability 0.893) than leaving them ungrouped. Which grouping would give the most parsimonious model, however, could not be determined with confidence. This result stands in contrast to the popular wisdom in ecology that species should form some sort of clear groups, be they the functional groups so often used (Blondel 2003; Kirwan et al. 2009; Boulangeat et al. 2012), or taxonomic groups per a relationship between phylogeny and ecological niche (Godoy et al. 2014; Wilcox et al. 2018). It also contrasts with previous studies that have

succeeded in finding clear groups in other ecological contexts, such as in food webs (Krause et al. 2003; Allesina and Pascual 2009) or plant–pollinator networks (Olesen et al. 2007).

One possible explanation for our results is that species simply do not group to the same extent in a quantitative competitive context as in binary interaction networks. That said, the results also suggest that this is unlikely as grouping was useful for increasing parsimony. In the present case, we cannot rule out the possibility that a lack of clarity in the groupings is instead a product of the study system. All the species in this study were members of the same genus, *Drosophila*, and as such were all rather similar. A study system with species spanning multiple functional groups, or a broader taxonomy, might have shown less equivocal groupings. Alternatively, the small number of species in the study might not have been sufficient for clear groups to be defined. Alas, the brute force method we employed in this study has a very poor algorithmic complexity, preventing an application of this approach to studies with more than seven or eight species (there are $(B_7)^2 = 769\,129$ possible models with 7 species and $(B_8)^2 = 17\,139\,600$ models with 8 species). As such, a better method would be required in order to test for grouping of species interactions among a larger set of species.

While the results could not definitely identify a single best grouping, it was clear that the model-supported groupings differed for the rows and columns. That is, species that are similar in their response to competition may not be similar in the competitive effects they have on other species, and vice versa. In hindsight, this should not be overly surprising since it is known that different traits govern species competitive effects than govern their responses (Goldberg 1990; Goldberg and Landa 1991); for example shading effect and shade tolerance in plants. In this system, the probability of column (competitive effect) coclassification was correlated with similarity in body size, and average competitive effect was directly correlated with body size, suggesting that body size is a trait governing competitive effect. However, body size had no correlation with competitive response, suggesting that a different trait is responsible for determining response (the original authors did not record any other species traits that could be tested). Yet the assumption that species should group equivalently in regards to competitive effects and responses is frequently made, often

implicitly and without comment (e.g. Ives et al. 2003; Martorell and Freckleton 2014; Gross and Edmunds 2015). Our results recommend more explicit questioning, and even testing, of this assumption when building future species interaction models.

The intransitivity of coclassification for some sets of species even suggests that there may not be a single "best" grouping of the species, but rather two or more quite distinct and mutually exclusive groupings of similar validity. A single trait like body size might therefore not be able to capture the multivariate nature of species' competitive niches (Blonder 2018). This could indicate that a Bayesian approach is worth exploring in the future, in which the outcome is not a definitive grouping but a distribution of valid groupings. Alternatively, it may suggest adopting a mixed-membership approach in which a species can belong to more than one group (Airoldi et al. 2008), reminiscent of the community-level drivers of Ovaskainen et al. (2017).

5 Data availability

All raw observation data used in the manuscript are archived on Zenodo `https://doi.org/10.5281/zenodo.4736655` and GitHub `https://github.com/jcdterry/TCL_DrosMCT/tree/1.0.0`. All code used to perform the analyses is also archived on Zenodo `https://doi.org/10.5281/zenodo.7677826` and GitHub `https://github.com/pyrrhicPachyderm/interaction-partitioning/tree/ms1-v2`.

Acknowledgements

We would like to thank Hao Ran Lai for feedback on a draft of the manuscript. DBS is also grateful for the support of the Marsden Fund Council, from New Zealand Government funding (grant 16-UOC-008).

Authors' contributions

DBS conceptualised the investigation. CRPB developed and implemented the analysis, with input and feedback from DBS. CRPB wrote the first manuscript draft. Both authors contributed to manuscript edits.

Conflicts of interest

The authors declare that they have no conflicts of interest.

Literature Cited

Aho, K., D. Derryberry, and T. Peterson. 2014. Model selection for ecologists: the worldviews of AIC and BIC. Ecology 95:631–636. DOI: 10.1890/13-1452.1.

Airoldi, E. M., D. M. Blei, S. E. Fienberg, and E. P. Xing. 2008. Mixed membership stochastic blockmodels. Journal of Machine Learning Research 9:1981–2014.

Allesina, S., and M. Pascual. 2009. Food web models: a plea for groups. Ecology Letters 12:652–662. DOI: 10.1111/j.1461-0248.2009.01321.x.

Bell, E. T. 1934. Exponential polynomials. Annals of Mathematics 35:258–277. DOI: 10.2307/1968431.

Blondel, J. 2003. Guilds or functional groups: does it matter? Oikos 100:223–231. DOI: 10.1034/j.1600-0706.2003.12152.x.

Blonder, B. 2018. Hypervolume concepts in niche-and trait-based ecology. Ecography 41:1441–1455. DOI: 10.1111/ecog.03187.

Boulangeat, I., P. Philippe, S. Abdulhak, R. Douzet, L. Garraud, S. Lavergne, S. Lavorel, J. Van Es, P. Vittoz, and W. Thuiller. 2012. Improving plant functional groups for dynamic models of biodiversity: at the crossroads between functional and community ecology. Global Change Biology 18:3464–3475. DOI: 10.1111/j.1365-2486.2012.02783.x.

Bret-Harte, M. S., M. C. Mack, G. R. Goldsmith, D. B. Sloan, J. DeMarco, G. R. Shaver, P. M. Ray, Z. Biesinger, and F. S. Chapin III. 2008. Plant functional types do not predict biomass responses to removal and fertilization in Alaskan tussock tundra. Journal of Ecology 96:713–726. DOI: 10.1111/j.1365-2745.2008.01378.x.

Cahill Jr., J. F., S. W. Kembel, E. G. Lamb, and P. A. Keddy. 2008. Does phylogenetic relatedness influence the strength of competition among vascular plants? Perspectives in Plant Ecology, Evolution and Systematics 10:41–50. DOI: 10.1016/j.ppees.2007.10.001.

Cameron, H., T. Coulson, and D. J. Marshall. 2019. Size and density mediate transitions between competition and facilitation. Ecology Letters 22:1879–1888. DOI: 10.1111/ele.13381.

Cornell, H. V. 1999. Unsaturation and regional influences on species richness in ecological communities: a review of the evidence. Ecoscience 6:303–315. DOI: 10.1080/11956860.1999.11682532.

Darwin, C. 1859. On the Origin of Species. John Murray, London, England.

Funk, J. L., J. E. Larson, G. M. Ames, B. J. Butterfield, J. Cavender-Bares, J. Firn, D. C. Laughlin, A. E. Sutton-Grier, L. Williams, and J. Wright. 2017. Revisiting the Holy Grail: using plant functional traits to understand ecological processes. Biological Reviews 92:1156–1173. DOI: 10.1111/brv.12275.

Gilman, S. E., M. C. Urban, J. Tewksbury, G. W. Gilchrist, and R. D. Holt. 2010. A framework for community interactions under climate change. Trends in Ecology & Evolution 25:325–331. DOI: 10.1016/j.tree.2010.03.002.

Girvan, M., and M. E. Newman. 2002. Community structure in social and biological networks. Proceedings of the National Academy of Sciences 99:7821–7826. DOI: 10.1073/pnas.122653799.

Godoy, O., N. J. B. Kraft, and J. M. Levine. 2014. Phylogenetic relatedness and the determinants of competitive outcomes. Ecology Letters 17:836–844. DOI: 10.1111/ele.12289.

Goldberg, D. E. 1990. Components of resource competition in plant communities. Perspectives on Plant Competition, ed. by J. B. Grace, and D. Tilman, 27–49. Academic Press, San Diego, CA.

Goldberg, D. E., and K. Landa. 1991. Competitive effect and response: hierarchies and correlated traits in the early stages of competition. Journal of Ecology 79:1013–1030. DOI: 10.2307/2261095.

Gross, K., and P. J. Edmunds. 2015. Stability of Caribbean coral communities quantified by long-term monitoring and autoregression models. Ecology 96:1812–1822. DOI: 10.1890/14-0941.1.

Guimerà, R., and L. A. Nunes Amaral. 2005. Functional cartography of complex metabolic networks. Nature 433:895–900. DOI: 10.1038/nature03288.

Hart, S. P., R. P. Freckleton, and J. M. Levine. 2018. How to quantify competitive ability. Journal of Ecology 106:1902–1909. DOI: 10.1111/1365-2745.12954.

Ives, A. R., B. Dennis, K. Cottingham, and S. R. Carpenter. 2003. Estimating community stability and ecological interactions from time-series data. Ecological Monographs 73:301–330. DOI: 10.1890/0012-9615(2003)073[0301:ECSAEI]2.0.CO;2.

Johnson, S. G. 2007. The NLopt nonlinear-optimization package. https://github.com/stevengj/nlopt.

Kirwan, L., J. Connolly, J. A. Finn, C. Brophy, A. Lüscher, D. Nyfeler, and M.-T. Sebastià. 2009. Diversity–interaction modeling: estimating contributions of species identities and interactions to ecosystem function. Ecology 90:2032–2038. DOI: 10.1890/08-1684.1.

Krause, A. E., K. A. Frank, D. M. Mason, R. E. Ulanowicz, and W. W. Taylor. 2003. Compartments revealed in food-web structure. Nature 426:282–285. DOI: 10.1038/nature02115.

Lotka, A. J. 1925. Elements of Physical Biology. Williams & Wilkins, Baltimore, MD.

Martorell, C., and R. P. Freckleton. 2014. Testing the roles of competition, facilitation and stochasticity on community structure in a species-rich assemblage. Journal of Ecology 102:74–85. DOI: 10.1111/1365-2745.12173.

Martyn, T. E., D. B. Stouffer, O. Godoy, I. Bartomeus, A. I. Pastore, and M. M. Mayfield. 2021. Identifying “useful” fitness models: balancing the benefits of added complexity with realistic

data requirements in models of individual plant fitness. The American Naturalist 197:415–433. DOI: 10.1086/713082.

Mayfield, M. M., and J. M. Levine. 2010. Opposing effects of competitive exclusion on the phylogenetic structure of communities. Ecology Letters 13:1085–1093. DOI: 10.1111/j.1461-0248.2010.01509.x.

Maynard, D. S., Z. R. Miller, and S. Allesina. 2020. Predicting coexistence in experimental ecological communities. Nature Ecology & Evolution 4:91–100. DOI: 10.1038/s41559-019-1059-z.

McCann, K., A. Hastings, and G. R. Huxel. 1998. Weak trophic interactions and the balance of nature. Nature 395:794–798. DOI: 10.1038/27427.

McPeek, M. A., and A. M. Siepielski. 2019. Disentangling ecologically equivalent from neutral species: the mechanisms of population regulation matter. Journal of Animal Ecology 88:1755–1765. DOI: 10.1111/1365-2656.13072.

Mutshinda, C. M., R. B. O'Hara, and I. P. Woiwod. 2009. What drives community dynamics? Proceedings of the Royal Society B: Biological Sciences 276:2923–2929. DOI: 10.1098/rspb.2009.0523.

Neutel, A.-M., J. A. Heesterbeek, and P. C. De Ruiter. 2002. Stability in real food webs: weak links in long loops. Science 296:1120–1123. DOI: 10.1126/science.1068326.

Newman, M. E. 2006. Modularity and community structure in networks. Proceedings of the National Academy of Sciences 103:8577–8582. DOI: 10.1073/pnas.0601602103.

Novak, M., J. D. Yeakel, A. E. Noble, D. F. Doak, M. Emmerson, J. A. Estes, U. Jacob, M. T. Tinker, and J. T. Wootton. 2016. Characterizing species interactions to understand press perturbations: what is the community matrix? Annual Review of Ecology, Evolution, and Systematics 47:409–432. DOI: 10.1146/annurev-ecolsys-032416-010215.

Olesen, J. M., J. Bascompte, Y. L. Dupont, and P. Jordano. 2007. The modularity of pollination networks. Proceedings of the National Academy of Sciences 104:19891–19896. DOI: 10.1073/pnas.0706375104.

Ovaskainen, O., G. Tikhonov, D. Dunson, V. Grøtan, S. Engen, B.-E. Sæther, and N. Abrego. 2017. How are species interactions structured in species-rich communities? A new method for analysing time-series data. Proceedings of the Royal Society B: Biological Sciences 284:20170768. DOI: 10.1098/rspb.2017.0768.

Rocha Filho, T. M., I. M. Gléria, A. Figueiredo, and L. Brenig. 2005. The Lotka–Volterra canonical format. Ecological Modelling 183:95–106. DOI: 10.1016/j.ecolmodel.2004.07.023.

Rowan, T. H. 1990. Functional stability analysis of numerical algorithms. PhD thesis. Department of Computer Science, University of Texas at Austin, Austin, TX.

Sutherst, R., G. Maywald, and A. Bourne. 2007. Including species interactions in risk assessments for global change. Global Change Biology 13:1843–1859. DOI: 10.1111/j.1365-2486.2007.01396.x.

Terry, J. C. D., J. Chen, and O. T. Lewis. 2021. Natural enemies have inconsistent impacts on the coexistence of competing species. Journal of Animal Ecology 90:2277–2288. DOI: 10.1111/1365-2656.13534.

Uriarte, M., R. Condit, C. D. Canham, and S. P. Hubbell. 2004. A spatially explicit model of sapling growth in a tropical forest: does the identity of neighbours matter? Journal of Ecology 92:348–360. DOI: 10.1111/j.0022-0477.2004.00867.x.

Violle, C., E. Garnier, J. Lecoeur, C. Roumet, C. Podeur, A. Blanchard, and M.-L. Navas. 2009. Competition, traits and resource depletion in plant communities. Oecologia 160:747–755. DOI: 10.1007/s00442-009-1333-x.

Wagenmakers, E.-J., and S. Farrell. 2004. AIC model selection using Akaike weights. Psychonomic Bulletin & Review 11:192–196. DOI: 10.3758/BF03206482.

Warton, D. I., F. G. Blanchet, R. B. O'Hara, O. Ovaskainen, S. Taskinen, S. C. Walker, and F. K. C. Hui. 2015. So many variables: joint modeling in community ecology. Trends in Ecology & Evolution 30:766–779. DOI: 10.1016/j.tree.2015.09.007.

Webb, C. O., D. D. Ackerly, M. A. McPeek, and M. J. Donoghue. 2002. Phylogenies and community ecology. Annual Review of Ecology and Systematics 33:475–505. DOI: 10.1146/annurev.ecolsys.33.010802.150448.

Weiss-Lehman, C. P. et al. 2022. Disentangling key species interactions in diverse and heterogeneous communities: A Bayesian sparse modelling approach. Ecology Letters 25:1263–1276. DOI: 10.1111/ele.13977.

Wilcox, T. M., M. K. Schwartz, and W. H. Lowe. 2018. Evolutionary community ecology: time to think outside the (taxonomic) box. Trends in Ecology & Evolution 33:240–250. DOI: 10.1016/j.tree.2018.01.014.

Wootton, K., and D. Stouffer. 2016. Many weak interactions and few strong; food-web feasibility depends on the combination of the strength of species' interactions and their correct arrangement. Theoretical Ecology 9:185–195. DOI: 10.1007/s12080-015-0279-3.

Zhang, Z., and M. van Kleunen. 2019. Common alien plants are more competitive than rare natives but not than common natives. Ecology Letters 22:1378–1386. DOI: 10.1111/ele.13320.

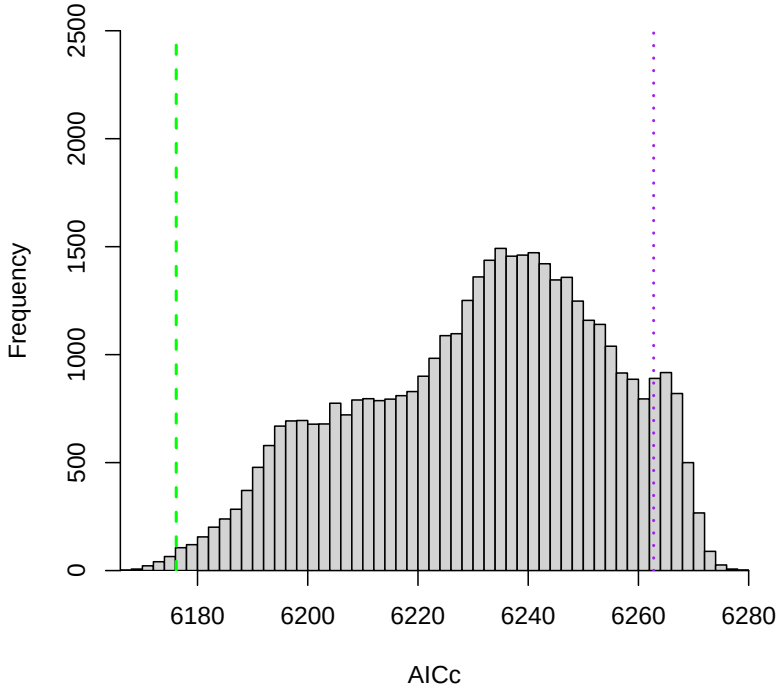


Figure 1: Histogram of the AICc values of Beverton–Holt model fits with all possible groupings of interaction coefficients. The dashed green line shows the AICc value (6176.14) of the ungrouped model in which all competition coefficients are fitted independently. The dotted purple line shows the AICc value (6262.76) of the fully grouped model in which every pair of species shares a single interaction coefficient.

	D. pallidifrons	*D. birchii*	*D. pandora*	*D. pseudoananassae*	*D. simulans*	*D. sulfurigaster*
D. pallidifrons	0.097	0.023	−0.007	0.130	0.040	
D. birchii						
D. pandora	0.047	0.004	0.021	−0.007	0.007	
D. pseudoananassae	0.096	0.002	0.027	0.063	0.160	
D. simulans	0.009	0.012	0.032	0.056	0.042	
D. sulfurigaster	0.073	0.045	0.003	0.026	0.095	

(a) Grouped alpha matrix

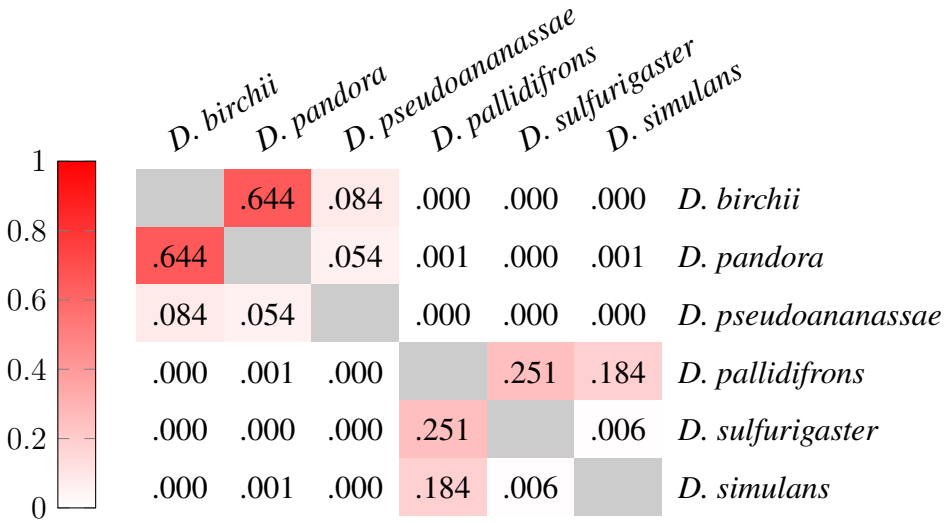
	0	0	0	0	0
0		1	0	0	0
0	1		0	0	0
0	0	0		0	0
0	0	0	0		0
0	0	0	0	0	

(b) Coclassification of competitive effects (columns)

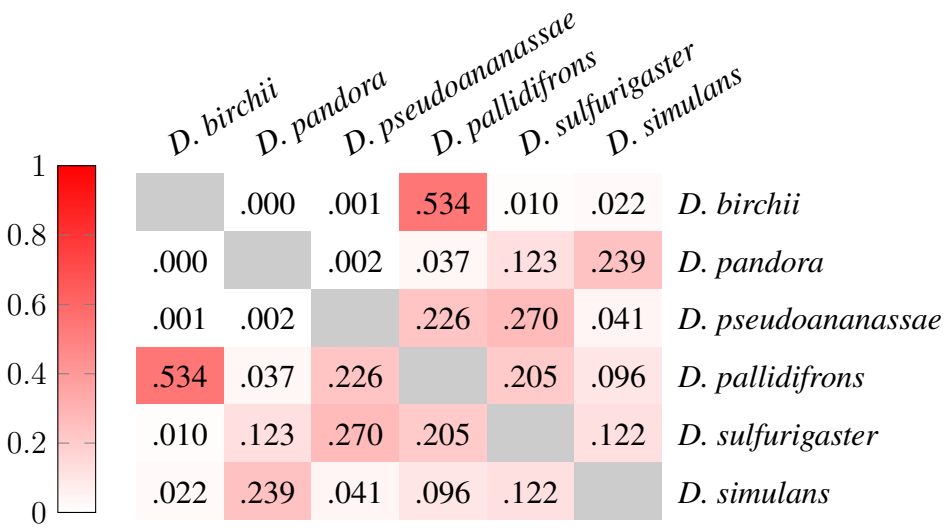
	1	0	0	0	0
1		0	0	0	0
0	0		0	0	0
0	0	0		0	0
0	0	0	0		0
0	0	0	0	0	

(c) Coclassification of competitive responses (rows)

Figure 2: The groupings and competition coefficients of a grouped Beverton–Holt model fitted to the six studied *Drosophila* species, where results are for the single model with the best AICc (6166.790). (a) We show the inferred matrix of competition coefficients, where a competition coefficient in a block represents the same coefficient being used for multiple interactions (per the identified groupings). (b) and (c) We show the coclassification matrices derived from the groupings of the columns and rows, respectively. Species have been reordered so that species that were grouped together appear adjacent, and the ordering is the same in all of (a), (b), and (c).



(a) Coclassification of competitive effects



(b) Coclassification of competitive responses

Figure 3: Coclassification matrices of the Beverton–Holt competitive effects and responses of the six studied *Drosophila* species, weighted by Akaike weights. The cell for a pair of species represents the probability that the best grouping, as determined by the Akaike weights of the grouped models, places those two species in the same group. (a) We show the grouping of species' competitive effects upon other species (columns of the interaction matrix). (b) We show the grouping of species' responses to competition (rows of the interaction matrix).