

Title: Bringing Elton and Grinnell together: a quantitative framework to represent the biogeography of ecological interaction networks

Authors: Dominique Gravel^{1,2,*}, Benjamin Baiser³, Jennifer A. Dunne⁴, Jens-Peter Kopelke⁵, Neo D. Martinez⁶, Tommi Nyman⁷, Timothée Poisot^{2,8}, Daniel B. Stouffer⁹, Jason M. Tylianakis^{9,10}, Spencer A. Wood¹¹, Tomas Roslin¹²

1: Canada Research Chair in Integrative Ecology. Département de biologie, Université de Sherbrooke, 2500 Boulevard l'Université, Sherbrooke (Québec). J1K 2R1

2: Québec Centre for Biodiversity Science

3: University of Florida, Department of Wildlife Ecology and Conservation, 110 Newins-Ziegler Hall, PO Box 110430, Gainesville, Fl. 32611-0430

4: Santa Fe Institute, 1399 Hyde Park Road, Santa Fe, NM 87501, USA.

5: Senckenberg Research Institute and Natural History Museum, Senckenberganlage 25, D-60325 Frankfurt am Main, Germany

6: Ecology and Evolutionary Biology Department, University of Arizona, P.O. Box 210088, Tucson, AZ 85721, USA

7: University of Eastern Finland, Department of Environmental and Biological Sciences, P.O. Box 111, FI-80101 Joensuu, Finland

8: Université de Montréal, Département des Sciences Biologiques, 90 Avenue Vincent d'Indy, Montréal, QC H2V3S9, Canada.

9: Centre for Integrative Ecology, School of Biological Sciences, University of Canterbury, Private Bag 4800, Christchurch, New Zealand

10: Department of Life Sciences, Imperial College London, Silwood Park Campus, Buckhurst Road, Ascot, Berkshire SL5 7PY, United Kingdom

11: University of Washington, School of Environmental and Forest Sciences, Box 352100, Seattle, WA 98195, USA

12: Department of Ecology, Swedish University of Agricultural Sciences, Box 7044, 750 07 Uppsala, Sweden

Keywords: networks, spatial ecology, co-occurrence, probability of interaction

Authorship: DG, BB, JAD, NDM, TP, SAW, DBS and JMT contributed to the concepts; JPK, TN and TR contributed the data; DG developed the probabilistic approach and performed the analysis; DG, TP and TR wrote the first draft and all of the authors edited the manuscript.

Abstract

Biogeography has traditionally focused on the spatial distribution and abundance of species. Both are driven by the way species interact with one another, but also by the way these interactions vary across time and space. Here, we call for an integrated approach, adopting the view that community structure is best represented as a network of ecological interactions, and show how it translates to biogeography questions. We propose that the ecological niche should encompass the effect of the environment on species distribution (the Grinnellian dimension of the niche) and on the ecological interactions among them (the Eltonian dimension). Starting from this concept, we develop a quantitative theory to explain turnover of interactions in space and time – *i.e.* a novel approach to interaction distribution modelling. We apply this framework to host–parasite interactions across Europe and find that two aspects of the environment (temperature and precipitation) exert a strong imprint on species co-occurrence, but not on species interactions. Even where species co-occur, interaction proves to be stochastic rather than deterministic, adding to variation in realized network structure. We also find that a large majority of host-parasite pairs are never found together, thus precluding any inferences regarding their probability to interact. This first attempt to explain variation of network structure at large spatial scales opens new perspectives at the interface of species distribution modelling and community ecology.

Introduction

Community ecology is *the study of the interactions that determine the distribution and abundance of organisms* (Krebs, 2009). Despite a general consensus on this definition (Scheiner & Willig, 2007), research on variation in community structure (beta-diversity) has mostly focused on the spatial and temporal turnover of species composition (Anderson *et al.*, 2011). Such research usually neglects variation in the way species interact with each other, despite accumulating empirical evidence that this is a major source of diversity (Poisot *et al.*, 2015b). Given this omission, it is perhaps not surprising that biogeographers are still struggling to establish whether interactions actually impact the distribution of species at large spatial scales (Kissling *et al.*, 2012). An interaction is conceived as the direct effect of the action of one species on the demography on another species; whether it scales up and impact the (co-)distribution of species remains matter of debate at the moment (Wisz *et al.*, 2012; ?; Cazelles, 2016; Harris, 2016; Godsoe *et al.*, 2017). Treating interactions as fixed events nonetheless neglects a large part of the complexity of empirical communities, and will most likely deliver underwhelming results (Poisot *et al.*, 2016a). Recent attempts at accounting for interactions in species distribution models (Pollock *et al.*, 2014; Pellissier *et al.*, 2013; Ovaskainen *et al.*) have brought some methodological advances, but are not sufficient for two reasons. First, these techniques are still based on a ‘species-based’ approach to communities, where interactions are merely treated as fixed covariates affecting distribution. Second, they failed to provide a conceptual step forward, both in their treatment of interactions and in the quality of the predictions they make.

Network approaches offer a convenient representation of communities because they simultaneously account for species composition and their interactions. Species are represented as nodes, so that networks already encompass all the information used by current approaches of species distribution modelling; in addition, interactions are represented

by links, so that networks provide additional information on community structure. To date, studies of network diversity have mostly been concerned with the distribution of interactions within locations, and less so with variation among locations (Dunne, 2006; Bascompte & Jordano, 2007; Ings *et al.*, 2009; Kéfi *et al.*, 2012). There is, however, ample evidence that interaction networks vary in space and, time (Laliberté & Tylianakis, 2010; Poisot *et al.*, 2012; Schleuning *et al.*, 2012; Albouy *et al.*, 2014; Poisot *et al.*, 2016b; Trøjelsgaard *et al.*, 2015), even though there is no common framework with which to generalize these results. Metacommunity theory provides explanations for variation in the distribution of the nodes (Gravel *et al.*, 2011; Pillai *et al.*, 2011; Cazelles *et al.*, 2015), but there is no such explanation to the variation of node and link occurrences. Consequently, we need theory to formalize these observations, as it is the only way towards fulfilling the goal of community ecology: providing cogent predictions about, and understanding of, the structure of ecological communities.

Given the historically different approaches to modelling the distributions of species vs. interactions, there is need to bring the two together. Here, we offer an integrated approach to do so, adopting the view that community structure is best represented as a network of ecological interactions. Based on this idea, we propose a description of the ecological niche that integrates the effect of the environment on species distribution and on the ecological interactions among them. Building on this concept, we develop a quantitative theory to explain turnover of interactions in space and time. We first present the concept and then formalize it mathematically, using a probabilistic model to represent the sampling of the regional pool of interactions. At the level of species pairs, the statistical approach could be conceived as an interaction distribution model. At the community level, the approach provides a likelihood-based method to compare different hypotheses of network turnover. As an illustrative example, we apply this novel framework to a large data set on host–parasite interactions across Europe and find that

two aspects of the environment (temperature and precipitation) exert a strong imprint on species co-occurrence, but not on species interactions. The network structure changes systematically across the latitudinal gradient, with a peak of connectance at intermediate latitudes.

The two dimensions of community structure

The problem of community assembly is often formulated as *how are species sampled from a regional pool to constitute a local community* (Götzenberger et al., 2012)? This question could be rewritten to address the problem of network assembly, as *how do samples from a regional pool of interactions constitute a local interaction network*? An illustration of this problem for a food web is provided in Fig. 1. The regional pool of interactions, the *metaweb*, represents potential interactions among all species that could be found in a given area. In this particular case, there are 275 nodes, and 1173 links among the plants (52 nodes), herbivores (96 nodes), and parasitoids (127 nodes) from Northern Europe. An instance of a local community is also illustrated, with 45 nodes and 93 interactions. Only 55.0% of all potential interactions (plant-herbivore or herbivore parasitoid combinations) are realized in this local network, revealing the stochastic nature of ecological interactions. Our objective here is to provide theory to explain the sampling of the regional pool of interactions, along with a quantitative method to predict it. The problem could be formalized sequentially by understanding first why only a fraction of the species co- occur locally and second why these species do or do not interact.

There are multiple causes of spatial turnover of species co-occurrence. The first and most-studied driver is the effect of variation in the abiotic environment on species performance. Combined with specific responses in demography, it generates variation among sites by selecting the locally fittest species (Leibold *et al.*, 2004). Stochasticity plays

132 an additional role, either because colonization and extinction events (Hanski, 1999) are
 133 inherently unpredictable or because strong non-linear feedbacks in community dynamics
 134 generate alternative transients and equilibria (Chase, 2007; Vellend *et al.*, 2014). In-
 135 teraction themselves may impact co-distribution, as hypothesized long ago by Diamond
 136 Diamond (1975). Analyses of community turnover are usually performed with data rep-
 137 resented in a table with rows corresponding to sites (or measurements) and columns to
 138 species. Metrics of beta diversity quantify the variance of this community data (Legendre
 139 *et al.*, 2005). Traditional approaches rely on measures of dissimilarity among communi-
 140 ties, such as the Jaccard or Bray–Curtis indices. More recent approaches decompose total
 141 variation of the community data into species and site contributions to beta diversity (Leg-
 142 endre & De Cáceres, 2013), and further partition it into dissimilarity due to changes in
 143 species richness and dissimilarity due to actual species turnover (Baselga, 2010). T Even
 144 though all of these methods compare whole lists of species among sites or measurements,
 145 they remain fundamentally ”species-based”, since they report variation within columns.
 146 None of them explicitly considers variation of associations (i.e., of pairs or higher-order
 147 motifs – Stouffer *et al.* 2007). The only exceptions are the Joint Species Distribution
 148 Models (Warton *et al.*, 2015; Ovaskainen *et al.*), which further account for the covariance
 149 among species after representing their response to the environment.

150 Similarly, we are now getting a better understanding of interaction turnover. As men-
 151 tioned above, in the network approach to community structure, species and interactions
 152 are represented by nodes and links, respectively. Associations can also be represented
 153 by matrices in which entries represent the occurrence or intensity of interactions among
 154 species (rows and columns). Network complexity is then computed as the number of
 155 interactions (in the case of binary networks) or interaction diversity (in the case of quan-
 156 titative networks, Bersier *et al.* 2002). Variability in community structure consequently
 157 arises from the turnover of species composition, along with turnover of interactions among

158 pairs of species. The occurrence and intensity of interactions could vary because of the
159 environment, species abundance, and higher-order ecological interactions (Poisot *et al.*,
160 2015b). Variation in community composition can be independent of variation of ecological
161 interactions, suggesting that species and interaction distribution may well respond to
162 different drivers (Poisot *et al.*, 2012).

163 The "niche" is by far the dominant concept invoked to explain species distributions
164 and community assembly, from the local to the global scale. Following Hutchinson 1957,
165 the niche is viewed as the set of environmental conditions allowing a population to es-
166 tablish and persist (see also Holt 2009). In other words, the niche is the location in that
167 multidimensional space that allow a species to have a positive growth rate when rare
168 (Godsoe *et al.*, 2017). Community turnover arises as a result of successive replacement of
169 species along an environmental gradient, in agreement with the Gleasonian view of com-
170 munities (Gleason, 1926). The concept is straightforward to put into practice with species
171 distribution models, as it maps naturally on available distributional and environmental
172 data. Consequently, a vast array of statistical tools have been developed to implement it
173 (e.g. BIOMOD Thuiller 2003, MaxEnt Phillips *et al.* 2006). It is however much harder
174 to account for ecological interactions within this approach (Townsend *et al.*, 2011). As
175 such, these interactions are often viewed as externalities constraining or expanding the
176 range of environmental conditions required for a species to maintain a viable population
177 (Pulliam, 2000; Soberón, 2007).

178 Interestingly, the food web literature also has its own "niche model" to position a
179 species in a community (Williams & Martinez, 2000), generalized later to other types of
180 interaction networks (Eklöf *et al.*, 2013). The niche of a species in this context represents
181 the multidimensional space of all of its interactions. Each species is characterized by a
182 niche position, an optimum and a range over three to five different niche axes (Williams
183 & Martinez, 2000; Eklöf *et al.*, 2013). The niche model of food web structure and its

184 variants have successfully explained the complexity of a variety of networks, from food
185 webs to plant–pollinator systems (Allesina *et al.*, 2008; Williams *et al.*, 2010; Eklöf *et al.*,
186 2013). This conceptual framework is, however, limited to local communities, and does
187 not provide any explanation for the turnover of network structure along environmental
188 gradients.

189 **The integrated niche**

190 Despite several attempts to update the concept of the ecological niche, ecologists have not
191 moved far beyond the "n-dimensional hypervolume" defined by Hutchinson. Despite its
192 intuitive interpretation and easy translation into species distribution models (Boulangeat
193 *et al.*, 2012; Blonder *et al.*, 2014), the concept has been frequently criticized (Hardin, 1960;
194 Peters, 1991; Silvertown, 2004), and several attempts have been made to expand and
195 improve it (Pulliam, 2000; Chase & Leibold, 2003; Soberón, 2007; Holt, 2009; McNerny
196 & Etienne, 2012b).

197 Part of the problem surrounding the niche concept has been clarified with the dis-
198 tinction between Eltonian and Grinnellian definitions (Chase & Leibold, 2003). The
199 Grinnellian dimension of the niche is the set of abiotic environmental conditions required
200 for a species to maintain a population in a location. The Grinnellian niche is intuitive
201 to apply, and constitutes the conceptual backbone of species distribution models. The
202 Eltonian niche, on the other hand, is 'the place of a species in its biotic environment, its
203 relations to food and enemies'. While this aspect of the niche is well known by community
204 ecologists, it is trickier to turn into predictive models. Nonetheless, the development of
205 the niche model of food web structure (Williams & Martinez, 2000) and its parameteri-
206 zation using functional traits (Gravel *et al.*, 2013; Bartomeus *et al.*, 2016) made it more
207 operational.

208 These perspectives are rather orthogonal to each other, and this has resulted in con-
 209 siderable confusion in the literature (McInerny & Etienne, 2012a). Chase & Leibold 2003
 210 attempted to reconcile with the following definition: "*[The niche is] the joint description*
 211 *of the environmental conditions that allow a species to satisfy its minimum requirements so*
 212 *that the birth rate of a local population is equal to or greater than its death rate along with*
 213 *the set of per capita effects of that species on these environmental conditions*". Their rep-
 214 resentation merges zero-net-growth isoclines delimiting the Grinnellian niche ("when does
 215 the population persists?") with impact vectors delimiting the Eltonian niche ("what is the
 216 per-capita impact?"). While this representation has been very influential in local-scale
 217 community ecology (the resource-ratio theory of coexistence, Tilman 1982), it remains
 218 impractical at larger spatial scales because of the difficulties in measuring it. The absence
 219 of any mathematical representation of the niche that can be easily fit to ecological data
 220 may explain why biogeographers are still struggling to develop species distribution mod-
 221 els that also consider ecological interactions. Thus, a more integrative description of the
 222 niche will be key to understand spatial and temporal turnover in community structure.

223 We propose to integrate the two perspectives of the niche using a visual representa-
 224 tion of both components (Fig. 2). The underlying rationale is that, in addition to the
 225 environmental constraints on demographic performance (Fig.2 top panel), any organism
 226 requires resources to meet its metabolic demands and to sustain reproduction (Fig. 2
 227 nodes in network of bottom panel). Abiotic environmental axes are any non-consumable
 228 factors affecting the demographic performance of an organism. Alternatively, the resource
 229 axes are traits of the resources that allow interactions with the consumers. By defini-
 230 tion of what an interaction is, these axes also influence the growth rate. The niche can
 231 therefore be viewed as the set of abiotic environmental conditions (the Grinnellian com-
 232 ponent) along with the set of traits (the Eltonian component) that allow a population
 233 to establish and to grow at a location. This visual representation make operational the

theoretical interpretation of the niche and species ranges by Godsoe et al. Godsoe *et al.* 2017. Accordingly, each species can be characterized by an optimal position along both the environmental (x-axis) and the trait (y-axis) plane. The integrated niche is then the hypervolume where interactions can occur and where a population has a positive growth rate.

This approach changes the representation of the niche, putting species distributions and ecological interactions into the same formalism. Moreover, it allows the limits of the niche axes to be independent of each other (as in the example in Fig. 2), or to interact. Some of these axes may not be independent. For instance, the optimal prey size for predatory fishes could decline with increasing temperature (Gibert & DeLong, 2014), which would make diet boundaries functions of the environment. Alternatively, we could also consider that the growth rate of the predator changes with the size of its prey items, thereby altering the environmental boundaries and changing the shape of the niche in the middle of this visual representation. It is also important that this conceptual representation depicts a reality that could be highly dimensional and sometimes difficult to represent statistically (?).

A probabilistic representation of interaction networks in space

We now formalize the integrated niche with a probabilistic approach to interactions and distributions. In particular, we seek to represent the probability that an interaction between species i and j occurs at location y . We define L_{ijy} as a stochastic process taking a value of 1 when an interaction occurs and a value of 0 when it does not, and focus on the probability that this event occurs, $P(L_{ijy} = 1)$, over a given sampling interval and space. We note that the occurrence of an interaction is dependent on the co-occurrence of

species i and j . This argument might seem trivial at first, but the explicit consideration of this condition in the probabilistic representation of ecological interactions will prove instrumental to understanding their variation. We note this phenomenon we aim to describe differs fundamentally from the effect that interactions do have on co-occurrence (Cazelles *et al.*, 2016). We similarly define X_{iy} as a stochastic process representing the occurrence of species i at location y , taking a value of 1 when the species occurs and a value of 0 when it is absent. The quantity we seek to understand is the probability of a joint event, conditional on the set of environmental conditions E_y :

$$P(X_{iy}, X_{jy}, L_{ijy} | E_y) \quad (1)$$

Or simply said, the probability of observing both species i and j plus an interaction between i and j given the conditions E_y at location y . This probability could be decomposed into two parts using the product rule of probabilities:

$$P(X_{iy}, X_{jy}, L_{ijy} | E_y) = P(X_{iy}, X_{jy} | E_y) P(L_{ijy} | X_{iy}, X_{jy}, E_y) \quad (2)$$

The first term on the right-hand side of the equation is the probability of observing the two species co-occurring at location y . It corresponds to the Grinnellian dimension of the niche. The second term represents the probability that an interaction occurs between species i and j , given that they are co-occurring. This predicate can be refined using information on trait distribution and trait matching rules ((Bartomeus *et al.*, 2016)). Above, we referred to this entity as the "metaweb" and it corresponds to the Eltonian dimension of the niche. Below, we will see how this formalism can be directly fit to empirical data. But before turning to an application, we will discuss the interpretation of different variants of these two terms.

278 Variants of co-occurrence

279 There are several variants to the co-occurrence probability, representing different hy-
 280 potheses concerning spatial variation in network structure (see the explicit formulations
 281 in Table 1). The simplest model relates the probability of co-occurrence directly to
 282 the environment, $P(X_{iy}, X_{jy}|E_y)$. In this situation, there are no underlying assump-
 283 tions about the ecological processes responsible for co-occurrence. It could arise because
 284 interactions constrain distribution, where in such case the co- occurrence would be con-
 285 ditional on L , or, alternatively, because of environmental requirements shared between
 286 i and j (Pollock *et al.*, 2014; Cazelles *et al.*, 2016). In the former case, species are
 287 not independent of each other and the conditional occurrence must be accounted for
 288 explicitly, $P(X_{iy}, X_{jy}|E_y) = P(X_{iy}|E_y, X_{jy})P(X_{jy}|E_y)$. In the latter case, species are
 289 independent, and only the marginal occurrence must be accounted for, $P(X_{ijy}|E_y) =$
 290 $P(X_{iy}|E_y)P(X_{jy}|E_y)$.

291 The co-occurrence probability itself could depend on ecological interactions. This
 292 should be viewed as the realized component of the niche (i.e. the distribution when
 293 accounting for species interactions). Direct pairwise interactions such as competition,
 294 facilitation, and predation have long been studied for their impact on co-distribution
 295 (e.g. Diamond 1975; Connor & Simberloff 1979; Gotelli 2000). Second- and higher-order
 296 interactions (e.g. trophic cascades) could also affect co-occurrence (Staniczenko *et al.*).
 297 Co-occurrence of multiple species embedded in ecological networks is a topic of its own,
 298 however, and is influenced by both network topology and species richness (Cazelles *et al.*,
 299 2016). Not only direct interactions influence co-occurrence, but indirect interactions do as
 300 well (e.g. plant species sharing an herbivore, or herbivores sharing parasitoids, could repel
 301 each other in space Holt & Lawton 1993). The impact of direct interactions and first-order
 302 indirect interactions on co-occurrence tends to vanish with increasing species richness in
 303 the community (Cazelles *et al.*, 2016). Further, co-occurrence is also influenced by the

covariance of interacting species to an environmental gradient (Cazelles *et al.*, 2015).

The interpretation that interactions may impact co-occurrence, and that the realization of an interaction requires co-occurrence are not mutually exclusive. But since here the object is to describe the variation in the occurrence of interactions, which absolutely requires co-occurrence, we will give attention to this part of the problem. Relating co-occurrence to the structure of interaction networks is a problem on its own (and vice-versa). We will therefore focus here on the variation of interactions and not on their distribution, and leave this specific issue for the Perspectives section and future research.

Variants of the metaweb

There are also variants of the metaweb. First, most documented metawebs have thus far considered ecological interactions to be deterministic, rather than probabilistic (e.g. Havens 1992; Wood *et al.* 2015). Species are assumed to interact whenever they are found together in a location, independent of their local abundance and the local environment. In other words, $P(L_{ijy}|X_{ijy} = 1) = 1$ and $P(L_{ijy}|X_{ijy} = 0) = 0$. This approach might be a reasonable approximation if the spatial or temporal scale of sampling and inference is so large that the probability of observing at least one interaction converges to unity. In this scenario, network variation arises solely from species distributions.

Second, ecological interactions could also vary with the environment, so that $P(L_{ijy}|E_y)$. Although it is rare to see a conditional representation of pairwise ecological interactions, experimental studies have frequently revealed interactions to be sensitive to the environment. For instance, (McKinnon *et al.*, 2010) showed that predation risks of shorebirds vary at the continental scale, decreasing from the south to the north. It is also common to see increasing top-down control with temperature (e.g. Shurin *et al.* 2012; Gray *et al.* 2015). Effects of the environment on interactions (Gibert & DeLong, 2014) also propagate up the community and influence network structure (Tylianakis *et al.*, 2007; Woodward

329 *et al.*, 2010; Petchey *et al.*, 2010).

330 **Application: continental-scale variation of host-parasite** 331 **community structure**

332 We now turn to an illustration of our theory with the analysis of an empirical dataset of
333 host-parasite networks sampled throughout the south-north environmental gradient in
334 continental Europe (Kopelke *et al.*, 2017). Our objective here is to illustrate potential
335 applications of the approach outlined above, rather than a specific attempt at describing
336 the biogeography of this particular community. The focal system consists of local food
337 webs of willows (genus *Salix*), their galling insects, and the natural enemies (parasitoids
338 and inquilines) of these gallers. We ask: i) how much does network structure vary across
339 the gradient, and ii) what is the primary driver of network turnover across the gradient?

340 **Data**

341 Communities of willows, gallers, and parasitoids are species-rich and widely distributed,
342 with pronounced variation in community composition across space. The genus *Salix* in-
343 cludes over 400 species, most of which are shrubs or small trees (Argus, 1997), and is
344 common in moist habitats across the Northern Hemisphere (Skvortsov, 1999). Willows
345 support a highly diverse community of herbivorous insects, with one of the main herbivore
346 groups being gall-inducing sawflies (Hymenoptera: Tenthredinidae: Nematinae: Euurina
347 (Kopelke, 1999)). Gall formation is induced by sawfly females during oviposition, and
348 includes marked manipulation of host-plant chemistry by the galler (Nyman & Julkunen-
349 Tiitto, 2000). The enemy community of the gallers includes nearly 100 species belonging
350 to 17 insect families of four orders (Kopelke, 2003). These encompass two main types:
351 inquiline larvae (Coleoptera, Lepidoptera, Diptera, and Hymenoptera) feed primarily on

352 gall tissue, but typically kill the galler larva in the process, while parasitoid larvae (repre-
353 senting many families in Hymenoptera) kill the galler larvae by direct feeding (Kopelke,
354 2003).

355 Local realizations of the willow–galler–parasitoid network were reconstructed from
356 community samples collected between 1982 and 2010. During this period, willow galls
357 were collected at 370 sites across Central and Northern Europe. In total, 52, 96 and
358 127 *Salix*, galler and parasitoid and inquiline taxa were distinguished, respectively. The
359 strength of this dataset is that observations were observed in situ, rather than inferred
360 from expert knowledge or other sources of information, thereby allowing the analysis of
361 their spatial variation. The drawback is that, because many species are rare, some of the
362 pairs of species may have been observed co-occurring only a few times, or never. As a
363 consequence, despite the extent of the sampling, there is significant uncertainty in the
364 quantification of some of the links. The above described methodology explicitly aims at
365 describing this uncertainty, and research should be done to develop methods to reduce it
366 (see Conclusion).

367 The current study represents the first analysis of the full data set from a spatial
368 perspective. Full details about the data collection and species identification are provided
369 in (Kopelke *et al.*, 2017). Annual mean temperature and precipitation were obtained
370 from WorldClim using the R package *dismo* (Hijmans, 2015) and GPS coordinates of the
371 sampling locations. While other covariates could have also been considered to represent
372 the high dimensionality of the niche, these two variables are likely representative of the
373 most important axes of the European climate, and are also more easily interpretable than
374 reduced variables obtained, for example, by principal component analysis.

Analysis

Computing the probability of observing an interaction involves fitting a set of binomial models and collecting their estimated probabilities. For the sake of illustration, we considered generalized linear models – although more flexible fitting algorithms (e.g. GAM or Random Forest) could equally well be used, as long as the algorithm can estimate the probability for each observation. The data consist of a simple (albeit large and sparse) table with the observation of each species, X_{iy} and X_{jy} , their co-occurrence, X_{ijy} , the observation of an interaction L_{ijy} , and environmental co- variates E_y . Thus, there is one row per pair of species per site. We considered that an absence of a record of an interaction between co-occurring species at a site means a true absence (see below for a discussion on this issue).

We compared three models for the co-occurrence probability. The first one directly represents the co-occurrence probability conditional on the local environment, $P(X_{iy}, X_{jy}|E_y)$ (models are listed at Table 1 and 2). Hence, this model makes no assumption about the mechanisms driving co- occurrence for any given environment, and instead uses the information directly available in the data. It thereby indirectly accounts for the effect of interactions on co-occurrence, if there are any. The second model considers independent occurrence of species. In this case, we independently fit $P(X_{iy}|E_y)$ and $P(X_{jy}|E_y)$, and we then take their product to derive the probability of co-occurrence. This model should be viewed as a null hypothesis with respect to the first model, since a comparison between the respective models will reveal if there is significant spatial association of the two species beyond a joint response to the shared environment (Cazelles *et al.*, 2016). Finally, the third model assumes that the probability of co-occurrence is independent of the environment and thus constant throughout the landscape. In other words, $P(X_{iy}, X_{jy})$ is obtained by simply counting the number of observed co-occurrences divided by the total number of observations. Thus, the comparison between the first and third model allows

us to test the hypothesis that co-occurrence is conditional on the environment. Whenever the environment was included as a covariate in the GLM, we considered a second-order polynomial term for both temperature and precipitation in order to account for optimal environmental conditions. There are consequently five parameters for the first model when fitting a given pair of species, 10 parameters for the second, and only one for the third model.

Following the same logic, we compared three models of the interaction probability. The first model conditions the interaction probability on the local environmental variables, $P(L_{ijy}|X_{iy}, X_{jy}, E_y)$. Consequently, the model was fit to the subset of the data where the two species co-occur. The second model fits the interaction probability independently of the local environmental variables, $P(L_{ijy}|X_{iy}, X_{jy})$. It corresponds to the number of times the two species were observed to interact when co-occurring, divided by the number of times that they co-occurred. The third model is an extreme case performed only to test the hypothesis that if two species are found to interact at least once, then they should interact whenever they co- occur, $P(L_{ijy}|X_{iy}, X_{jy}) = 1$. While not necessarily realistic, this model tests an assumption commonly invoked in the representation of local networks from the knowledge of a deterministic metaweb. It represents potential interactions, rather than realized ones. There are consequently five parameters for the first model, a single parameter for the second model and no parameter to evaluate for the third model (where the interaction probability is fixed by the hypothesis).

We fitted the different models to each pair of species and recorded the predicted probabilities. The joint probability $P(L_{ijy}, X_{iy}, X_{jy})$ was then computed from Eq. 2, and the likelihood of each observation D_{ijy} was computed as $\mathcal{L}(\theta_{ijy}|D_{ijy}) = P(L_{ijy}, X_{iy}, X_{jy})$ if an interaction was observed, and as $\mathcal{L}(\theta_{ijy}|D_{ijy}) = 1 - P(L_{ijy}, X_{iy}, X_{jy})$ if no interaction was observed. The log-likelihood was summed over the entire dataset to compare the different models by AIC. We therefore evaluate the likelihood of all local networks, given

the model. Not surprisingly, there was a very large number of species pairs for which this model could not be computed, as they simply never co-occurred. For these pairs, we have no information of the interaction probability, and they were consequently removed from the analysis. The log-likelihood reported across the entire dataset was summed over all pairs of species observed to co- occur at least once. Interactions between the first (*Salix*) and second (gallers) trophic layers and those between the second and third (parasitoids) were considered separately. Finally, we used the full model (in which both co-occurrence and the interaction are conditional on the environment) to interpolate species distributions and interaction probabilities across the entire European continent. We reconstructed the expected network for each location in a 1 X 1 km grid and computed the probabilistic connectance following (Poisot *et al.*, 2016a).

All of the data are openly available in the database *mangal* (Poisot *et al.*, 2015a) and all R scripts for running the analysis, are provided in the github repository https://github.com/DominiqueGravel/ms_probaweb.

Results

Despite the extensive sampling, many pairs of species were observed to co- occur only a few times. This made it difficult to evaluate interaction probabilities with any reasonable confidence interval. Thus, we start with an example of a single pair of species selected because of its high number of co- occurrences ($N_{ij} = 38$): the leaf folder *Phyllocolpa prussica* and the parasitoid *Chrysocharis elongata*. These two fairly abundant species were observed $N_i = 49$ and $N_j = 121$ times, respectively, across the 370 sites, and they were found to interact with a marginal probability $P(L_{ij}) = 0.55$, which means they interacted at 21 different locations. Here, a comparison of model fit (Table 1) reveals that conditioning the interaction probability on local environmental conditions adds no explanatory power beyond a model assuming the same probability of interaction

anywhere in space (Model 1 vs Model 2). Moreover, when the two species co-occur, the occurrence of the interaction was insensitive to the environment (Model 2 vs Model 3). Alternatively, climatic variables significantly impacted co-occurrence (Model 3 vs Model 4). The neutral model performed worse than the non-random co-occurrence model (Model 3 vs Model 6). The full model revealed that the greatest interaction probability occurred at intermediate temperature and precipitation, simply because this is where the two species most frequently co-occur (Fig. 3). The probabilities of co-occurrence and interaction can be represented in space, where we found that the highest interaction probability occurred in Central Europe (Fig. 4).

We evaluated each model for all pairs of species in order to better understand the large-scale drivers of network turnover. The results were highly consistent among trophic layers (*Salix*-gallers and gallers-parasitoids; Table 2), despite very significant variability in the fit of the models to the different pairs of species. This variability was particularly important for the models of interactions, given co-occurrence. Across all pairs of species, the conditional representation of interactions performed better than the marginal one (Model 1 vs Model 2); that is, interactions did not occur systematically whenever the two species were found co-occurring. Hence, in addition to species turnover, the stochastic nature of interactions contributes to network variability. In total, we recorded 1,173 pairs of interactions, only 290 of which occurred more than five times. Out of these 290 interactions, 143 were systematically detected whenever the two species co-occurred. In the instances when species co-occurred, the two environmental variables considered proved relatively poor predictors of their interactions (Model 2 vs Model 3). Not surprisingly, for both types of interactions (*Salix*-galler and galler-parasitoid), the log-likelihood increased when the environment was considered. However, the extra number of parameters exceeded the gain in log-likelihood and inflated AIC. Therefore, the most parsimonious model excluded the effect of the environment. On the basis of log-likelihood only, co-occurrence

was non-neutral for both *Salix*–galler and galler–parasitoid interactions. Thus, according to AIC, the best model was the one of non-random co-occurrence (Model 3 vs Model 6) for both types of interactions.

Once we had selected the best model based on AIC (Model 3, Table 2), we used it to reconstruct the expected species richness, along with the most likely network for each location. Using this approach, we mapped the expected distribution of network properties across Europe (Fig. 6). For simplicity, we chose to consider connectance as our descriptor of network configuration, as this metric can be easily computed from probabilistic networks (Poisot *et al.*, 2016a) and is also a good proxy for many other network properties (Poisot & Gravel, 2014). Overall, we found a peak in *Salix*, gallers and parasitoid diversity in Northern Europe. The expected number of interactions roughly followed the distribution of species richness, but accumulated at a rate different from species numbers. Connectance likewise peaked in Northern Europe (Fig. 6).

Interpretation

We have proposed that the representation of community structure and its variation in space is best captured by the formalism of ecological networks, as both the distribution of species and their interspecific interactions can then be accounted for. We consequently revised the niche concept in order to integrate its abiotic and biotic components. This integrated niche was represented visually with an ordination of species into an environmental space and a trait space. The fundamental niche of a species is represented as the set of environmental conditions and resources that allow a species to establish in a location, thereby integrating the Eltonian and the Grinnellian components of the niche. We then translated the concept mathematically by investigating the probability of the joint occurrences of species and their interaction, which should be interpreted as an in-

502 teraction distribution model. We used this approach to characterize the turnover of the
503 structure of ecological interactions in a species-rich tri-trophic network across Western
504 Europe, finding that the primary driver of network variation is the turnover in species
505 composition.

506 Applying the framework to our large data set on host–parasite interactions across
507 Europe revealed key features in the interaction between *Salix* taxa, their herbivores,
508 and the natural enemies of these herbivores. Consistent with a general increase in the
509 diversity of *Salix* towards boreal areas (Cronk *et al.*, 2015), overall species richness of
510 the networks increased towards the north. The distribution of *Salix* species richness
511 largely matched those of gallers and parasitoids. These observations within Europe are
512 also matched by the ones found at a global scale for *Salix* (Argus, 1997; Cronk *et al.*,
513 2015; Wu *et al.*, 2015) and sawflies (Kouki *et al.*, 1994; Kouki, 1999). Species richness
514 in a common group of parasitic wasps, the Ichneumonidae, was originally presumed to
515 show a similar "reversed latitudinal gradient", but this observation has been recently
516 challenged by findings of rather high ichneumonid diversity in the tropics (Veijalainen
517 *et al.*, 2013). Nevertheless, ichneumonid subfamilies specifically associated with sawflies
518 (Ctenopelmatinae, Tryphoninae) are clearly less diverse in the south.

519 Exactly what processes are responsible for the distribution of species richness at dif-
520 ferent trophic levels is yet to be established (but see e.g. Roininen *et al.* 2005; Nyman
521 *et al.* 2010; Leppänen *et al.* 2014), but as a net outcome of different latitudinal trends
522 across trophic levels, the distribution of co- occurrence and therefore of potential in-
523 teractions differed between the first and second layers of feeding links. The correlation
524 between expected *Salix* and gallers richness was 0.73, while it was 0.58 between gallers
525 and their parasitoids. Therefore, the ratio of herbivores to *Salix* species is essentially
526 constant across Europe, whereas each herbivore species is potentially attacked by a richer
527 enemy community higher latitudes. Consequently, overall connectance peaks in Northern

Europe (Fig. 6).

In terms of species interacting with each other, our analysis suggests that the environment leaves a detectable imprint on species co-occurrence, but only a slight mark on the occurrence of realized links among species in a specific place: the probability of finding a given combination of species at a higher and a lower trophic level at the same site was clearly affected by the environment, whereas the probability of observing an interaction between the two was, given co-occurrence, not detectably so. The interactions were highly uncertain, as only few pairs of species were systematically interacting when co-occurring but most not. This applies to the example species *Phyllocolpa prussica* and *Chrysocharis elongata* (Figs 2 and 3), but also to all species pairs more generally. For the example species pair, the full model revealed that the joint probability event (interaction and co-occurrence) peaks at intermediate temperature and precipitation, simply because this is where the two species co-occur most often. This does not imply that species will always interact when they meet – although this is a basic assumption in most documented metawebs to date (e.g. Havens 1992; Wood *et al.* 2015). Rather, an interaction is better represented as a stochastic process whose probability is also influenced by the probability with which species co-occur. What we cannot reliably know is how this stochasticity splits into different sources of uncertainty – i.e., the extent to which a species at the higher trophic level runs into a species at the lower level co-occurring at the site, the extent to which this interaction is detected by an observer collecting a finite sample, or simply the uncertainty arising from incomplete description of a highly dimensional niche. Future work will be required to document the relative importance of these sources of uncertainty in the occurrence of interactions.

Perspectives

Evidence that the structure of ecological networks does vary across habitats (e.g. Tylianakis *et al.* 2007), over environmental gradients (Lurgi *et al.*, 2012) and in time (Trøjelsgaard *et al.*, 2015) is accumulating rapidly. It is not clear, however, to what extent the turnover of network structure is driven by a systematic change in species composition or of pair-wise interactions (Poisot *et al.*, 2012, 2015b). Our model comparison of host-parasite interactions revealed that most of the turnover is driven by species-specific responses to the environment, impacting species richness, and that co-occurrence was mostly neutral. Further, the occurrence of interactions among host and parasite is highly stochastic even when both are present, and not predictable by the variables considered by us. We know that interactions vary with the environment in other systems, for instance, herbivory (Shurin *et al.*, 2012; ?) and predation (McKinnon *et al.*, 2010; Legagneux *et al.*, 2014) are often found to increase with temperature, resulting in spatial variation of trophic cascades (Gray *et al.*, 2015). What remains unclear, however, is the extent to which such variation is driven by a turnover of species composition along gradients, or a turnover of the interactions. Here we found that interactions vary substantially but non-predictably along the annual temperature and the precipitation gradient. The lack of detectable signal may be due to our choice of covariates. Indeed, a previous study on a similar system identified habitat characteristics as the primary drivers of interactions (Nyman *et al.*, 2015). New investigations with other systems will thus be required to challenge this result. Under all circumstances, documenting the relationship between the environment and the occurrence of interactions at continental scales is critical for understanding how large-scale variation of trophic regulation influences community dynamics and ecosystem functioning (Harfoot *et al.*, 2014).

We restricted our analysis to the effect of co-occurrence on ecological interactions, neglecting the inverse of the problem. We did not investigate in depth the drivers of

577 co-occurrence and simply took it for granted from the data. Co-occurrence was indeed
 578 different from the expectation of independent species distributions. It thus begs the
 579 question of whether, once environmental effects on species-specific distribution have been
 580 accounted for, interactions come with significant effects on co-occurrence? We could
 581 rephrase this problem by asking whether the fundamental niche differs from the real-
 582 ized niche, and how this applies to our framework. For example, we have considered
 583 above simply the co-occurrence probability, $P(X_{iy}, X_{jy}|E_y)$, which could be expanded as
 584 $P(X_{iy}|X_{jy}, E_y)P(X_{jy}|E_y)$. After some re-arrangement of Eq. 2, the marginal occurrence
 585 probability, $P(X_{jy}|E_y)$, could be considered as a species distribution model taking into
 586 account the interaction between these species. This derivation would however critically
 587 depend on a strong *a priori* expectation of the conditional probability of observing a
 588 species given the distribution of the other species. This assumption seems reasonable
 589 for some situations, such as a parasitoid species that requires a host to develop. On the
 590 other hand, Cazelles 2016 found that the strength of this association is often rather weak
 591 if not neutral (for instance, with the example pair analysed at Table 1). The lack of
 592 an association could simply arise when the parasitoid is generalist enough that it is not
 593 obligated to track the distribution of any single/given host (Cazelles *et al.*, 2015).

594 At present, there is only indirect support for the hypothesis that interacting species
 595 are conditionally distributed but this possibility should be the topic of theoretical inves-
 596 tigation with dynamical metacommunity models (Cazelles *et al.*, 2015) and empirical hy-
 597 pothesis testing. The impact of ecological interactions on the distribution of co-occurrence
 598 has been the topic of many publications since Diamond 1975 seminal study on competi-
 599 tion and "checkerboard" distribution, but pairwise approaches have only recently received
 600 attention (Veech, 2013). Whether two interacting species are more closely associated in
 601 space remains unclear, since most approaches based on null models consider community-
 602 level metrics (e.g. Gotelli 2000), such as the C-score, thereby making it hard to evaluate

603 if specific interactions do indeed affect co-occurrence. The expansion of the framework
604 we describe to account for the difference between the realized and the fundamental niche
605 will therefore require further investigation of the impact of interactions on co-occurrence.

606 Ecological networks are known to be extremely sparse, *i.e.* they have far more absences
607 than presences of interactions. Absences of interactions, however, can come from different
608 sources. The fact that unequal sampling at the local scale can affect our understanding
609 of network structure is well documented (Martinez *et al.*, 1999). In a spatial context,
610 however, some interactions may be undocumented simply because the species involved
611 have never been observed to co-occur. Although these cases are reported as a lack of
612 interactions, in actuality we cannot make any reliable inference from them: since the
613 species have never been observed together, it remains possible that they would interact
614 if they did. A fundamentally different category of absences of interactions are then those
615 reported after multiple observations of species co-occurrence. Thus, to gain confidence
616 that the probability of an interaction is low, extensive sampling (that is, several records
617 of co- occurrence) is needed. Generally, our confidence that the interaction is indeed
618 impossible will increase with the number of observations of the species pair. Seeing that
619 this is essentially a Bernoulli process (the probability that the species will interact given
620 their presence), the breadth of the confidence interval is expected to saturate after a fixed
621 number of observations, which can be set as a threshold above which a species pair has
622 finally been observed "often enough". This will allow us to deal with both confirmed
623 absences of interactions and mere absence of evidence.

624 Conclusion

625 Our representation of spatial variation of community structure offers a new approach for
626 the study of the biogeography of ecological networks. We see the following key challenges

and opportunities ahead in this exciting area of research:

1. New generation of network data. Investigating spatial variation of network structure will require high quality and highly replicated network data. We have investigated one the most comprehensive spatial network datasets we are aware of and nonetheless found immense gaps of knowledge in its resolution. Species richness accumulates much faster than observations of ecological interactions (Poisot *et al.*, 2012). Each pair of species must be observed several times in order to obtain reliable estimates of their interaction probability.

2. Estimation of the reliability of interactions. We need quantitative tools to estimate the confidence intervals around inferred interaction probabilities, as well as estimators of the frequency of false absences. Bayesian methods are promising to that end because we could use information on the target species (e.g. if they are known as specialists or generalists) to provide prior estimates of the interaction probability.

3. From interaction probabilities to a distribution of network properties. Metrics are available to analyse the structure of probabilistic networks (Poisot *et al.*, 2016a). These metrics are useful as first approximation, but they assume independence among interactions. This might not be the case in nature because of the role of co-occurrence and shared environmental requirements. We also need to better understand the distribution of network properties arising from probabilistic interactions.

4. Investigation of the environmental-dependence of ecological interactions. There is evidence that interactions can vary in space, but this problem has not been investigated in a systematic fashion. The paucity of currently available data precludes an extensive analysis of this question at present.

5. Effects of ecological interactions on co-occurrence. We have intentionally omitted the feedback of ecological interactions on co-occurrence in this framework. As abundance can impact the occurrence of interactions, and conversely since interactions

653 impact abundance (Canard *et al.*, 2014), we could reasonably expect that interactions
654 will also influence co-occurrence. Theory in this regard does exist for simple three-species
655 modules (Cazelles *et al.*, 2015), but its extension to entire co-occurrence networks will
656 prove critical in the future, especially given the interest in using co- occurrence to infer
657 ecological interactions (Morales-Castilla *et al.*, 2015; Morueta-Holme *et al.*, 2016).

658 **Acknowledgements**

659 This is a contribution to the working groups *Continental-scale variation of ecological net-*
660 *works* supported by the Canadian Institute for Ecology and Evolution), and the *Next*
661 *Generation Data, Models, and Theory Working Group*, supported by the Santa Fe Insti-
662 tute, the Betsy and Jesse Fink Foundation, the ASU-SFI Center for Biosocial Complex
663 Systems, and NSF Grant PHY-1240192. DG also acknowledges financial support from
664 NSERC- Discovery grant program and Canada Research Chair program.

References

- Albouy, C., Velez, L., Coll, M., Colloca, F., Le Loc'h, F., Mouillot, D. & Gravel, D. (2014) From projected species distribution to food-web structure under climate change. *Global Change Biology* **20**, 730–741.
- Allesina, S., Alonso, D. & Pascual, M. (2008) A general model for food web structure. *Science* **320**, 658–61.
- Anderson, M.J., Crist, T.O., Chase, J.M., Vellend, M., Inouye, B.D., Freestone, A.L., Sanders, N.J., Cornell, H.V., Comita, L.S., Davies, K.F., Harrison, S.P., Kraft, N.J.B., Stegen, J.C. & Swenson, N.G. (2011) Navigating the multiple meanings of beta-diversity: a roadmap for the practicing ecologist. *Ecology Letters* **14**, 19–28.
- Argus, G. (1997) Infrageneric classification of *Salix* (Salicaceae) in the New World. *Systematic Botany Monographs* **52**, 1–121.
- Bartomeus, I., Gravel, D., Tylianakis, J., Aizen, M., Dickie, I. & Bernard-Verdier, M. (2016) A common framework for identifying linkage rules across different types of interactions. *Functional Ecology* pp. n/a–n/a.
- Bascompte, J. & Jordano, P. (2007) Plant-Animal Mutualistic Networks: The Architecture of Biodiversity. *Annual Review of Ecology, Evolution, and Systematics* **38**, 567–593.
- Baselga, A. (2010) Partitioning the turnover and nestedness components of beta diversity. *Global Ecology and Biogeography* **19**, 134–143.
- Bersier, L.F., Banašek-Richter, C. & Cattin, M.F. (2002) Quantitative descriptors of food-web matrices. *Ecology* **83**, 2394–2407.

- Blonder, B., Lamanna, C., Violle, C. & Enquist, B.J. (2014) The n-dimensional hyper-volume. *Global Ecology and Biogeography* **23**, 595–609.
- Boulangeat, I., Gravel, D. & Thuiller, W. (2012) Accounting for dispersal and biotic interactions to disentangle the drivers of species distributions and their abundances. *Ecology Letters* **15**, 584–593.
- Canard, E., Mouquet, N., Mouillot, D., Stanko, M., Miklisova, D. & Gravel, D. (2014) Empirical evaluation of neutral interactions in host-parasite networks. *The American naturalist* **183**, 468–79.
- Cazelles, K. (2016) *Influence des interactions biotiques sur la répartition géographique des espèces*. PhD. Dissertation. Université du Québec à Rimouski, Rimouski.
- Cazelles, K., Araújo, M.B., Mouquet, N. & Gravel, D. (2016) A theory for species co-occurrence in interaction networks. *Theoretical Ecology* **9**, 39–48.
- Cazelles, K., Mouquet, N., Mouillot, D. & Gravel, D. (2015) On the integration of biotic interaction and environmental constraints at the biogeographical scale. *Ecography* **39**, 921–931.
- Chase, J. & Leibold, M. (2003) *Ecological niches*. Chicago University Press, Chicago.
- Chase, J.M. (2007) Drought mediates the importance of stochastic community assembly. *Proceedings of the National Academy of Sciences of the United States of America* **104**, 17430–4.
- Connor, E. & Simberloff, D. (1979) The assembly of species communities: chance or competition? *Ecology* **60**, 1132–1140.

- 708 Cronk, Q., Ruzzier, E., Belyaeva, I. & Percy, D. (2015) Salix transect of Europe: lati-
709 tudinal patterns in willow diversity from Greece to arctic Norway. *Biodiversity Data*
710 *Journal* **3**.
- 711 Diamond, J. (1975) Assembly of species communities. *Ecology and evolution of communi-*
712 *ties* (eds. M. Cody & J. Diamond), pp. 342–444, Harvard University Press, Cambridge.
- 713 Dunne, J.A. (2006) The network structure of food webs. *Ecological networks: Linking*
714 *structure and dynamics* (eds. M. Pascual & J.A. Dunne), pp. 27–86, Oxford University
715 Press, Oxford.
- 716 Eklöf, A., Jacob, U., Kopp, J., Bosch, J., Castro-Urgal, R., Chacoff, N.P., Dalsgaard, B.,
717 de Sassi, C., Galetti, M., Guimarães, P.R., Lomáscolo, S.B., Martín González, A.M.,
718 Pizo, M.A., Rader, R., Rodrigo, A., Tylianakis, J.M., Vázquez, D.P. & Allesina, S.
719 (2013) The dimensionality of ecological networks. *Ecology letters* **16**, 577–583.
- 720 Gibert, J.P. & DeLong, J.P. (2014) Temperature alters food web body-size structure.
721 *Biology Letters* **10**.
- 722 Gleason, H. (1926) The individualistic concept of the plant association. *Bulletin of the*
723 *Torrey Botanical Club* **53**, 7–26.
- 724 Godsoe, W., Jankowski, J., Holt, R.D. & Gravel, D. (2017) Integrating biogeography with
725 contemporary niche theory. *Trends in Ecology and Evolution* **32**, 488–499.
- 726 Gotelli, N.J. (2000) Null Model Analysis of Species Co-Occurrence Patterns. *Ecology* **81**,
727 2606.
- 728 Götzenberger, L., de Bello, F., Bråthen, K.A., Davison, J., Dubuis, A., Guisan, A., Lepš,
729 J., Lindborg, R., Moora, M., Pärtel, M., Pellissier, L., Pottier, J., Vittoz, P., Zobel,
730 K. & Zobel, M. (2012) Ecological assembly rules in plant communities—approaches,

731 patterns and prospects. *Biological reviews of the Cambridge Philosophical Society* **87**,
732 111–27.

733 Gravel, D., Massol, F., Canard, E., Mouillot, D. & Mouquet, N. (2011) Trophic theory
734 of island biogeography. *Ecology letters* **14**, 1010–6.

735 Gravel, D., Poisot, T., Albouy, C., Velez, L. & Mouillot, D. (2013) Inferring food web
736 structure from predator-prey body size relationships. *Methods in Ecology and Evolution*
737 **4**, 1083–1090.

738 Gray, S.M., Poisot, T., Harvey, E., Mouquet, N., Miller, T.E. & Gravel, D. (2015) Temper-
739 ature and trophic structure are driving microbial productivity along a biogeographical
740 gradient. *Ecography* pp. n/a–n/a.

741 Hanski, I. (1999) *Metapopulation Ecology*. Oxford University Press, Oxford.

742 Hardin, G. (1960) The competitive exclusion principle. *Science* **131**, 1292–1297.

743 Harfoot, M.B.J., Newbold, T., Tittensor, D.P., Emmott, S., Hutton, J., Lyutsarev, V.,
744 Smith, M.J., Scharlemann, J.P.W. & Purves, D.W. (2014) Emergent global patterns of
745 ecosystem structure and function from a mechanistic general ecosystem model. *PLoS*
746 *Biol* **12**, 1–24.

747 Harris, D.J. (2016) Inferring species interactions from cooccurrence data with markov
748 networks. *Ecology* **97**, 3308–3314.

749 Havens, K. (1992) Scale and structure in natural food webs. *Science* **257**, 1107–9.

750 Hijmans, R.J. (2015) *raster: Geographic Data Analysis and Modeling*. R package version
751 2.5-2.

752 Holt, R.D. (2009) Bringing the Hutchinsonian niche into the 21st century : Ecological
753 and evolutionary perspectives. *Proceedings of the National Academy of Sciences* **106**,
754 19659–19665.

755 Holt, R.D. & Lawton, J.H. (1993) Apparent competition and enemy-free space in insect
756 host-parasitoid communities. *The American Naturalist* **142**, 623–645, PMID: 19425963.

757 Hutchinson, G. (1957) Concluding remarks. *Cold Spring Harbour Symposium* **22**, 415–
758 427.

759 Ings, T.C., Montoya, J.M., Bascompte, J., Blüthgen, N., Brown, L., Dormann, C.F.,
760 Edwards, F., Figueroa, D., Jacob, U., Jones, J.I., Lauridsen, R.B., Ledger, M.E.,
761 Lewis, H.M., Olesen, J.M., van Veen, F.J.F., Warren, P.H. & Woodward, G. (2009)
762 Ecological networks–beyond food webs. *The Journal of animal ecology* **78**, 253–69.

763 Kéfi, S., Berlow, E.L., Wieters, E.a., Navarrete, S.a., Petchey, O.L., Wood, S.a., Boit, A.,
764 Joppa, L.N., Lafferty, K.D., Williams, R.J., Martinez, N.D., Menge, B.a., Blanchette,
765 C.a., Iles, A.C. & Brose, U. (2012) More than a meal... integrating non-feeding inter-
766 actions into food webs. *Ecology letters* pp. 291–300.

767 Kissling, W.D., Dormann, C.F., Groeneveld, J., Hickler, T., Kühn, I., McNerny, G.J.,
768 Montoya, J.M., Römermann, C., Schiffers, K., Schurr, F.M., Singer, A., Svenning,
769 J.C., Zimmermann, N.E. & O'Hara, R.B. (2012) Towards novel approaches to mod-
770 elling biotic interactions in multispecies assemblages at large spatial extents. *Journal*
771 *of Biogeography* **39**, 2163–2178.

772 Kopelke, J., Nyman, T., Cazelles, K., Gravel, D., Vissault, S. & Roslin, T. (2017) Foodweb
773 structure of willowgalling sawflies and their natural enemies across europe. *Ecology* **98**,
774 1730–1730.

- 775 Kopelke, J.P. (1999) Gallenerzeugende Blattwespen Europas - Taxonomische Grundlagen,
776 Biologie und Ökologie (Tenthredinidae: Nematinae: Euura, Phyllocolpa, Pontania).
777 *Courier Forschungsinstitut Senckenberg* **212**, 1–183.
- 778 Kopelke, J.P. (2003) Natural enemies of gall-forming sawflies on willows (*Salix* spp.).
779 *Entomologia Generalis* **26**, 277–312.
- 780 Kouki, J. (1999) Latitudinal gradients in species richness in northern areas: some excep-
781 tional patterns. *Ecological Bulletins* **47**, 30–37.
- 782 Kouki, J., Niemelä, P. & Viitasaari, M. (1994)) Reversed latitudinal gradient in species
783 richness of sawflies (Hymenoptera, Symphyta). *Annales Zoologici Fennici* **31**, 83–88.
- 784 Krebs, C. (2009) *Ecology: The Experimental Analysis of Distribution and Abundance. 6th*
785 *ed.* Pearson, San Francisco.
- 786 Laliberté, E. & Tylianakis, J.M. (2010) Deforestation homogenizes tropical parasitoid-
787 host networks. *Ecology* **91**, 1740–1747.
- 788 Legagneux, P., Gauthier, G., Lecomte, N., Schmidt, N.M., Reid, D., Cadieux, M.C.,
789 Berteaux, D., Bêty, J., Krebs, C.J., Ims, R.a., Yoccoz, N.G., Morrison, R.I.G., Leroux,
790 S.J., Loreau, M. & Gravel, D. (2014) Arctic ecosystem structure and functioning shaped
791 by climate and herbivore body size. *Nature Climate Change* **E2168**, 1–5.
- 792 Legendre, P., Borcard, D. & Peres-Neto, P.R. (2005) Analyzing beta diversity: partition-
793 ing the spatial variation of composition data. *Ecological Monographs* **75**, 435–450.
- 794 Legendre, P. & De Cáceres, M. (2013) Beta diversity as the variance of community data:
795 dissimilarity coefficients and partitioning. *Ecology Letters* **16**, 951–963.
- 796 Leibold, M., Holyoak, M., Mouquet, N., Amarasekare, P., Chase, J.M., Hoopes, M.,
797 Holt, R.D., Shurin, J.B., Law, R., Tilman, D., Loreau, M. & Gonzalez, A. (2004) The

798 metacommunity concept : a framework for multi-scale community ecology. *Ecology*
799 *Letters* **7**, 601–613.

800 Leppänen, S., Malm, T., Värri, K. & Nyman, T. (2014) A comparative analysis of genetic
801 differentiation across six shared willow host species in leaf- and bud-galling sawflies.
802 *PLoS ONE* **9**, e116286.

803 Lurgi, M., López, B.C. & Montoya, J.M. (2012) Novel communities from climate change.
804 *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*
805 **367**, 2913–22.

806 Martinez, N.D., Hawkins, B.a., Dawah, H.A. & Feifarek, B.P. (1999) Effects of Sampling
807 Effort on Characterization of Food-Web Structure. *Ecology* **80**, 1044–1055.

808 McNerny, G.J. & Etienne, R.S. (2012a) Ditch the niche - is the niche a useful concept in
809 ecology or species distribution modelling? *Journal of Biogeography* **39**, 2096–2102.

810 McNerny, G.J. & Etienne, R.S. (2012b) Stitch the niche - a practical philosophy and
811 visual schematic for the niche concept. *Journal of Biogeography* **39**, 2103–2111.

812 McKinnon, L., Smith, P.A., Nol, E., Martin, J.L., Doyle, F.I., Abraham, K.F., Gilchrist,
813 H.G., Morrison, R.I.G. & Bêty, J. (2010) Lower predation risk for migratory birds at
814 high latitudes. *Science* **327**, 326–327.

815 Morales-Castilla, I., Matias, M.G., Gravel, D. & Araújo, M.B. (2015) Inferring biotic
816 interactions from proxies. *Trends in Ecology & Evolution* **30**, 347–356.

817 Morueta-Holme, N., Blonder, B., Sandel, B., McGill, B.J., Peet, R.K., Ott, J.E., Violle,
818 C., Enquist, B.J., Jørgensen, P.M. & Svenning, J.C. (2016) A network approach for
819 inferring species associations from co-occurrence data. *Ecography* pp. n/a–n/a.

- 820 Nyman, T. & Julkunen-Tiitto, R. (2000) Manipulation of the phenolic chemistry of wil-
 821 lows by gall-inducing sawflies. *Proceedings of the National Academy of Sciences* **97**,
 822 13184–13187.
- 823 Nyman, T., Leppänen, S.A., Várkonyi, G., Shaw, M.R., Koivisto, R., Barstad, T.E.,
 824 Vikberg, V. & Roininen, H. (2015) Determinants of parasitoid communities of willow-
 825 galling sawflies: habitat overrides physiology, host plant and space. *Molecular Ecology*
 826 **24**, 5059–5074.
- 827 Nyman, T., Vikberg, V., Smith, D. & BoevéJ-L (2010) How common is ecological speci-
 828 ation in plant-feeding insects? A “Higher” Nematinae perspective. *BMC Evolutionary*
 829 *Biology* **10**, e266.
- 830 Ovaskainen, O., Tikhonov, G., Norberg, A., Blanchet, F.G., Duan, L., Dunson, D., Roslin,
 831 T., Abrego, N. & Chave, J. (????) How to make more out of community data? a
 832 conceptual framework and its implementation as models and software. *Ecology Letters*
 833 **20**, 561–576.
- 834 Pellissier, L., Rohr, R., Ndiribe, C., Pradervand, J.N., Salamin, N., Guisan, A. & Wisz, M.
 835 (2013) Combining food web and species distribution models for improved community
 836 projections. *Ecology and Evolution* **3**, 4572–4583.
- 837 Petchey, O.L., Brose, U. & Rall, B.C. (2010) Predicting the effects of temperature on food
 838 web connectance. *Philosophical transactions of the Royal Society of London. Series B*,
 839 *Biological sciences* **365**, 2081–91.
- 840 Peters, R. (1991) *A critique for ecology*. Cambridge University Press, Cambridge.
- 841 Phillips, S.J., Anderson, R.P. & Schapire, R.E. (2006) Maximum entropy modeling of
 842 species geographic distributions. *Ecological Modelling* **190**, 231 – 259.

843 Pillai, P., Gonzalez, A. & Loreau, M. (2011) Metacommunity theory explains the emer-
844 gence of food web complexity. *Proceedings of the National Academy of Sciences of the*
845 *United States of America* **108**, 19293–19298.

846 Poisot, T., Baiser, B., Dunne, J.A., Kéfi, S., Massol, F., Mouquet, N., Romanuk, T.N.,
847 Stouffer, D.B., Wood, S.A. & Gravel, D. (2015a) mangal – making ecological network
848 analysis simple. *Ecography* **37**.

849 Poisot, T., Canard, E., Mouillot, D., Mouquet, N. & Gravel, D. (2012) The dissimilarity
850 of species interaction networks. *Ecology letters* **XX**, XX–XX.

851 Poisot, T., Cirtwill, A.R., Cazelles, K., Gravel, D., Fortin, M.J. & Stouffer, D.B. (2016a)
852 The structure of probabilistic networks. *Methods in Ecology and Evolution* **7**, 303–312.

853 Poisot, T. & Gravel, D. (2014) When is an ecological network complex? connectance
854 drives degree distribution and emerging network properties. *PeerJ* **2**, e251.

855 Poisot, T., Gravel, D., Leroux, S., Wood, S.A., Fortin, M.J., Baiser, B., Cirtwill, A.R.,
856 Araújo, M.B. & Stouffer, D.B. (2016b) Synthetic datasets and community tools for the
857 rapid testing of ecological hypotheses. *Ecography* **39**, 402–408.

858 Poisot, T., Stouffer, D.B. & Gravel, D. (2015b) Beyond species: why ecological interac-
859 tions vary through space and time. *Oikos* **124**, 243–251.

860 Pollock, L.J., Tingley, R., Morris, W.K., Golding, N., O’Hara, R.B., Parris, K.M., Vesk,
861 P.A. & McCarthy, M.A. (2014) Understanding co-occurrence by modelling species si-
862 multaneously with a joint species distribution model (jsdm). *Methods in Ecology and*
863 *Evolution* **5**, 397–406.

864 Pulliam, H.R. (2000) On the relationship between niche and distribution. *Ecology Letters*
865 **3**, 349–361.

- 866 Roininen, H., Nyman, T. & Zinovjev, A. (2005) Biology, ecology, and evolution of gall-
867 inducing sawflies (Hymenoptera: Tenthredinidae and Xyelidae). *Biology, Ecology, and*
868 *Evolution of Gall-Inducing Arthropods* (eds. A. Raman, C.W. Schaefer & T.M. With-
869 ers), pp. 467–494, Science Publishers, Inc.
- 870 Scheiner, S.M. & Willig, M.R. (2007) A general theory of ecology. *Vital And Health*
871 *Statistics. Series 20 Data From The National Vitalstatistics System Vital Health Stat*
872 *20 Data Natl Vital Sta .*
- 873 Schleuning, M., Fründ, J., Klein, A.M., Abrahamczyk, S., Alarcón, R., Albrecht, M., An-
874 dersson, G., Bazarian, S., Böhning-Gaese, K., Bommarco, R., Dalsgaard, B., Dehling,
875 D., Gotlieb, A., Hagen, M., Hickler, T., Holzschuh, A., Kaiser-Bunbury, C., Kreft,
876 H., Morris, R., Sandel, B., Sutherland, W., Svenning, J.C., Tschardtke, T., Watts,
877 S., Weiner, C., Werner, M., Williams, N., Winqvist, C., Dormann, C. & Blüthgen,
878 N. (2012) Specialization of mutualistic interaction networks decreases toward tropical
879 latitudes. *Current Biology* **22**, 1925 – 1931.
- 880 Shurin, J.B., Clasen, J.L., Greig, H.S., Kratina, P. & Thompson, P.L. (2012) Warm-
881 ing shifts top-down and bottom-up control of pond food web structure and function.
882 *Philosophical Transactions of the Royal Society of London B: Biological Sciences* **367**,
883 3008–3017.
- 884 Silvertown, J. (2004) Plant coexistence and the niche. *Trends in Ecology and Evolution*
885 **19**, 605–611.
- 886 Skvortsov, A.K. (1999) Willows of Russia and adjacent countries. Taxonomical and geo-
887 graphical revision. *Univ. Joensuu Fac. Math. Natl. Sci. Rep. Ser* **39**, 1–307.
- 888 Soberón, J. (2007) Grinnellian and eltonian niches and geographic distributions of species.
889 *Ecology Letters* **10**, 1115–1123.

890 Staniczenko, P.P., Sivasubramaniam, P., Suttle, K.B., Pearson, R.G. & Lawler, J. (????)
 891 Linking macroecology and community ecology: refining predictions of species distribu-
 892 tions using biotic interaction networks. *Ecology Letters* **20**, 693–707.

893 Stouffer, D.B., Camacho, J., Jiang, W. & Amaral, L.a.N. (2007) Evidence for the existence
 894 of a robust pattern of prey selection in food webs. *Proceedings of the Royal Society B:*
 895 *Biological Sciences* **274**, 1931–40.

896 Thuiller, W. (2003) Biomod – optimizing predictions of species distributions and project-
 897 ing potential future shifts under global change. *Global Change Biology* **9**, 1353–1362.

898 Tilman, D. (1982) *Resource competition and community structure*. Princeton University
 899 Press, Princeton.

900 Townsend, P.A., Soberón, S., Pearson, R.G., Anderson, R.P., Martínez-Meyer, E., Miguel,
 901 N. & Araújo, M.B. (2011) *Ecological Niches and Geographic Distributions*. Princeton
 902 University Press, Princeton.

903 Trøjelsgaard, K., Jordano, P., Carstensen, D.W. & Olesen, J.M. (2015) Geographical
 904 variation in mutualistic networks: similarity, turnover and partner fidelity. *Proceedings*
 905 *of the Royal Society of London B: Biological Sciences* **282**, 20142925.

906 Tylianakis, J.M., Tscharntke, T. & Lewis, O.T. (2007) Habitat modification alters the
 907 structure of tropical host–parasitoid food webs. *Nature* **445**, 202–205.

908 Veech, J.A. (2013) A probabilistic model for analysing species co-occurrence. *Global Ecol-*
 909 *ogy and Biogeography* **22**, 252–260.

910 Veijalainen, A., Sääksjärvi, I., Erwin, T., Gómez, I. & Longino, J. (2013) Subfamily
 911 composition of Ichneumonidae (Hymenoptera) from western Amazonia: insights into
 912 diversity of tropical parasitoid wasps. *Insect Conservation and Diversity* **6**, 28–37.

913 Vellend, M., Srivastava, D.S., Anderson, K.M., Brown, C.D., Jankowski, J.E., Kleyn-
914 hans, E.J., Kraft, N.J.B., Letaw, A.D., Macdonald, A.A.M., Maclean, J.E., Myers-
915 Smith, I.H., Norris, A.R. & Xue, X. (2014) Assessing the relative importance of neutral
916 stochasticity in ecological communities. *Oikos* **123**, 1420–1430.

917 Warton, D.I., Blanchet, G.F., O’Hara, R.B., Ovaskainen, O., Taskinen, S., Walker, S.C.
918 & Hui, F.K. (2015) So many variables: Joint modeling in community ecology. *Trends*
919 *in Ecology and Evolution* **30**, 766–779.

920 Williams, R. & Martinez, N. (2000) Simple rules yield complex food webs. *Nature* **404**,
921 180–183.

922 Williams, R.J., Anandanadesan, A. & Purves, D. (2010) The probabilistic niche model
923 reveals the niche structure and role of body size in a complex food web. *PloS One* **5**,
924 e12092.

925 Wisz, M.S., Pottier, J., Kissling, W.D., Pellissier, L., Lenoir, J., Damgaard, C.F., Dor-
926 mann, C.F., Forchhammer, M.C., Grytnes, J.A., Guisan, A., Heikkinen, R.K., Høye,
927 T.T., Kühn, I., Luoto, M., Maiorano, L., Nilsson, M.C., Normand, S., Ockinger, E.,
928 Schmidt, N.M., Termansen, M., Timmermann, A., Wardle, D.a., Aastrup, P. & Sven-
929 ning, J.C. (2012) The role of biotic interactions in shaping distributions and realised
930 assemblages of species: implications for species distribution modelling. *Biological Re-*
931 *views* .

932 Wood, S.A., Russell, R., Hanson, D., Williams, R.J. & Dunne, J.A. (2015) Effects of
933 spatial scale of sampling on food web structure. *Ecology and Evolution* **5**, 3769–3782.

934 Woodward, G., Perkins, D.M. & Brown, L.E. (2010) Climate change and freshwater
935 ecosystems: impacts across multiple levels of organization. *Philosophical transactions*
936 *of the Royal Society of London. Series B, Biological sciences* **365**, 2093–106.

937 Wu, J., Nyman, T., Wang, D.C., Argus, G., Yang, Y.P. & Chen, J.H. (2015) Phylogeny
938 of *Salix* subgenus *Salix* s.l. (Salicaceae): delimitation, biogeography, and reticulate
939 evolution. *BMC Evolutionary Biology* **15**, e31.

Table 1: Summary of model comparison for the interaction between the leaf folder *Phyllocolpa prussica*) and the parasitoid *Chrysocharis elongata*

#	Metaweb model	Co-occurrence model	LL	npars	AIC
1	$P(L_{ijy})$	$P(X_{iy}, X_{jy} E_y)$	-71.1	6	154.2
2	$P(L_{ijy} X_{iy}, X_{jy})$	$P(X_{iy}, X_{jy} E_y)$	-65.7	6	143.4
3	$P(L_{ijy} X_{iy}, X_{jy}, E_y)$	$P(X_{iy}, X_{jy} E_y)$	-65.6	10	151.3
4	$P(L_{ijy} X_{iy}, X_{jy}, E_y)$	$P(X_{iy}, X_{jy})$	-84.5	6	183
5	$P(L_{ijy} X_{iy}, X_{jy}, E_y)$	$P(X_{iy})P(X_{jy})$	-80.7	7	173.4
6	$P(L_{ijy} X_{iy}, X_{jy}, E_y)$	$P(X_{iy} E_y)P(X_{jy} E_y)$	-68.8	15	167.6

Table 2: Summary of model comparison for the interaction across all pairs of *Salix*, gallers and parasitoids.

Interaction	#	Metaweb model	Co-occurrence model	LL	npars	AIC
<i>Salix</i> -Galler	1	$P(L_{ijy})$	$P(X_{iy}, X_{jy} E_y)$	-6022.1	7548	27140.3
	2	$P(L_{ijy} X_{iy}, X_{jy})$	$P(X_{iy}, X_{jy} E_y)$	-5547.9	7548	26191.8
	3	$P(L_{ijy} X_{iy}, X_{jy}, E_y)$	$P(X_{iy}, X_{jy} E_y)$	-5364.0	12580	35888.0
	4	$P(L_{ijy} X_{iy}, X_{jy}, E_y)$	$P(X_{iy}, X_{jy})$	-5998.4	8806	30287.2
	5	$P(L_{ijy} X_{iy}, X_{jy}, E_y)$	$P(X_{iy})(X_{jy})$	-6636.1	7548	27092.7
	6	$P(L_{ijy} X_{iy}, X_{jy}, E_y)$	$P(X_{iy} E_y)P(X_{jy} E_y)$	-6002.9	18870	49745.7
Galler-Parasitoid	1	$P(L_{ijy})$	$P(X_{iy}, X_{jy} E_y)$	-22068.1	19206	82548.2
	2	$P(L_{ijy} X_{iy}, X_{jy})$	$P(X_{iy}, X_{jy} E_y)$	-19504.8	19206	77421.6
	3	$P(L_{ijy} X_{iy}, X_{jy}, E_y)$	$P(X_{iy}, X_{jy} E_y)$	-20217.0	32010	104454.1
	4	$P(L_{ijy} X_{iy}, X_{jy}, E_y)$	$P(X_{iy}, X_{jy})$	-19591.3	22407	77594.5
	5	$P(L_{ijy} X_{iy}, X_{jy}, E_y)$	$P(X_{iy})P(X_{jy})$	-22491.5	19206	89796.9
	6	$P(L_{ijy} X_{iy}, X_{jy}, E_y)$	$P(X_{iy} E_y)P(X_{jy} E_y)$	-18936.9	48015	133903.7

Figure legends

Figure 1

Non-random sampling of the metaweb. Network assembly can be viewed as a sampling process of the regional pool of potential interactions. Species (indicated by coloured nodes) are sampled first, and among the species found in the local network, only some interactions (indicated by blue links) occur. We characterize these sampling processes with the quantitative framework proposed in this paper. As a concrete illustration of metaweb sampling, we here show a local interaction network among *Salix* (left/green), gallers (center/red), and parasitoids (red/blue). The metaweb was constructed by aggregating interactions observed across 370 local networks.

Figure 2

Visual representation of the integrated niche. In biogeography, the niche is considered the set of environmental conditions where the intrinsic growth rate r is positive (Holt, 2009). The horizontal axis represents an environmental gradient impacting the growth of a focal species (red point). The location of each species (grey points) along this gradient represents their optimum, and the vertical dotted lines represent the limits of the Grinnellian niche of the focal species. In food web ecology, the Eltonian niche represents the location of a species in the food web, as determined by its niche position (n) and its niche optimum (c). The vertical axis represents a niche gradient, for example a trait such as body size. The location of each species along this gradient represents their niche position. The focal species will feed only on prey species occupying niche locations within a given interval around the optimum, represented by the horizontal lines. The integrated Grinnellian and Eltonian niche corresponds to the square in the middle where an interaction is possible owing to a match of traits and spatial distribution. The central

964 square represents the area where the joint probability of observing co-occurrence and
965 interactions is positive.

966 **Figure 3**

967 **Probabilistic representation of the interaction probability between a leaf folder**
968 **(*Phyllocolpa prussica*) and a parasitoid (*Chrysocharis elongata*) across gra-**
969 **dients of annual average temperature and annual precipitation.** The colour
970 gradient represents the probability of observing the event, from 0 (white) to 1 (black).
971 The representation is based on predictions from Model 3 (see Table 1). In the left panel,
972 open circles represent the absence of both species, whereas closed circles represent co-
973 occurrence and plus signs the occurrence of only one of the two species. In the other
974 two panels, open circles represent co-occurrence but an absence of interaction and closed
975 circles the occurrence of an interaction.

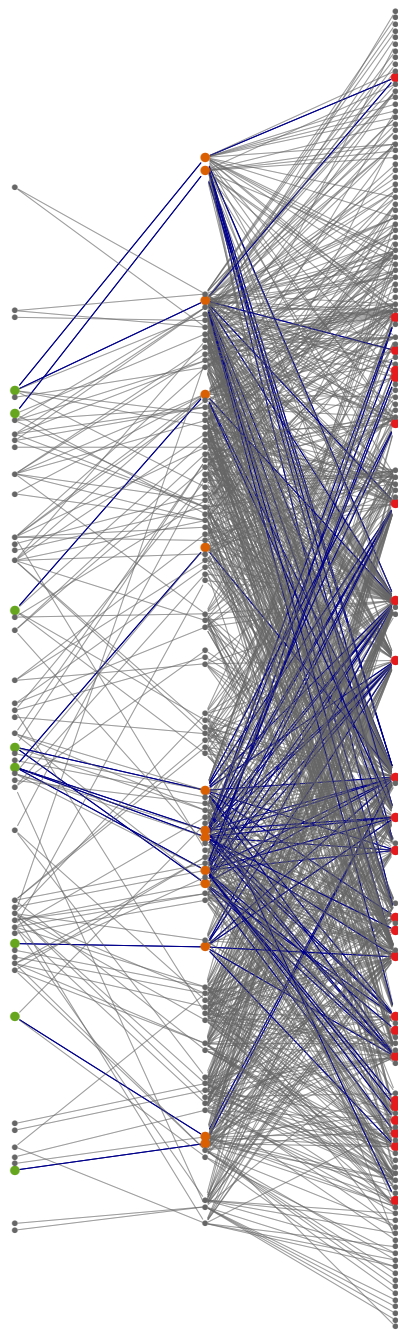
976 **Figure 4**

977 **Probabilistic representation of the interaction probability between a leaf folder**
978 **(*Phyllocolpa prussica*) and a parasitoid (*Chrysocharis elongata*) across Eu-**
979 **rope.** Climate data are extracted from WorldClim over a 1km x 1km grid and predictions
980 made according to the model illustrated at Fig. 3. Colour gradient ranges from low values
981 (dark blue) to high values (yellow).

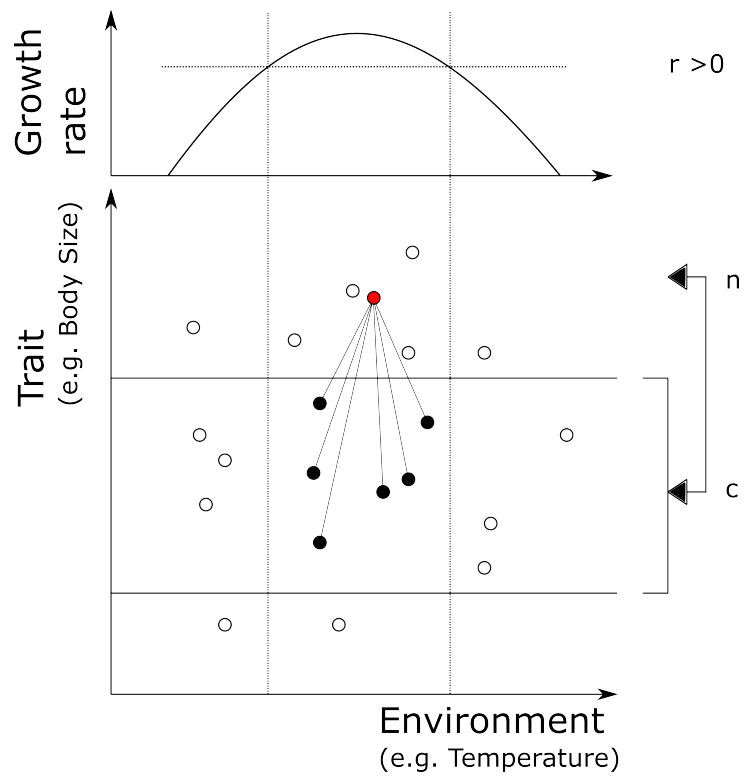
982 **Figure 5**

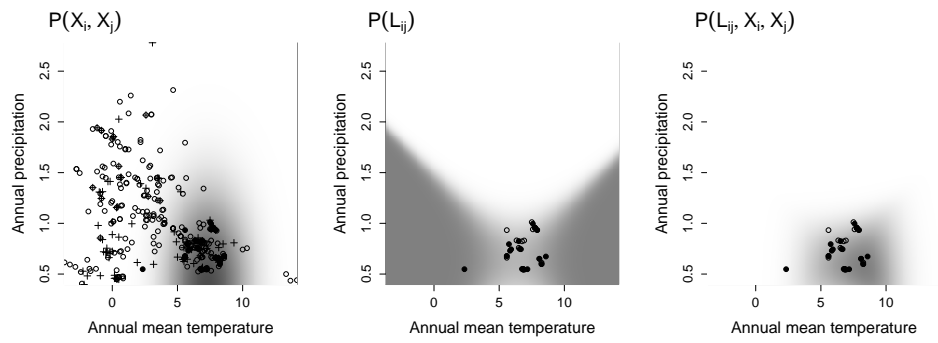
983 **Mapping the distribution of species richness, the number of links and con-**
984 **nectance across Europe.** The representation is based on predictions from Model 3
985 (see Table 2) using climate data extracted from WorldClim over a 1km x 1km grid.
986 Species richness is obtained by summation of individual occurrence probabilities, and

987 link density by summation of interaction probabilities. Colour gradient ranges from low
988 values (dark blue) to high values (yellow).



990 **Figure 2**





992 **Figure 4**

