

Fundamental contradictions among observational and experimental estimates of non-trophic species interactions

ALLISON K. BARNER ¹, KYLE E. COBLENTZ , SALLY D. HACKER, AND BRUCE A. MENGE

Department of Integrative Biology, Oregon State University, 3029 Cordley Hall, Corvallis, Oregon 97331 USA

Abstract. The difficulty of experimentally quantifying non-trophic species interactions has long troubled ecologists. Increasingly, a new application of the classic “checkerboard distribution” approach is used to infer interactions by examining the pairwise frequency at which species are found to spatially co-occur. However, the link between spatial associations, as estimated from observational co-occurrence, and species interactions has not been tested. Here we used nine common statistical methods to estimate associations from surveys of rocky intertidal communities in the Northeast Pacific Ocean. We compared those inferred associations with a new data set of experimentally determined net and direct species interactions. Although association methods generated networks with aggregate structure similar to previously published interaction networks, each method detected a different set of species associations from the same data set. Moreover, although association methods generally performed better than a random model, associations rarely matched empirical net or direct species interactions, with high rates of false positives and true positives, and many false negatives. Our findings cast doubt on studies that equate species co-occurrences to species interactions and highlight a persistent, unanswered question: how do we interpret spatial patterns in communities? We suggest future research directions to unify the observational and experimental study of species interactions, and discuss the need for community standards and best practices in association analysis.

Key words: assembly; checkerboard patterns; competition; co-occurrence; joint species distribution models; null models; pairwise dissimilarity; species associations.

INTRODUCTION

Non-trophic species interactions are difficult ecological relationships to quantify, often requiring experiments that are logistically infeasible. However, new applications of a classic concept (Diamond 1975, Connor and Simberloff 1979) facilitate the estimation of non-trophic interactions from widely available spatial occurrence data. Specifically, the strengths and types of species interactions within a community are being increasingly estimated from patterns in species spatial co-occurrences (Sfenthourakis et al. 2006, Faisal et al. 2010, Gotelli and Ulrich 2010, Ovaskainen et al. 2010, Faust and Raes 2012, Veech 2013, Lane et al. 2014, Pollock et al. 2014, Harris 2016). Using these methods, if a pair of species is found co-occurring in a community less (or more) often than expected by chance, then they are thought to be competitively (or positively) interacting.

Many new methods have been developed to detect species interactions from analyses of spatial co-occurrence data (e.g., Sfenthourakis et al. 2006, Faisal et al. 2010, Gotelli and Ulrich 2010, Ovaskainen et al. 2010, Faust

and Raes 2012, Veech 2013, Lane et al. 2014, Pollock et al. 2014, Harris 2016). These new methods, broadly termed “association methods” for their ability to detect statistical spatial associations among species, build on the classic inference of competition from “checkerboard distributions” (Diamond 1975), which assumes that non-trophic species interactions create spatial patterns distinct from other assembly processes. This classic approach is now paired with newly available large community data sets, increasing computational power, and growing interest in applying network methods across systems (Faust and Raes 2012). However, with many of the newly available association methods, no consensus has emerged about the best-performing method, making it difficult to compare across studies, and thus difficult to interpret the meaning of any one “association.”

As with association methods, how to experimentally estimate species interactions has been long debated. Different experimental designs quantify different interaction metrics associated with a range of technical and mathematical definitions (Laska and Wootton 1998, Wootton and Emmerson 2005). Experiments may ultimately reveal little about how interactions influence community structure; two species may co-occur but not interact, two species may interact experimentally but never co-occur. As such, the metric of interest in the study of community assembly may not be the direct interactions among species, but their net effect (Cazelles et al. 2016, Harris 2016). Or, emergent community patterns may be the unit

Manuscript received 3 October 2017; revised 14 December 2017; accepted 18 December 2017. Corresponding Editor: Nicholas J. Gotelli.

¹ Present address: Department of Environmental Science, Policy, and Management, University of California, 130 Mulford Hall, Berkeley, California, 94720 USA. E-mail: allisonkbarner@gmail.com

of interest, as they are in functional trait and community phylogenetic approaches to assembly. Further, a pairwise, experimental approach likely underestimates the total number of interactions in ecosystems due to logistical constraints. In many cases, experimentation is impossible and alternative approaches are necessary (e.g., microbial communities [Steele et al. 2011, Faust and Raes 2012, Levy and Borenstein 2013, Zelezniak et al. 2015], paleoecological communities [Blois et al. 2014, Lyons et al. 2016]). Finally, experiments may not capture spatial or temporal variation in species interactions (Chamberlain et al. 2014).

However, although most association methods are proposed as hypothesis-generation tools to circumvent some of the problems with interaction experiments, associations estimated using these methods are rarely subsequently tested. To our knowledge, only two studies have compared spatially estimated interactions with observational interaction data, both with some success (insect predation [Bell et al. 2010], microbial interactions [Levy and Borenstein 2013]). Association methods are often validated against simulated community data with known underlying species interactions (e.g., Cazelles et al. 2016, Harris 2016). However, models that simulate community membership as a function of species interactions are tautologically likely to show that species interactions are detectable from community membership. Overall, spatial co-occurrence has been proposed as a method of quantifying pairwise species interactions without definitive tests of predictive power and critiques of association methods primarily address null model approaches (Connor and Simberloff 1979, Schamp et al. 2015, Harris 2016).

Nevertheless, association methods are used to draw strong inferences about species interactions and the nature of ecological processes. Classic co-occurrence theory was derived explicitly on the assumption of competition among the species of interest (Diamond and Gilpin 1982). However, modern co-occurrence theory implemented in association analysis has been used to identify interactions among species for which no a priori functional or trophic information exists (e.g., Lima-Mendez et al. 2015). This approach has also been used to detect trophic interactions (Bell et al. 2010, Faisal et al. 2010), though with uncertain predictions about whether predation should result in a positive or negative associations (Schluter 1984, Morales-Castilla et al. 2015). Further, association methods have been proposed for use in resource management and human health, including for pest management (Bell et al. 2010), in conservation decision-making (Araújo et al. 2011), as indicators of climate-driven ecosystem collapse (Griffith et al. 2018), and to understand how human-associated microbial interactions impact health (Faust and Raes 2012).

In this study, we examined the coherence among experimentally estimated and occurrence-estimated species interactions. We specifically distinguish *species associations* and *species interactions*. *Species associations* measure the degree to which the occurrence of two

species is correlated, the significance of which is determined using statistical analysis of co-occurrence. *Species interactions* measure the numerical effect that one species has on the abundance or distribution of another species, either directly or through the net effect of both direct and indirect pathways. Following studies showing that the effects of species removal experiments may be predicted from observed temporal association analysis (Wootton 2004, Sander et al. 2017), we ask whether spatial co-occurrence can predict the general structure and specific links in empirical species interaction networks. We implemented a suite of existing methods using observational data in the species-rich Northeast Pacific rocky intertidal ecosystem, and compared associations against a new data set on experimentally determined direct and net interactions from the same ecosystem.

MATERIAL AND METHODS

Community survey for association analysis

In May–September 2012, we conducted surveys at 15 locations along Oregon coast (Appendix S1). In each low intertidal zone, we conducted nine 25 × 25 cm quadrat surveys (except for the Manipulation Bay site, where only seven surveys were conducted due to logistical constraints), for a total of 133 survey plots (see Appendix S2 for survey design sensitivity analysis). Within each plot, we recorded presence–absence data for all species (invertebrates and macrophytes). Organisms were identified to the lowest taxonomic level possible (referred to hereafter as “species”) and recorded in a way that considers their functional role in the community (Appendix S1). For analysis, 82 of the 184 species occurred fewer than three times across all samples and were removed (56 singletons, 26 doubletons). Further, five unknown taxa were removed (all occurring fewer than four times across all samples; see *Sensitivity analyses*). Without more information, associations with/between unknown taxa could not be compared to empirical interactions, and thus were excluded from further analysis, resulting in 97 species.

Summary and implementation of association methods

Given that there is no consensus “best” method for association analysis, we selected the most widely used approaches and several new ones that claim to make significant improvements over previous methods (for more information, see reviews by Faust and Raes [2012], Veech [2014]). To promote reproducibility, we chose methods implemented in a freely available programming language, R (R Core Team 2017). These criteria resulted in nine methods that we generalized to three categories (Table 1). We applied each method to the community data described above. We followed all implementation recommendations and defaults as they are described in each original method paper and as are operationalized in the literature. We conduct several sensitivity analyses below, though it is not

TABLE 1. Comparison of model performance among association methods, comparing associations with all empirically determined species interactions.

Category and association method	No. species associations	Network connectance	False positive rate	False negative rate	Model precision	Model recall
Null model						
Pairwise co-occurrence	160	0.045	0.017*	1.00	0.006	0
Frequency distribution	9,044	0.97	0.497*	0.645*	0.018+	0.355+
Combinatorics	676	0.11	0.068*	0.998*	0.030+	0.002+
Odds ratio	903	0.52	0.090*	0.998	0.017+	0.002+
Correlation						
Spearman's	54	0.098	0.006*	1.00+	0*	0*
Pearson's	50	0.12	0.006*	1.00+	0*	0*
NC score	68	0.097	0.007*	1.00+	0.029+	0*
Partial correlation						
JSDM residuals	68	0.19	0.007*	1.00+	0.061+	0*
Partial correlation	8,646	0.93	0.486*	0.907*	0.018+	0.193+

Notes: Each method was compared to a simulated network that randomly assigned interactions among pairs, controlling for the connectance of each association network. Methods that had lower (higher) rates than expected by chance are designated with * (+), at $\alpha = 0.05$. See Appendix S2: Table S5 for association method performance relative to net and direct empirical interactions.

within the scope of this paper to improve previous methods, nor is it the goal to develop a “best practices” for association analysis (though see *Discussion*).

We summarize each method here. Further details are available in Appendix S2 and all data to run these models are available (<https://doi.org/10.6084/m9.figshare.5727051>). *Null model methods* are based on permutation of the observed data, to determine whether the permuted data differ from the observed data. We implemented four null model methods: pairwise co-occurrence (Gotelli and Ulrich 2010, von Gager et al. 2015), frequency distribution (Sfenthourakis et al. 2006) combinatorics (Veech 2013), and odds ratio (Lane et al. 2014). *Correlation-based methods* infer an association between two species if their occurrences are significantly correlated. Largely used in microbial studies, and based on network reconstruction methods common to studies of gene expression (Faust and Raes 2012), correlation statistics are calculated among the distributions of all taxa, with corrections for false discovery. In this study, we implemented a general form of these correlation methods using three different metrics: Pearson's and Spearman's correlation (Steele et al. 2011, Faust and Raes 2012) and the NC score (Schwager et al. 2014), which uses a metric of interaction similar to that of Gotelli and Ulrich (2010). Finally, *partial correlation methods* take into account multiple drivers, which may be responsible for apparent observed correlations among species occurrences. After accounting for the effect that the environment or indirect species interactions may have in determining the occurrence of any one species, the remaining association among occurrences is used to infer a direct association (Ovaskainen et al. 2010, Pollock et al. 2014, Morueta-Holme et al. 2016). We implemented two methods: residual covariance from a joint species distribution model (JSDM; Golding and Harris 2015) and partial correlation with significance evaluated by a randomized null model.

Analysis of co-occurrence results

We first compared the structure of species association networks generated by association methods to the structure of previously published species interaction networks. We quantified structure using two aggregate properties: connectance, a metric describing the relationship between number of species and number of interactions among the species, and the proportion of positive and negative links in the association network, previously interpreted to be a signal of assembly mechanism (e.g., Lyons et al. 2016). We calculated connectance (C) from the number of species (S) and the total number of links in each network (L) such that $C = L/(S \times [S - 1])$. We compared these estimates with connectance of previously published food webs (Dunne et al. 2002), plant–pollinator networks (Olesen et al. 2006), and whole-community networks (all non-trophic and trophic interactions among all species; Sander et al. 2015, Kéfi et al. 2016). We then compared the proportion of positive and negative links to those observed in two fully parameterized non-trophic interaction networks (Sander et al. 2015, Kéfi et al. 2016).

We then quantified the differences among association networks using an analog of beta-diversity. We calculated the turnover in overall network structure among all methods as a function of (1) turnover in which species interact (i.e., different methods detected different species as significantly associated), or (2) turnover in the sign of species interactions (i.e., different methods estimated different associations between the same species pairs; Poisot et al. 2012). Network turnover ranges from 0 to 1, representing networks that range from sharing no species or interactions in common, to networks that are identical in their species composition or in sign of interactions. Network turnover statistics were calculated using the “betalink” R package (Poisot et al. 2012).

Comparison of associations and experimentally determined interactions

After assessing structure of networks and turnover among networks, we compared species associations from each method with those from a data set of experimental species interactions. To generate this data set, we searched the scientific literature for non-trophic species interactions between any two of the species in our observational surveys, using each scientific name as a keyword in the ISI Web of Science Database and supplemented with data from archived theses. Non-trophic interactions in this ecosystem include both negative interactions (e.g., interference and exploitative competition, biological disturbances) and positive interactions (e.g., recruitment facilitation, stress amelioration, habitat provisioning). We included only non-trophic effects in this study, as the theory of inferring interactions from species co-occurrence was derived only for species engaged in non-feeding interactions (Diamond 1975, Diamond and Gilpin 1982). Although some species included in our study are known consumers, consumers in the Oregon low intertidal zone have little impact on spatial interactions among algae and invertebrates (Menge et al. 2005). To ensure that the species pool for the studies would be consistent, we restricted the geographic extent of relevant studies to the Mendocinian biogeographic province within which our surveys were conducted (Fenberg et al. 2015).

We included both field and laboratory/mesocosm studies, but only included manipulative experiments or mensurative studies of species interactions. Mensurative studies were included only if there was an a priori proposed interaction mechanism. An example mensurative study for this system is examining the strength of recruitment facilitation of species *i* by species *j* by counting the number of recruits of species *i* on different facilitators, including on species *j* (e.g., Barner et al. 2016). For each species pair, we recorded the sign of the interaction (1, 0, -1), whether the interaction was measured reciprocally, and the type of study (mensurative/manipulative). We also recorded when experiments estimated direct effects (only two species in the experiment) or net effects (one species manipulated, effect tracked for entire community). With these criteria, we found 366 non-trophic interactions among 66 taxa. A summary of these data is found in Appendix S2: Table S2, and the entire data set is available at (<https://doi.org/10.6084/m9.figshare.5727051>). Unlike the whole-community networks of Sander et al. (2015) and Kéfi et al. (2016), our non-trophic interaction network includes only experimental data and thus, despite including data from 39 studies, likely under-samples the interactions in this community. In particular, the connectance of the empirical network is 0.039, an order of magnitude lower than the connectance of published whole-community networks (see *Discussion*).

We then asked, do any of the species associations and experimental interactions match, either in terms of the species pair and the sign of the interaction? For each

method, we calculated the false positive and false negative rates. A false positive rate is the rate of inferring of an association not found in the experimental data set (false positives), scaled by the sum of false positives and true negatives (calculated as total possible number of interactions minus number of empirical interactions). False negative rate is the number of empirical interactions an association method failed to detect (false negatives), scaled by the sum of false negatives and true positives (calculated as the number of empirical interactions). For comparison with similar studies (e.g., Sander et al. 2017), we calculated model precision (true positives/true positives + false positives) and model recall (true positives/true positives + false negatives). If most of the interactions identified in the association network are included in the empirical network, the association method would have high model precision. Association methods with high model recall would both identify most of the interactions in the empirical network.

We compared these metrics against those generated by an Erdős-Rényi model, which randomly assigns interactions between species with a fixed probability (Gilbert 1959). We generated 999 random networks of the same size as our species pool ($S = 97$), using each association network connectance (Table 1) as the interaction probability. Each association network was compared with three versions of the empirical data set: the full data set, net effects experiments, and pairwise experiments.

Sensitivity analyses

We tested for the sensitivity of our results to several factors. First, as environmental conditions could drive observed co-occurrences (“habitat filtering”), we tested for the sensitivity of our results to the inclusion of environmental covariates in each co-occurrence model. Second, we tested for the sensitivity of our results to the spatial grain size of our surveys (Appendix S1). We compared association network inference for all methods using data at 5×5 , 10×10 , 15×15 , and 20×20 against the 25×25 cm² results presented in the *Results*. Third, as co-occurrence inference can be sensitive to species inclusion or exclusion (e.g., Diamond and Gilpin 1982), we tested for the sensitivity of our inclusion criteria by (1) including the five unidentifiable taxa, and (2) splitting our survey data into two sets: only competing species, only facilitators (Appendix S1). Finally, for the association methods that rely on comparison to a null model, we tested whether our results were sensitive to the choice of null model algorithm. All the sensitivity analysis details and results can be found in Appendix S2, as these methodological considerations did not qualitatively change the results and conclusions we present.

RESULTS

Of the 97 species in the survey, association analysis identified a large range of associated species, from 19 to

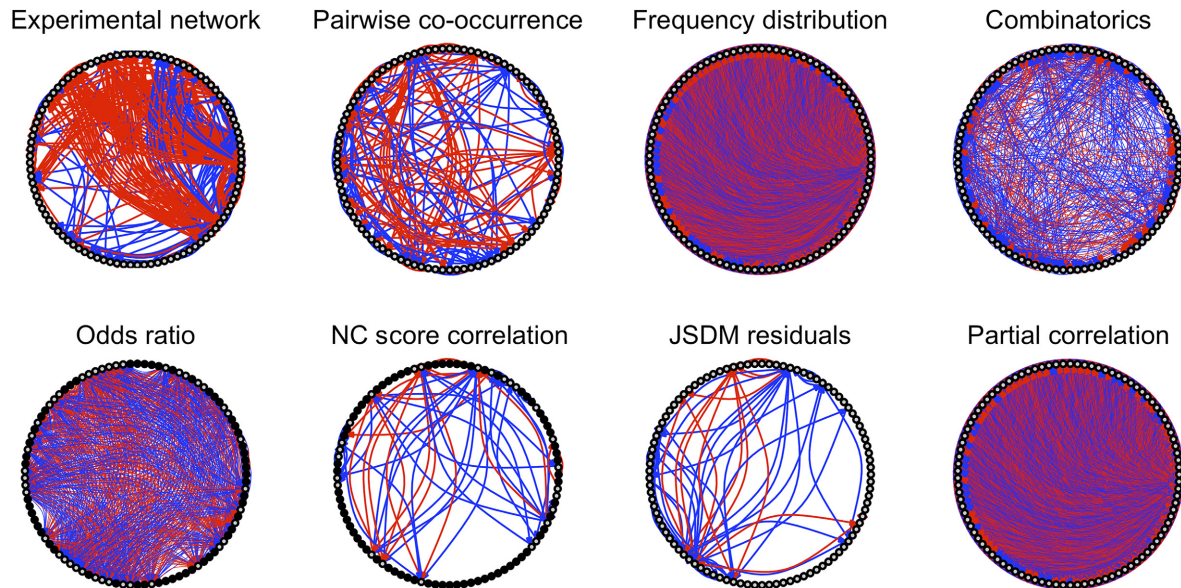


FIG. 1. High turnover among species interaction and association networks. Each network shows the same 97 taxa. Arrow direction indicates the effect of species i on species j . Links are coded blue (red) for positive (negative) effects. Both the frequency distribution and the partial correlation methods detected close to the maximum number of pairs, leading to a high density of links. The interaction network depicted includes all interactions in the empirical database. See Appendix S2: Fig. S1 for Spearman's and Pearson's correlation networks and Appendix S2: Fig. S2 for node labels. Species coded in black were removed before association analysis, based on the default implementation of odds ratio and correlation methods (Appendix S2). NC score is described in NC Score Schwager et al. (2014); JSDM residuals, joint species distribution model.

all 97, with the number of associations among those species ranging between 44 and 9,016 (Fig. 1). Compared with previously quantified species interaction networks, most association networks had structure that fell within the range of previous studies, both in the number of inferred associations and the proportional sign of those

interactions (Fig. 2). However, two of the nine association methods tested (frequency distribution and partial correlation with randomized null) generated networks with hundreds to thousands more interactions than found in even fully quantified species interaction networks (indicated with arrow in Fig. 2, Table 1).

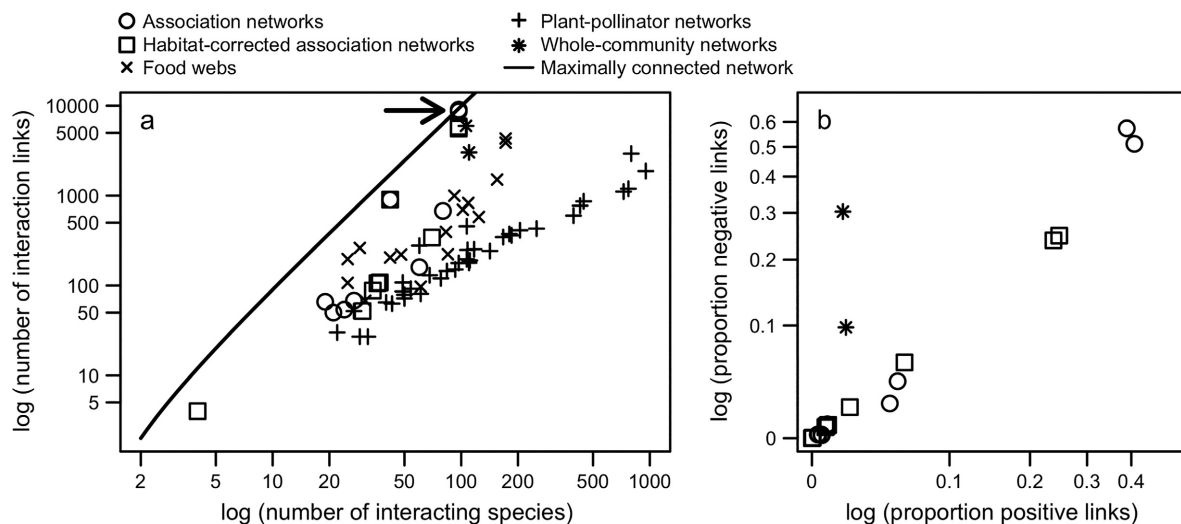


FIG. 2. (a) Most association methods estimated numbers of interaction per network size within the range of published food webs (Dunne et al. 2002), plant–pollinator networks (Olesen et al. 2006), and whole-community networks (Sander et al. 2015, Kéfi et al. 2016). (b) Association methods generally predicted similar proportions of positive links but under-predicted the number of negative links when compared to the non-trophic links of whole-community networks. Arrow indicates the two association methods with higher connectance than whole-community networks.

Species association networks differed widely among methods (mean overall network turnover = 0.78; 95% CI: 0.70–0.87; Fig. 1). Each method estimated different pairs of associated species and different signs of associations among species, even among methodologically similar approaches (Appendix S2). The mean turnover in species identity was 0.40 among pairs of association networks (95% CI 0.33–0.46), though only five species pairs were identified by all methods, out of a possible 9,312 possible pairs. This inconsistency is reflected in the overall high dissimilarity among methods in the sign of the associations between species (mean sign turnover 0.60, 95% CI 0.52–0.69). Among species that were identified

by multiple methods, the sign of associations between these shared species were up to 92% dissimilar. Two species pairs were identified as having the same sign of association by all methods: both *Corallina vancouveriensis* (a coralline algal turf)/*Dilsea californica* (an algal turf) and *Balanus glandula*/*Semibalanus cariosus* (barnacles) were identified as being positively associated.

Although most association methods performed better than models that randomly assigned interactions among species pairs (Table 1; Appendix S2: Table S5), species associations matched few species interactions in the experimental data set, relative to the total number of inferred associations (low model precision; Fig. 3). For example,

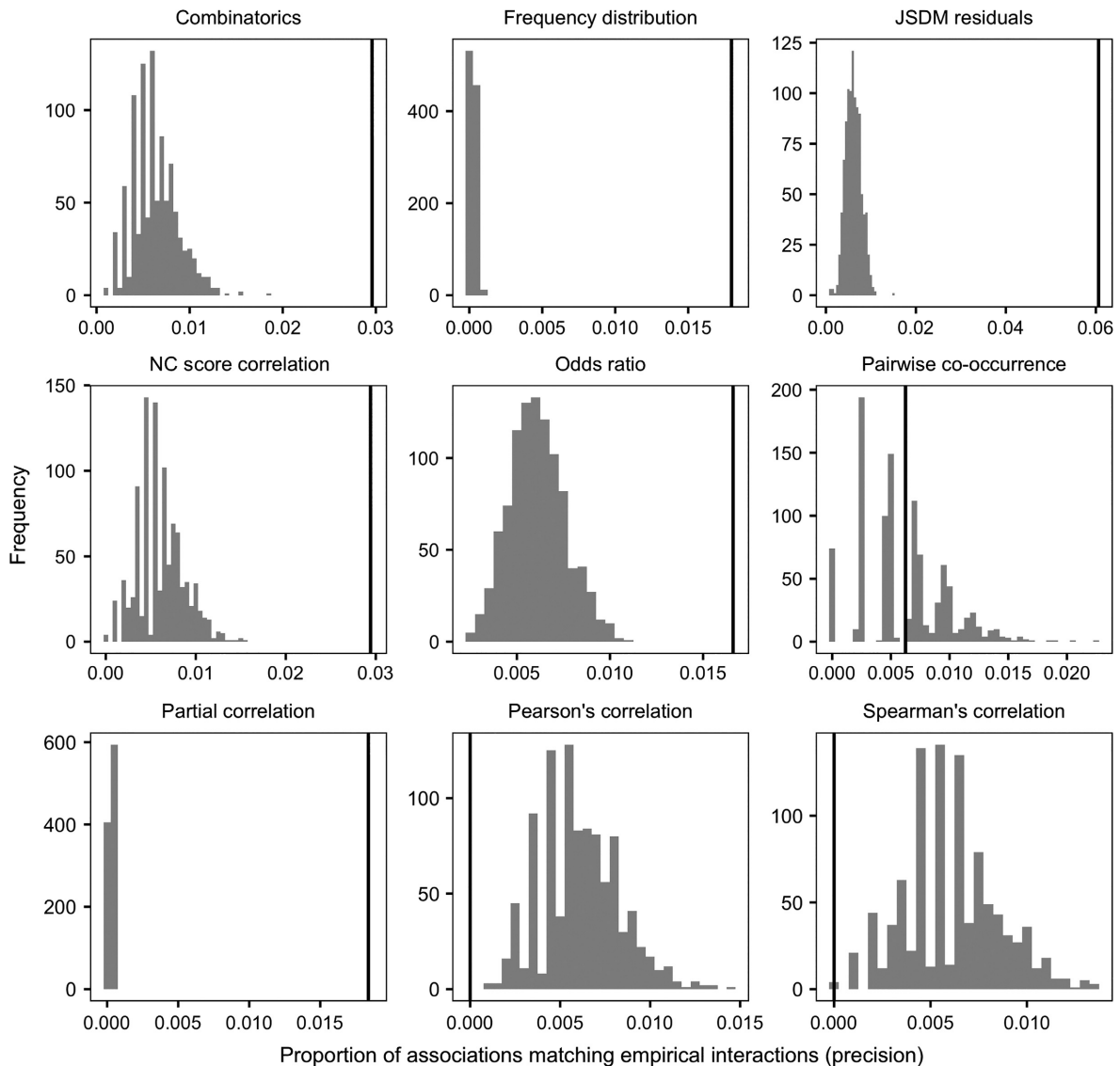


FIG. 3. The proportion of associations matching species experimental interactions (model precision) ranged between 0% and 6% across all association methods. For each association method, the model precision of the association method is plotted as a vertical line, against the distribution of precisions of 999 random models. Six methods performed better than random models (higher model precision), two methods performed worse (lower model precision than random). See Appendix S2: Table S5 for comparison of associations and interactions from net or direct interaction experiments.

although the frequency distribution method matched 162 associations with empirical interactions, model precision was low, due to the number of false positives (8,882 false positives; Fig. 3). Further, methods estimated many associations that were not in our experimental data set (low model recall) (Fig. 1, Table 1). No method matched more than 27% of associations with empirical interactions, and representing only 1.8% of the total associations detected by that method (partial correlation).

Some methods estimated many species associations, some estimated few (Fig. 2, Table 1), but no approach increased model precision or model recall. Even among species pairs shared by an association network and the experimental data set, association methods detected the opposite interaction sign 76% of the time (95% CI 0.66, 0.85; Appendix S2: Table S4). In other words, association methods frequently estimated competing species pairs as mutualistic and vice-versa. *Balanus glandula* and *Semibalanus cariosus*, although identified by all nine association methods above as being positively associated, experimentally compete for space (Dayton 1971). Only the odds ratio method could estimate different reciprocal interactions, but detected associations in the opposite direction from experimental interactions 60% of the time (Appendix S2: Table S4). This high dissimilarity between association networks and the experimentally determined interaction was not sensitive to the interaction type (net effects experiments vs. pairwise experiments; Appendix S2: Table S5).

DISCUSSION

In this study, we assessed a set of approaches to estimate non-trophic species interactions from easily collected and widely available spatial occurrence data. We found that inference of species interactions is highly sensitive to the statistical method used, with no “best” method. We suggest that analysis of spatial co-occurrence infers a fundamentally different relationship among species (an association) than is estimated with species interaction experiments.

Each of the most commonly used association methods detected a different set of species associations. A large part of this inconsistency arises out of the ways in which species associations are estimated, though association networks varied widely even among a suite of conceptually similar methods (Weiss et al. 2016; Appendix S2). With no consensus method, the results of any one study in this growing field of research will be difficult to compare to, and may conflict with, a similar study that uses a different association method. For this reason, we asked whether any one association method was better at reproducing general properties of ecological networks than any other method. We first considered whether association methods generally followed empirical network scaling relationships. Promisingly, most methods estimated similar proportions of positive interactions to those found in fully parameterized non-trophic species

interaction networks from rocky intertidal systems in Chile and Washington, USA (Sander et al. 2015, Kéfi et al. 2016) and estimated numbers of interactions that fall within ranges expected from the literature. For two association methods, connectance was much larger than in previous studies, nearly the maximum possible number of interspecific interactions for a fully connected network (frequency distribution, partial correlation). For comparison, connectance in empirical interaction networks is generally around 0.1 and rarely >0.3 (Dunne et al. 2002, Olesen et al. 2006, Sander et al. 2015, Kéfi et al. 2016). However, because so few fully parameterized interaction networks exist, the question remains as to whether high connectance in these association networks reflects reality or statistical artifact.

Unlike previous modeling studies using simulated species interactions (Faisal et al. 2010, Cazelles et al. 2016, Harris 2016), we found little overlap between estimates of species associations and experimental estimates of species interactions. The few experimentally determined species interactions that did match those generated by co-occurrence analyses tended to be positive interactions. For example, association analysis identified the positive thermal amelioration of *Katharina tunicata* (a mobile invertebrate chiton) by *Saccharina sessilis* (a canopy-forming alga) and the recruitment facilitation of *Phyllospadix* spp. (surfgrass) by *Neorhodomela larix* (a turf-like alga; Appendix S2: Table S2). However, when an association matched an interaction in the empirical data set, the empirical reciprocal interaction had an opposite sign most of the time. These example “positive-positive” interactions are in reality “positive-negative”: *K. tunicata* bulldozes *S. sessilis* and *Phyllospadix* outcompetes *N. larix*.

The reason for the discrepancies between species association estimates and species interaction estimates in this study could include a mix of limitations both of the association methods and of our study design, or both. We conducted a series of sensitivity analyses to address these limitations. Inference was not improved by adjusting the spatial grain size of co-occurrence observation, species inclusion criteria, or null model choice, nor by accounting for habitat filtering (Appendix S2). However, discrepancies may still be a function of (1) limitations of the empirical interaction data set, (2) dilution of interaction signals, or (3) factors that neither association analysis nor experiments were designed to detect. First, a full experimental accounting of all direct interactions in this community, though logistically infeasible, might improve matching between associations and interactions. A similar study found higher precision and recall than in the present study, when comparing temporal associations with a full empirical interaction network (Sander et al. 2017). However, to generate a full interaction network, interactions were assigned with a mix of observational data, experimental results, and expert knowledge (Sander et al. 2015). We do show that certain association methods estimate a realistic number of interactions and

match empirical interactions at a rate higher than random models (Figs. 2 and 3), suggesting that our comparisons are reasonable. We tested whether species associations reflect net species interactions rather than direct interactions (Cazelles et al. 2016, Harris 2016). Model performance did not improve when we restricted the interaction data set to include only net interaction experiments (by the removal of one species from a community) or only pairwise experiments (Appendix S2). Second, the signal of spatial interactions may have been “diluted” by the inclusion of so many species (Diamond and Gilpin 1982), but our results were not changed when we included species expected to have positive (or negative) interactions (Appendix S2). Finally, both association analysis and experiments are limited in their ability to detect interactions that change in space and time, common in natural systems (Chamberlain et al. 2014). Further, neither approach accounts for the non-pairwise, higher-order interactions thought to be strong in driving community dynamics (Mayfield and Stouffer 2017).

In our choice of study system, we expected to maximize the probability of detecting a signal of spatial interactions in spatial co-occurrence. Specifically, the Oregon rocky intertidal ecosystem is characterized by sharp spatial zonation produced by abiotic stress and predation. Within the low intertidal zone, non-trophic interactions have demonstrated spatial effects (Menge et al. 2005), but the spatial imprints of these interactions were not consistently distinguishable as associations in the present study. Such a result is in stark contrast to the effect of keystone predation in this system, which generates a clear transition between the diverse low intertidal zone and the less diverse mid intertidal zone dominated by the mussel *Mytilus californianus* (Paine 1966). Association analysis across zones, at the scale of tens of meters, likely would have detected competitive exclusion by *Mytilus* as a negative association. However, our coarsest analysis was conducted at the scale of centimeters, a relevant spatial scale for intertidal organisms with relatively little mobility and the standard scale at which experimental manipulations occur. Although no one spatial grain size performed better than another, when we combined all associations estimated by all methods across all grain sizes, 71% of empirical interactions were matched. This suggests that interactions between species play out at different spatial scales, and only with a priori knowledge of the best spatial or temporal scale at which to detect an interaction, may association analysis correctly detect that interaction.

Pattern and process in community ecology

Association methods are increasingly being used to estimate species interactions, with multiple applications across disparate disciplines. In this study, we show that there is no relationship between experimentally determined species interactions and species associations. This may be unsurprising, given that these two modes of

inference, association analysis and species interaction experiments, have carried on independently for decades and reflect differences in philosophy of ecology, spatial scales of study, and available tools.

If associations do not map directly to pairwise empirical species interactions, can broader mechanisms of community organization can be inferred from associations? Although our study suggests that associations may not accurately predict interaction links, we found that association analysis may predict aggregate community statistics such as: the total number of interacting species, number of interactions, and the proportion of positive and negative interactions. As such, association analysis may be a way to estimate higher-level properties, if such properties can be rigorously linked to function or mechanism. Recent papers suggest that frequencies of positive or negative associations reflect assembly mechanism (Levy and Borenstein 2013, Zelezniak et al. 2015, Lyons et al. 2016), but no study to date has determined which processes lead to more positive or negative associations. In other studies, changes in associations are thought to signal shifts in system stability (Griffith et al. 2018, Kay et al. 2018).

The difficulty in inferring structuring processes from community pattern is not isolated to association analysis, but is widespread and deeply embedded in the history of ecology. Similar inferential challenges plague the analysis of functional trait dispersion and phylogenetic community analysis (Mayfield and Levine 2010, Adler et al. 2013), where competition is inferred to be a primary driver of community assembly if species in the community are more “similar” (evolutionarily or functionally) than expected by chance. Observational assembly analyses operate on the assumption that the influence of species interactions on community structure is strong enough to be detectable over the influence of other drivers. Such analyses often rely on a snapshot of extant community structure to infer the processes that generate such structure (though temporally explicit methods may not improve inference; Sander et al. 2017). Community formation and maintenance is dynamic and involves many processes that may not be easily disentangled: trophic and non-trophic interactions, habitat filtering, regional and evolutionary constraints, priority effects, and dispersal. These other drivers could either nullify community patterns produced by interactions, or could themselves produce patterns indistinguishable from those expected to be produced by interactions (Mayfield and Levine 2010, Adler et al. 2013). For example, dispersal can produce checkerboard co-occurrence patterns, the same pattern that is interpreted as a signal of competition (Schamp et al. 2015).

The challenge of reconciling experimental results with statistical tests for pattern in nature is not easily overcome. When, as in this study, observed spatial patterns do not seem to manifest underlying mechanisms, what are the next steps? We found that no one method outperformed another; even statistically sophisticated tools did

not improve model performance, suggesting that model development may not lead to estimates of associations that match empirical interactions. As emphasized in the observational study of community assembly (Mayfield and Levine 2010, Adler et al. 2013), we lack theory to distinguish among a myriad of drivers that may produce similar patterns in community structure, limiting our ability to mechanistically interpret pairwise associations. Further, although methodological best practices are well defined for null-model analysis of whole co-occurrence matrices (e.g., Gotelli 2000), no such common best practices exist for the practice of association analysis (but see Blois et al. 2014, Morueta-Holme et al. 2016). Such recommendations require further development outside the scope of this study, but important next steps include testing for the influence of spatial or temporal scale on inference, development of independent simulation models that reflect a variety of assembly mechanisms including the role of trophic interactions, and examining sensitivity to species inclusion or exclusion.

ACKNOWLEDGMENTS

We are grateful to K. Beem, K. Belica, A. Brown, E. Cerny-Chipman, S. Cook, M. Frenock, S. Gerrity, D. Howe, N. McKee, J. Reimer, J. Robinson, and J. Sapp for assistance collecting and entering data and with computation. Feedback from L. Thurman, D. Harris, A. Iles, T. Dallas, K. Norman, N. Martinez, M. Novak, and multiple anonymous reviewers were of particular help in shaping this manuscript. Funding for this study was provided by an EPA STAR Graduate Research Fellowship (A. K. Barner), the Oregon State University Integrative Biology Graduate Research Fund (A. K. Barner), an NSF Graduate Research Fellowship (KEC), NSF grants OCE-0726983 and OCE-1061233 (B. A. Menge, S. D. Hacker), the David and Lucile Packard Foundation (B. A. Menge), and the Wayne and Gladys Valley Foundation (B. A. Menge). This is publication 480 from PISCO, the Partnership for Interdisciplinary Studies of Coastal Oceans, funded primarily by the Packard Foundation.

LITERATURE CITED

- Adler, P. B., A. Fajardo, A. R. Kleinhesselink, and N. J. B. Kraft. 2013. Trait-based tests of coexistence mechanisms. *Ecology Letters* 16:1294–1306.
- Araújo, M. B., A. Rozenfeld, C. Rahbek, and P. A. Marquet. 2011. Using species co-occurrence networks to assess the impacts of climate change. *Ecography* 34:897–908.
- Barner, A. K., S. D. Hacker, B. A. Menge, and K. J. Nielsen. 2016. The complex net effect of reciprocal interactions and recruitment facilitation maintains an intertidal kelp community. *Journal of Ecology* 104:33–43.
- Bell, J. R., R. Andrew King, D. A. Bohan, and W. O. C. Symondson. 2010. Spatial co-occurrence networks predict the feeding histories of polyphagous arthropod predators at field scales. *Ecography* 33:64–72.
- Blois, J. L., et al. 2014. A framework for evaluating the influence of climate, dispersal limitation, and biotic interactions using fossil pollen associations across the late Quaternary. *Ecography* 37:1095–1108.
- Cazelles, K., M. B. Araújo, N. Mouquet, and D. Gravel. 2016. A theory for species co-occurrence in interaction networks. *Theoretical Ecology* 9:39–48.
- Chamberlain, S. A., J. L. Bronstein, and J. A. Rudgers. 2014. How context dependent are species interactions? *Ecology Letters* 17:881–890.
- Connor, E. F., and D. Simberloff. 1979. The assembly of species communities: Chance or competition? *Ecology* 60:1132–1140.
- Dayton, P. K. 1971. Competition, disturbance, and community organization: the provision and subsequent utilization of space in a rocky intertidal community. *Ecological Monographs* 41:351–389.
- Diamond, J. M. 1975. Assembly of species communities. Pages 342–344 in J. M. Diamond and M. L. Cody, editors. *Ecology and evolution of communities*. Belknap Press, Cambridge, Massachusetts, USA.
- Diamond, J. M., and M. E. Gilpin. 1982. Examination of the “null” model of Connor and Simberloff for species co-occurrences on islands. *Oecologia* 52:64–74.
- Dunne, J. A., R. J. Williams, and N. D. Martinez. 2002. Food-web structure and network theory: the role of connectance and size. *Proceedings of the National Academy of Sciences USA* 99:12917–12922.
- Faisal, A., F. Dondelinger, D. Husmeier, and C. M. Beale. 2010. Inferring species interaction networks from species abundance data: A comparative evaluation of various statistical and machine learning methods. *Ecological Informatics* 5:451–464.
- Faust, K., and J. Raes. 2012. Microbial interactions: from networks to models. *Nature Reviews Microbiology* 10:538–550.
- Fenberg, P. B., B. A. Menge, P. T. Raimondi, and M. M. Rivadeneira. 2015. Biogeographic structure of the northeastern Pacific rocky intertidal: the role of upwelling and dispersal to drive patterns. *Ecography* 38:83–95.
- Gilbert, E. N. 1959. Random graphs. *Annals of Mathematical Statistics* 30:1141–1144.
- Golding, N., and D. J. Harris. 2015. BayesComm: Bayesian community ecology analysis. R package. Version 0.1-2. <https://cran.r-project.org/web/packages/BayesComm/index.html>
- Gotelli, N. J. 2000. Null model analysis of species co-occurrence patterns. *Ecology* 81:2606–2621.
- Gotelli, N. J., and W. Ulrich. 2010. The empirical Bayes approach as a tool to identify non-random species associations. *Oecologia* 162:463–477.
- Griffith, G. P., P. G. Strutton, and J. M. Semmens. 2018. Climate change alters stability and species potential interactions in a large marine ecosystem. *Global Change Biology* 24:e90–e100.
- Harris, D. J. 2016. Inferring species interactions from observational data with Markov networks. *Ecology* 97:3308–3314.
- Kay, G. M., A. Tulloch, P. S. Barton, S. A. Cunningham, D. A. Driscoll, and D. B. Linden. 2018. Species co-occurrence networks show reptile community reorganization under agricultural transformation. *Ecography* 41:113–125.
- Kéfi, S., V. Miele, E. A. Wieters, S. A. Navarrete, and E. L. Berlow. 2016. How structured is the entangled bank? The surprisingly simple organization of multiplex ecological networks leads to increased persistence and resilience. *PLoS Biology* 14:e1002527.
- Lane, P. W., D. B. Lindenmayer, P. S. Barton, W. Blanchard, and M. J. Westgate. 2014. Visualization of species pairwise associations: a case study of surrogacy in bird assemblages. *Ecology and Evolution* 4:3279–3289.
- Laska, M. S., and J. T. Wootton. 1998. Theoretical concepts and empirical approaches to measuring interaction strength. *Ecology* 79:461–476.
- Levy, R., and E. Borenstein. 2013. Metabolic modeling of species interaction in the human microbiome elucidates community-level assembly rules. *Proceedings of the National Academy of Sciences USA* 110:12804–12809.

- Lima-Mendez, G., et al. 2015. Determinants of community structure in the global plankton interactome. *Science* 348: 1262073-1–1262073-9.
- Lyons, S. K., et al. 2016. Holocene shifts in the assembly of plant and animal communities implicate human impacts. *Nature* 529:80–83.
- Mayfield, M. M., and J. M. Levine. 2010. Opposing effects of competitive exclusion on the phylogenetic structure of communities. *Ecology Letters* 13:1085–1093.
- Mayfield, M. M., and D. B. Stouffer. 2017. Higher-order interactions capture unexplained complexity in diverse communities. *Nature Ecology and Evolution* 1:62.
- Menge, B. A., G. W. Allison, C. A. Blanchette, T. M. Farrell, A. M. Olson, T. A. Turner, and P. van Tamelen. 2005. Stasis or kinesis? Hidden dynamics of a rocky intertidal macrophyte mosaic revealed by a spatially explicit approach. *Journal of Experimental Marine Biology and Ecology* 314:3–39.
- Morales-Castilla, I., M. G. Matias, D. Gravel, and M. B. Araújo. 2015. Inferring biotic interactions from proxies. *Trends in Ecology and Evolution* 30:349–356.
- Moruela-Holme, N., B. Blonder, B. Sandel, B. J. McGill, R. K. Peet, J. E. Ott, C. Violle, B. J. Enquist, P. M. Jørgensen, and J.-C. Svenning. 2016. A network approach for inferring species associations from co-occurrence data. *Ecography* 39: 1–12.
- Olesen, J. M., J. Bascompte, Y. L. Dupont, and P. Jordano. 2006. The smallest of all worlds: pollination networks. *Journal of Theoretical Biology* 240:270–276.
- Ovaskainen, O., J. Hottola, and J. Siitonen. 2010. Modeling species co-occurrence by multivariate logistic regression generates new hypotheses on fungal interactions. *Ecology* 91: 2514–2521.
- Paine, R. T. 1966. Food web complexity and species diversity. *American Naturalist* 100:65–75.
- Poisot, T., E. Canard, D. Mouillot, N. Mouquet, D. Gravel, and F. Jordan. 2012. The dissimilarity of species interaction networks. *Ecology Letters* 15:1353–1361.
- Pollock, L. J., R. Tingley, W. K. Morris, N. Golding, R. B. O'Hara, K. M. Parris, P. A. Vesk, and M. A. McCarthy. 2014. Understanding co-occurrence by modelling species simultaneously with a Joint Species Distribution Model (JSDM). *Methods in Ecology and Evolution* 5:397–406.
- R Core Team. 2017. R: A language and environment for statistical computing. Version 3.4.0. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>
- Sander, E. L., J. T. Wootton, and S. Allesina. 2015. What can interaction webs tell us about species roles? *PLoS Computational Biology* 11:e1004330.
- Sander, E. L., J. T. Wootton, and S. Allesina. 2017. Ecological network inference from long-term presence-absence data. *Scientific Reports* 7:7154.
- Schamp, B. S., S. E. Arnott, and K. L. Joslin. 2015. Dispersal strength influences zooplankton co-occurrence patterns in experimental mesocosms. *Ecology* 96:1074–1083.
- Schluter, D. 1984. A variance test for detecting species associations, with some example applications. *Ecology* 65: 998–1005.
- Schwager, E., C. Bielski, and G. Weingart. 2014. ccrepe: compositionality corrected by renormalization and permutation. Bioconductor R Package. Version 3.5. <https://bioconductor.org/packages/release/bioc/html/ccrepe.html>
- Sfenthourakis, S., E. Tzanatos, and S. Giokas. 2006. Species co-occurrence: the case of congeneric species and a causal approach to patterns of species association. *Global Ecology and Biogeography* 15:39–49.
- Steele, J. A., et al. 2011. Marine bacterial, archaeal and protistan association networks reveal ecological linkages. *ISME Journal* 5:1414–1425.
- von Gagn, M., M. von Gagn, and A. Schmitz Ornés. 2015. Problems with bins: A critical reassessment of Gotelli and Ulrich's Bayes approach using bird data. *Acta Oecologica* 69:137–145.
- Veech, J. A. 2013. A probabilistic model for analysing species co-occurrence. *Global Ecology and Biogeography* 22:252–260.
- Veech, J. A. 2014. The pairwise approach to analysing species co-occurrence. *Journal of Biogeography* 41:1029–1035.
- Weiss, S., et al. 2016. Correlation detection strategies in microbial data sets vary widely in sensitivity and precision. *ISME Journal* 10:1669–1681.
- Wootton, J. T. 2004. Markov chain models predict the consequences of experimental extinctions. *Ecology Letters* 7:653–660.
- Wootton, J. T., and M. Emmerson. 2005. Measurement of interaction strength in nature. *Annual Review of Ecology, Evolution, and Systematics* 36:419–444.
- Zelezniak, A., S. Andrejev, O. Ponomarova, D. R. Mende, P. Bork, and K. R. Patil. 2015. Metabolic dependencies drive species co-occurrence in diverse microbial communities. *Proceedings of the National Academy of Sciences USA* 112: 6449–6454.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of this article at <http://onlinelibrary.wiley.com/doi/10.1002/ecs.2133/supinfo>

DATA AVAILABILITY

Survey and experimental data associated with this study are available from Figshare at <https://doi.org/10.6084/m9.figshare.5727051.v1>. Example R code to aggregate these data are available from Zenodo at <https://doi.org/10.5281/zenodo.996879>.