

¹ **Title:** Bringing Elton and Grinnell together: a quantitative framework to represent
² the biogeography of ecological interaction networks

³

⁴ Abstract

⁵ Biogeography has traditionally focused on the spatial distribution and abundance of
⁶ species. Both are driven by the way species interact with one another, but only recently
⁷ community ecologists realized the need to document their spatial and temporal variation.
⁸ Here, we call for an integrated approach, adopting the view that community structure
⁹ is best represented as a network of ecological interactions, and show how it translates to
¹⁰ biogeography questions. We propose that the ecological niche should encompass the effect
¹¹ of the environment on species distribution (the Grinnellian dimension of the niche) and
¹² on the ecological interactions among them (the Eltonian dimension). Starting from this
¹³ concept, we develop a quantitative theory to explain turnover of interactions in space
¹⁴ and time – *i.e.* a novel approach to interaction distribution modeling. We apply this
¹⁵ framework to host–parasite interactions across Europe and find that two aspects of the
¹⁶ environment (temperature and precipitation) exert a strong imprint on species co- occur-
¹⁷ rence, but not on species interactions. Even where species co-occur, interaction proves to
¹⁸ be stochastic rather than deterministic, adding to variation in realized network structure.
¹⁹ We also find that a large majority of host-parasite pairs are never found together, thus
²⁰ precluding any inferences regarding their probability to interact. This first attempt to
²¹ explain variation of network structure at large spatial scales opens new perspectives at
²² the interface of species distribution modeling and community ecology.

²³ **Keywords:** networks, spatial ecology, co-occurrence, probability of interaction

²⁴

25 **Introduction**

26 Community ecology is *the study of the interactions that determine the distribution and*
27 *abundance of organisms* (Krebs, 2009). Despite a general consensus on this definition
28 (Scheiner & Willig, 2007), research on variation in community structure (beta-diversity)
29 has mostly focused on the spatial and temporal turnover of species composition (Anderson
30 *et al.*, 2011). Such research usually neglects variation in the way species interact with
31 each other, despite accumulating empirical evidence that is a major source of diversity
32 (Poisot *et al.*, 2015). Given this omission, it is perhaps not surprising that biogeographers
33 are still struggling to establish whether interactions actually impacts the distribution of
34 species at large spatial scales (Kissling *et al.*, 2012). An interaction is conceived as
35 the direct effect of the action of one species on the demography on another species;
36 whether it scales up and impact the (co-)distribution of species remains matter of debate
37 at the moment (Wisz *et al.*, 2013; González-Salazar *et al.*, 2013; Cazelles, 2016; Harris,
38 2016; Godsoe *et al.*, 2017). Treating interactions as fixed events nonetheless neglects a
39 large part of the complexity of empirical communities, and will most likely deliver biased
40 metrics (Poisot *et al.*, 2016b). Recent attempts at accounting for interactions in species
41 distribution models (Pollock *et al.*, 2014; Pellissier *et al.*, 2013; Ovaskainen *et al.*, 2017)
42 have brought some methodological advances, but are not sufficient for two reasons. First,
43 these techniques are still based on a ‘species-based’ approach to communities, where
44 interactions are merely treated as fixed covariates affecting distribution. Second, they fail
45 to provide a conceptual step forward, both in their treatment of interactions and in the
46 quality of the predictions they make.

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

51 by links, so that networks provide additional information on community structure. To
52 date, studies of network diversity have mostly been concerned with the distribution of
53 interactions within locations, and less so with variation among locations (Dunne, 2006;
54 Bascompte & Jordano, 2007; Ings *et al.*, 2009; Kéfi *et al.*, 2012). There is, however, am-
55 ple evidence that interaction networks vary in space and time (Laliberté & Tylianakis,
56 2010; Poisot *et al.*, 2012; Schleuning *et al.*, 2012; Albouy *et al.*, 2014; Poisot *et al.*, 2016c;
57 Trøjelsgaard *et al.*, 2015), even though there is no common framework with which to gen-
58 eralize these results. Metacommunity theory provides explanations for variation in the
59 distribution of the nodes (Gravel *et al.*, 2011; Pillai *et al.*, 2011; Cazelles *et al.*, 2015), but
60 there are no such explanation to the variation of node and link occurrences. Consequently,
61 we need theory to formalize these observations, as it is the only way towards fulfilling
62 the goal of community ecology: providing cogent predictions about, and understanding
63 of, the structure of ecological communities.

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

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

81 The two dimensions of community structure

82 The problem of community assembly is often formulated as *how are species sampled from*
83 *a regional pool to constitute a local community* (Götzenberger *et al.*, 2012)? This question
84 could be rewritten to address the problem of network assembly, as *how do samples from*
85 *a regional pool of interactions constitute a local interaction network?* An illustration of
86 this problem for a food web is provided in Fig. 1. The regional pool of interactions,
87 the *metaweb*, represents potential interactions among all species that could be found in a
88 given area. In this particular case, there are 275 nodes, and 1173 links among the plants
89 (52 nodes), herbivores (96 nodes), and parasitoids (127 nodes) from Northern Europe. An
90 instance of a local community is also illustrated, with 45 nodes and 93 interactions. Only
91 55.0% of all potential interactions (plant-herbivore or herbivore-parasitoid combinations)
92 are realized in this local network, revealing the stochastic nature of ecological interactions.
93 Our objective here is to provide theory to explain the sampling of the regional pool of
94 interactions, along with a quantitative method to predict it. The problem could be
95 formalized sequentially by understanding first why only a fraction of the species co-occur
96 locally and second why these species do or do not interact.

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

an additional role, either because colonization and extinction events (Hanski, 1999) are inherently unpredictable or because strong non-linear feedbacks in community dynamics generate alternative transients and equilibria (Chase, 2007; Vellend *et al.*, 2014). Interaction themselves may impact co-distribution, as hypothesized long ago by Diamond (1975). Analyses of community turnover are usually performed with data represented in a table with rows corresponding to sites (or measurements) and columns to species. Metrics of beta diversity quantify the variance of this community data (Legendre *et al.*, 2005). Traditional approaches rely on measures of dissimilarity among communities, such as the Jaccard or Bray–Curtis indices. More recent approaches decompose total variation of the community data into species and site contributions to beta diversity (Legendre & De Cáceres, 2013), and further partition it into dissimilarity due to changes in species richness and dissimilarity due to actual species turnover (Baselga, 2010). Even though all of these methods compare whole lists of species among sites or measurements, they remain fundamentally "species-based", since they report variation within columns. None of them explicitly considers variation of associations (i.e., of pairs or higher-order motifs – Stouffer *et al.* 2007). Some previous studies have considered how species distribution could be influenced by the joint effects of the abiotic and biotic environment (Stephens & Heau, 2009; ?; Cazelles *et al.*, 2015; Ovaskainen *et al.*, 2017), here we inverse the problem and describe how the distribution of biotic interactions is influenced by species distribution and the environment.

We are now getting a better understanding of interaction turnover. As mentioned above, in the network approach to community structure, species and interactions are represented by nodes and links, respectively. Associations can also be represented by matrices in which entries represent the occurrence or intensity of interactions among species (rows and columns). Network complexity is then computed as the number of interactions (in the case of binary networks) or interaction diversity (in the case of quantitative net-

127 works, Bersier *et al.* 2002). Variability in community structure consequently arises from
128 the turnover of species composition, along with turnover of interactions among pairs of
129 species. The occurrence and intensity of interactions could vary because of the environ-
130 ment, species abundance, and higher-order ecological interactions (Poisot *et al.*, 2015).
131 Variation in community composition can be independent of variation of ecological inter-
132 actions, suggesting that species and interaction distribution may well respond to different
133 drivers (Poisot *et al.*, 2012).

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

149 Interestingly, the food web literature also has its own "niche model" to position a
150 species in a community (Williams & Martinez, 2000), generalized later to other types of
151 interaction networks (Eklöf *et al.*, 2013). The niche of a species in this context represents
152 the multidimensional space of all of its interactions. Each species is characterized by a

niche position, an optimum and a range over three to five different niche axes (Williams & Martinez, 2000; Eklöf *et al.*, 2013). The niche model of food web structure and its variants have successfully explained the complexity of a variety of networks, from food webs to plant-pollinator systems (Allesina *et al.*, 2008; Williams *et al.*, 2010; Eklöf *et al.*, 2013). This conceptual framework is, however, limited to local communities, and does not provide any explanation for the turnover of network structure along environmental gradients.

The integrated niche

Despite several attempts to update the concept of the ecological niche, ecologists have not moved far beyond the "n-dimensional hypervolume" defined by Hutchinson (1957). Despite its intuitive interpretation and easy translation into species distribution models (Boulangeat *et al.*, 2012; Blonder *et al.*, 2014), the concept has been frequently criticized (Hardin, 1960; Peters, 1991; Silvertown, 2004), and several attempts have been made to expand and improve it (Pulliam, 2000; Chase & Leibold, 2003; Soberón, 2007; Holt, 2009; McInerny & Etienne, 2012b).

Part of the problem surrounding the niche concept has been clarified with the distinction between Eltonian and Grinnellian definitions (Chase & Leibold, 2003). The Grinnellian dimension of the niche is the set of abiotic environmental conditions required for a species to maintain a population in a location. The Grinnellian niche is intuitive to apply, and constitutes the conceptual backbone of species distribution models. The Eltonian niche, on the other hand, is 'the place of a species in its biotic environment, its relations to food and enemies'. While this aspect of the niche is well known by community ecologists, it is trickier to turn into predictive models. Nonetheless, the development of the niche model of food web structure (Williams & Martinez, 2000) and its parameteri-

¹⁷⁷ zation using functional traits (Gravel *et al.*, 2013; Bartomeus *et al.*, 2016) made it more
¹⁷⁸ operational.

¹⁷⁹ These perspectives are rather orthogonal to each other, and this has resulted in con-
¹⁸⁰ siderable confusion in the literature (McInerny & Etienne, 2012a). Chase & Leibold 2003
¹⁸¹ attempted to reconcile with the following definition: "*[The niche is] the joint description*
¹⁸² *of the environmental conditions that allow a species to satisfy its minimum requirements*
¹⁸³ *so that the birth rate of a local population is equal to or greater than its death rate along*
¹⁸⁴ *with the set of per capita effects of that species on these environmental conditions*". Their
¹⁸⁵ representation merges zero-net-growth isolines delimiting the Grinnellian niche ("when
¹⁸⁶ does the population persist?") with impact vectors delimiting the Eltonian niche ("what is
¹⁸⁷ the per-capita impact?"). While this representation has been very influential in local-scale
¹⁸⁸ community ecology (the resource-ratio theory of coexistence, Tilman 1982), it remains
¹⁸⁹ impractical at larger spatial scales because of the difficulties in measuring it. The absence
¹⁹⁰ of any mathematical representation of the niche that can be easily fit to ecological data
¹⁹¹ may explain why biogeographers are still struggling to develop species distribution mod-
¹⁹² els that also consider ecological interactions. Thus, a more integrative description of the
¹⁹³ niche will be key to understand spatial and temporal turnover in community structure.

¹⁹⁴ We propose to integrate the two perspectives of the niche using a visual representa-
¹⁹⁵ tion of both components (Fig. 2). The underlying rationale is that, in addition to the
¹⁹⁶ environmental constraints on demographic performance (Fig. 2, top panel), any organism
¹⁹⁷ requires resources to meet its metabolic demands and to sustain reproduction (Fig. 2,
¹⁹⁸ nodes in network of bottom panel). Abiotic environmental axes are any non-consumable
¹⁹⁹ factors affecting the demographic performance of an organism. Alternatively, the resource
²⁰⁰ axes are traits of the resources that allow interactions with the consumers. By definition
²⁰¹ of what an interaction is, these axes also influence the growth rate. The niche can there-
²⁰² fore be viewed as the set of abiotic environmental conditions (the Grinnellian component)

203 along with the set of traits (the Eltonian component) that allow a population to establish
204 and to grow at a location. This visual representation make operational the theoretical
205 interpretation of the niche and species ranges by Godsoe *et al.* (2017). Accordingly, each
206 species can be characterized by an optimal position along both the environmental (x-
207 axis) and the trait (y-axis) plane. The integrated niche is then the hypervolume where
208 interactions can occur and where a population has a positive growth rate.

209 This approach changes the representation of the niche, putting species distributions
210 and ecological interactions into the same formalism. Moreover, it allows the limits of
211 the niche axes to be independent of each other (as in the example in Fig. 2), or to
212 interact. Some of these axes may not be independent. For instance, the optimal prey
213 size for predatory fishes could decline with increasing temperature (Gibert & DeLong,
214 2014), which would make diet boundaries functions of the environment. Alternatively,
215 we could also consider that the growth rate of the predator changes with the size of its
216 prey items, thereby altering the environmental boundaries and changing the shape of the
217 niche in the middle of this visual representation. It is also important that this conceptual
218 representation depicts a reality that could be highly dimensional and sometimes difficult
219 to represent statistically (Clark *et al.*, 2007).

220 A probabilistic representation of interaction networks in 221 space

222 We now formalize the integrated niche with a probabilistic approach to interactions and
223 distributions. In particular, we seek to represent the probability that an interaction
224 between species i and j occurs at location y . We define L_{ijy} as a stochastic process taking
225 a value of 1 when an interaction occurs and a value of 0 when it does not, and focus on
226 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 . For brevity, we have here avoided specific references to traits and considered that the interactions are provided in the data. But one could develop the model further to integrate such traits, inspired by the framework developed by Gravel et al. 2016. They proposed to represent the occurrence of an interaction between species as a joint probability event $P(L_{ij}, \mathbf{T}_i, \mathbf{T}_j | E)$, where \mathbf{T}_i

248 and \mathbf{T}_j are trait vectors for species i and j respectively. In order to develop a trait-based
249 representation of the spatial variation of interactions, one may thereafter consider how
250 these traits are related to the environment and how they co-occur. Above, we referred
251 to this entity as the "metaweb" and note that it corresponds to the Eltonian dimension
252 of the niche. Below, we will see how this formalism can be directly fit to empirical data,
253 restricting ourselves to occurrences only and referring to the above citations for trait-
254 matching and trait-environment relationships. But before turning to an application, we
255 will discuss the interpretation of different variants of these two terms.

256 Variants of co-occurrence

257 There are several variants to the co-occurrence probability, representing different hy-
258 potheses concerning spatial variation in network structure (see the explicit formulations
259 in Table 1). The simplest model relates the probability of co-occurrence directly to the
260 environment, $P(X_{iy}, X_{jy}|E_y)$. In this situation, there are no underlying assumptions
261 about the ecological processes responsible for co-occurrence. Spatial associations be-
262 tween species could arise because interactions constrain distribution, where in such case
263 the co-occurrence would be conditional on L , or, alternatively, because of environmental
264 requirements shared between i and j (Pollock *et al.*, 2014; Cazelles *et al.*, 2016). In the
265 former case, species are not independent of each other and the conditional occurrence
266 must be accounted for explicitly, $P(X_{iy}, X_{jy}|E_y) = P(X_{iy}|E_y, X_{jy})P(X_{jy}|E_y)$. In the lat-
267 ter case, species are independent, and only the marginal occurrence must be accounted
268 for, $P(X_{ijy}|E_y) = P(X_{iy}|E_y)P(X_{jy}|E_y)$.

269 The co-occurrence probability itself could depend on ecological interactions. This
270 should be viewed as the realized component of the niche (i.e. the distribution when
271 accounting for species interactions). Direct pairwise interactions such as competition,
272 facilitation, and predation have long been studied for their impact on co-distribution

(e.g. Diamond 1975; Connor & Simberloff 1979; Gotelli 2000). Second- and higher-order interactions (e.g. trophic cascades) could also affect co-occurrence (Harris, 2016; Staniczenko *et al.*, 2017). Co-occurrence of multiple species embedded in ecological networks is a topic of its own, however, and is influenced by both network topology and species richness (Cazelles *et al.*, 2016). Not only direct interactions influence co-occurrence, but indirect interactions do as well (e.g. plant species sharing an herbivore, or herbivores sharing parasitoids, could repel each other in space (Holt & Lawton, 1993)). The impact of direct interactions and first-order indirect interactions on co-occurrence tends to vanish with increasing species richness in 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, $P(L_{ijy}|X_{iy} = 1, X_{jy} = 1) = 1$. This approach might be a reasonable approximation if the spatial or temporal scale of sampling and inference is so large that

298 the probability of observing at least one interaction converges to unity. In this scenario,
299 network variation arises solely from species distributions.

300 Second, ecological interactions could also vary with the environment, so that $P(L_{ijy}|E_y)$.
301 Although it is rare to see a conditional representation of pairwise ecological interactions,
302 experimental studies have frequently revealed interactions to be sensitive to the environ-
303 ment. For instance, McKinnon *et al.* (2010) showed that predation risks of shorebirds
304 vary at the continental scale, decreasing from the south to the north. It is also common
305 to see increasing top-down control with temperature (e.g. Shurin *et al.* 2012; Gray *et al.*
306 2016). Effects of the environment on interactions (Gibert & DeLong, 2014) also propagate
307 up the community and influence network structure (Tylianakis *et al.*, 2007; Woodward
308 *et al.*, 2010; Petchey *et al.*, 2010).

309 **Application: continental-scale variation of host-parasite 310 community structure**

311 We now turn to an illustration of our theory with the analysis of an empirical dataset
312 of host-parasite networks sampled throughout the south-north environmental gradient in
313 continental Europe (Kopelke *et al.*, 2017). Our objective here is to illustrate potential
314 applications of the approach outlined above, rather than to specifically describe the bio-
315 geography of this particular community. The focal system consists of local food webs of
316 willows (genus *Salix*), their galling insects, and the natural enemies (parasitoids and in-
317 quilines) of gallers. We ask: i) how much does network structure vary across the gradient,
318 and ii) what is the primary driver of network turnover across the gradient?

319 **Data**

320 Communities of willows, gallers, and parasitoids are species-rich and widely distributed,
321 with pronounced variation in community composition across space. The genus *Salix* in-
322 cludes over 400 species, most of which are shrubs or small trees (Argus, 1997), and is
323 common in moist habitats across the Northern Hemisphere (Skvortsov, 1999). Willows
324 support a highly diverse community of herbivorous insects, with one of the main herbivore
325 groups being gall-inducing sawflies (Hymenoptera: Tenthredinidae: Nematinae: Euurina
326 (Kopelke, 1999). Gall formation is induced by sawfly females during oviposition, and
327 includes marked manipulation of host-plant chemistry by the galler (Nyman & Julkunen-
328 Tiitto, 2000). The enemy community of the gallers includes nearly 100 species belonging
329 to 17 insect families of four orders (Kopelke, 2003). These encompass two main types:
330 inquiline larvae (Coleoptera, Lepidoptera, Diptera, and Hymenoptera) feed primarily on
331 gall tissue, but typically kill the galler larva in the process, while parasitoid larvae (repre-
332 senting many families in Hymenoptera) kill the galler larvae by direct feeding (Kopelke,
333 2003).

334 Local realizations of the willow-galler-parasitoid network were reconstructed from
335 community samples collected between 1982 and 2010. During this period, willow galls
336 were collected at 370 sites across Central and Northern Europe. In total, 52, 96 and
337 127 *Salix*, galler and parasitoid and inquiline taxa were distinguished, respectively. The
338 strength of this dataset is that observations were observed *in situ*, rather than inferred
339 from expert knowledge or other sources of information, thereby allowing the analysis of
340 their spatial variation. The drawback is that, because many species are rare, some of the
341 pairs of species may have been observed co-occurring only a few times, or never. As a
342 consequence, despite the extent of the sampling, there is significant uncertainty in the
343 quantification of some of the links. The above described methodology explicitly aims at
344 describing this uncertainty, and research should be done to develop methods to reduce it

345 (see Conclusion).

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

354 Analysis

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

365 We compared three models for the co-occurrence probability. The first one directly
366 represents the co-occurrence probability conditional on the local environment, $P(X_{iy}, X_{jy}|E_y)$
367 (models are listed in Tables 1 and 2). Hence, this model makes no assumption about the
368 mechanisms driving co- occurrence for any given environment, and instead uses the in-
369 formation 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 optima in 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 evaluated 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

396 from the knowledge of a deterministic metaweb. It represents potential interactions,
397 rather than realized ones. There are consequently five parameters for the first model, a
398 single parameter for the second model and no parameter to evaluate for the third model
399 (where the interaction probability is fixed by the hypothesis).

400 We fitted the different models to each pair of species and recorded the predicted
401 probabilities. The joint probability $P(L_{ijy}, X_{iy}, X_{jy})$ was then computed from Eq. 2, and
402 the likelihood of each observation D_{ijy} given the model θ was computed as $\mathcal{L}(D_{ijy}|\theta) =$
403 $P(L_{ij}, X_{iy}, X_{jy})$ if an interaction was observed, and as $\mathcal{L}(D_{ijy}|\theta) = 1 - P(L_{ij}, X_{iy}, X_{jy})$
404 if no interaction was observed. The log-likelihood was summed over the entire dataset
405 to compare the different models by AIC. We therefore evaluate the likelihood of all local
406 networks, given the model. Not surprisingly, there was a very large number of species pairs
407 for which this model could not be computed, as they simply never co-occurred. For these
408 pairs, we have no information of the interaction probability, and they were consequently
409 removed from the analysis. The log-likelihood reported across the entire dataset was
410 therefore summed over all pairs of species observed to co-occur at least once. Interactions
411 between the first (*Salix*) and second (gallers) trophic layers and those between the second
412 and third (parasitoids) were considered separately. Finally, we used the full model (in
413 which both co-occurrence and the interaction are conditional on the environment) to
414 interpolate species distributions and interaction probabilities across the entire European
415 continent. We reconstructed the expected network for each location in a 1 X 1 km grid
416 and computed the probabilistic connectance following Poisot *et al.* (2016b).

417 All of the data are openly available in the database *mangal* (Poisot *et al.*, 2016a)
418 and all R scripts for running the analyses, are provided in the github repository https://github.com/DominiqueGravel/ms_probaweb.
419

420 **Results**

421 Despite the extensive sampling, many pairs of species were observed to co- occur only a
422 few times. This made it difficult to evaluate interaction probabilities with any reasonable
423 confidence. Thus, we start with an example of a single pair of species selected because
424 of its high number of co- occurrences ($N_{ij} = 38$): the leaf folder *Phyllocolpa prussica* and
425 the parasitoid *Chrysocharis elongata*. These two fairly abundant species were observed
426 $N_i = 49$ and $N_j = 121$ times, respectively, across the 370 sites, and they were found to
427 interact with a marginal probability $P(L_{ij}) = 0.55$, which means they interacted at 21
428 different locations. Here, a comparison of model fit (Table 2) reveals that conditioning
429 the interaction probability on local environmental conditions adds no explanatory power
430 beyond a model assuming the same probability of interaction anywhere in space (Model
431 2 vs Model 3). Moreover, when the two species co-occur, the occurrence of the interac-
432 tion was insensitive to the environment (Model 2 vs Model 3). Alternatively, climatic
433 variables significantly impacted co-occurrence (Model 3 vs Model 4). The independent
434 model performed worse than the non-random co- occurrence model (Model 3 vs Model
435 6). The full model revealed that the greatest interaction probability occurred at interme-
436 diate temperature and precipitation, simply because this is where the two species most
437 frequently co-occur (Fig. 3). The probabilities of co-occurrence and interaction can be
438 represented in space, where we find that the highest interaction probability occurred in
439 Central Europe (Fig. 4).

440 We evaluated each model for all pairs of species in order to better understand the
441 large-scale drivers of network turnover. The results were highly consistent among trophic
442 levels (*Salix*-gallers and gallers-parasitoids; Table 3), despite large variability in the fit of
443 the models to the different pairs of species. This variability was particularly important for
444 the models of interactions given co-occurrence. Across all pairs of species, the conditional
445 representation of interactions performed (Model 2) better than the marginal one (Model

446 1); that is, interactions did not occur systematically whenever the two species were found
447 co-occurring. Hence, in addition to species turnover, the stochastic nature of interactions
448 contributes to network variability. In total, we recorded 1,173 pairs of interactions, only
449 290 of which occurred more than five times. Out of these 290 interactions, 143 were
450 systematically detected whenever the two species co-occurred. In the instances when
451 species co-occurred, the two environmental variables considered proved relatively poor
452 predictors of their interactions (Model 2 vs Model 3). Not surprisingly, for both types
453 of interactions (*Salix*-galler and galler-parasitoid), the log-likelihood increased when the
454 environment was considered. However, the extra number of parameters exceeded the gain
455 in log-likelihood and inflated AIC. Therefore, the most parsimonious model excluded the
456 effect of the environment. On the basis of log-likelihood only, co-occurrence was non-
457 independent for both *Salix*-galler and galler-parasitoid interactions. Thus, according to
458 AIC, the best model was the one of non-random co-occurrence (Model 3 vs Model 6) for
459 both types of interactions.

460 Once we had selected the best model based on AIC (Model 3, Table 2), we used
461 it to reconstruct the expected species richness, along with the most likely network for
462 each location. Using this approach, we mapped the expected distribution of network
463 properties across Europe (Fig. 5). For simplicity, we chose to consider connectance
464 as our descriptor of network configuration, as this metric can be easily computed from
465 probabilistic networks (Poisot *et al.*, 2016b) and is also a good proxy for many other
466 network properties (Poisot & Gravel, 2014). Overall, we found a peak in *Salix*, gallers and
467 parasitoid diversity in Northern Europe. The expected number of interactions roughly
468 followed the distribution of species richness, but accumulated at a rate different from
469 species numbers. Connectance likewise peaked in Northern Europe (Fig. 5).

⁴⁷⁰ Interpretation

⁴⁷¹ We have proposed that the representation of community structure and its variation in
⁴⁷² space is best captured by the formalism of ecological networks, as both the distribution
⁴⁷³ of species and their interspecific interactions can then be accounted for. We consequently
⁴⁷⁴ revised the niche concept in order to integrate its abiotic and biotic components. This
⁴⁷⁵ integrated niche was represented visually with an ordination of species into an environ-
⁴⁷⁶ mental space and a trait space. The fundamental niche of a species is represented as
⁴⁷⁷ the set of environmental conditions and resources that allow a species to establish in a
⁴⁷⁸ location, thereby integrating the Eltonian and the Grinnellian components of the niche.
⁴⁷⁹ We then translated the concept mathematically by investigating the probability of the
⁴⁸⁰ joint occurrences of species and their interaction, which should be interpreted as an in-
⁴⁸¹ teraction distribution model. We used this approach to characterize the turnover of the
⁴⁸² structure of ecological interactions in a species-rich tri-trophic network across Western
⁴⁸³ Europe, finding that the primary driver of network variation is the turnover in species
⁴⁸⁴ composition.

⁴⁸⁵ Applying the framework to our large data set on host-parasite interactions across
⁴⁸⁶ Europe revealed key features in the interaction between *Salix* taxa, their herbivores,
⁴⁸⁷ and the natural enemies of these herbivores. Consistent with a general increase in the
⁴⁸⁸ diversity of *Salix* towards boreal areas (Cronk *et al.*, 2015), overall species richness of
⁴⁸⁹ the networks increased towards the north. The distribution of *Salix* species richness
⁴⁹⁰ largely matched those of gallers and parasitoids. These observations within Europe are
⁴⁹¹ also matched by the ones found at a global scale for *Salix* (Argus, 1997; Cronk *et al.*,
⁴⁹² 2015; Wu *et al.*, 2015) and sawflies (Kouki *et al.*, 1994; Kouki, 1999). Species richness
⁴⁹³ in a common group of parasitic wasps, the Ichneumonidae, was originally presumed to
⁴⁹⁴ show a similar "reversed latitudinal gradient", but this observation has been recently
⁴⁹⁵ challenged by findings of rather high ichneumonid diversity in the tropics (Veijalainen

496 *et al.*, 2013). Nevertheless, ichneumonid subfamilies specifically associated with sawflies
497 (Ctenopelmatinae, Tryphoninae) are clearly less diverse in the south.

498 Exactly what processes are responsible for the distribution of species richness at dif-
499 ferent trophic levels is yet to be established (but see e.g. Roininen *et al.* 2005; Nyman
500 *et al.* 2010; Leppänen *et al.* 2014), but as a net outcome of different latitudinal trends
501 across trophic levels, the distribution of co-occurrence and therefore of potential inter-
502 actions differed between the first and second layers of feeding links. The correlation
503 between expected *Salix* and galler richness was 0.73, while it was 0.58 between gallers
504 and their parasitoids. Therefore, the ratio of herbivores to *Salix* species is essentially
505 constant across Europe, whereas each herbivore species is potentially attacked by a richer
506 enemy community higher latitudes. Consequently, overall connectance peaks in Northern
507 Europe (Fig. 5).

508 In terms of species interacting with each other, our analysis suggests that the en-
509 vironment leaves a detectable imprint on species co-occurrence, but only a slight mark
510 on the occurrence of realized links among species in a specific place: the probability of
511 finding a given combination of species at a higher and a lower trophic level at the same
512 site was clearly affected by the environment, whereas the probability of observing an in-
513 teraction between the two was, given co-occurrence, not significantly so. The interactions
514 were highly uncertain, as only few pairs of species were systematically interacting when
515 co-occurring but most not. This applies to the example species *Phyllocolpa prussica* and
516 *Chrysocharis elongata* (Figs. 3 and 4), but also to all species pairs more generally. For
517 the example species pair, the full model revealed that the joint probability event (inter-
518 action and co-occurrence) peaks at intermediate temperature and precipitation, simply
519 because this is where the two species co-occur most often. This does not imply that
520 species will always interact when they meet – although this is a basic assumption in most
521 documented metawebs to date (e.g. Havens 1992; Wood *et al.* 2015). Rather, an inter-

action 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.

530 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 (Simanonok & Burkle, 2014) 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 pairwise interactions (Poisot *et al.*, 2012, 2015). 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 independent. 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; Baskett *et al.*, 2018) 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.*, 2016). 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

546 but non-predictably along the annual temperature and the precipitation gradient. The
547 lack of detectable signal may be due to our choice of covariates. Indeed, a previous study
548 focusing on a subset of the system studied here identified habitat characteristics as the
549 primary drivers of interactions (Nyman *et al.*, 2015). New investigations with other sys-
550 tems will thus be required to challenge this result. Under all circumstances, documenting
551 the relationship between the environment and the occurrence of interactions at conti-
552 nental scales is critical for understanding how large- scale variation of trophic regulation
553 influences community dynamics and ecosystem functioning (Harfoot *et al.*, 2014).

554 The framework we provide complements methods to compute network beta-diversity
555 presented by Poisot *et al.* 2012 and Poisot *et al.* 2018. The total network turnover is
556 partitioned into interaction turnover and species turnover – which in our approach would
557 correspond to the three terms of Eq. 2, respectively. Importantly, one could easily derive
558 the different β of Poisot *et al.* 2012 using the fitted probabilities. For instance, for a
559 pair of species i and j and sites x and y , the expectation for the fraction a for the
560 interactions would be $P(L_{ijx}|X_{ix}, X_{ijx})P(L_{ijx}|X_{ix}, X_{ijx})$. The same could be computed
561 for fractions b , and therefore the expected beta-diversity for an entire network could be
562 recomputed this way. The novel feature of the approach presented here is then that the
563 different contributions to network turnover can be attributed to different hypotheses (as
564 performed here), in order to evaluate for instance the role of environmental heterogeneity
565 or species co- distribution on network turnover.

566 We restricted our analysis to the effect of co-occurrence on ecological interactions,
567 neglecting the inverse of the problem. We did not investigate in depth the drivers of
568 co-occurrence and simply took it for granted from the data. Co-occurrence was indeed
569 different from the expectation of independent species distributions. It thus begs the
570 question of whether, once environmental effects on species-specific distribution have been
571 accounted for, interactions come with significant effects on co-occurrence? We could

572 rephrase this problem by asking whether the fundamental niche differs from the real-
573 ized niche, and how this applies to our framework. For example, we have considered
574 above simply the co-occurrence probability, $P(X_{iy}, X_{jy}|E_y)$, which could be expanded as
575 $P(X_{iy}|X_{jy}, E_y)P(X_{jy}|E_y)$. After some re-arrangement of Eq. 2, the marginal occurrence
576 probability, $P(X_{jy}|E_y)$, could be considered as a species distribution model taking into
577 account the interaction between these species (in the same spirit as done by Gonzalez-
578 Salazar et al. 2013). This derivation would however critically depend on a strong *a priori*
579 expectation of the conditional probability of observing a species given the distribution
580 of the other species. This assumption seems reasonable for some situations, such as a
581 parasitoid species that requires a host to develop. On the other hand, Cazelles (2016)
582 found that the strength of this association is often rather weak if not inexistant (for in-
583 stance, with the example pair presented in Table 1). The lack of an association could
584 simply arise when the parasitoid is generalist enough that it is not constrained to track
585 the distribution of any single/given host (Cazelles *et al.*, 2015).

586 The approach we presented could easily be expanded to account for traits, in order to
587 derive a complete and more mechanistic representation of the niche. In this context, both
588 the Eltonian and Grinnellian dimensions of the niche could be represented. As mentioned
589 earlier, Gravel *et al.* 2016 proposed to represent the occurrence of interactions based on
590 some trait-matching rules (given by a function of the type $P(L_{ij})|\mathbf{T}_i, \mathbf{T}_j, E)$) and the co-
591 distribution of traits (corresponding to $P(\mathbf{T}_i, \mathbf{T}_j|E)$). The key insight from the formalism
592 proposed by Gravel *et al.* 2016 is that the co-distribution of traits will impact variation in
593 network structure. As a result, it was found that network properties often co-vary with
594 functional structure and functional diversity Laigle *et al.*, 2018. The next step in the
595 development of this framework will be to develop the models for the co-distribution of
596 traits, for instance following Ovaskainen et al. 2017. This may facilitate the accessibility
597 and the applicability of the framework, since the basic unit of analysis is no longer the

598 species, but rather a combination of traits, thereby facilitating the statistical analysis by
599 sharing information among species.

600 At present, there is only indirect support for the hypothesis that interacting species
601 are conditionally distributed, but this possibility should be the topic of theoretical inves-
602 tigation with dynamical metacommunity models (Cazelles *et al.*, 2015) and empirical hy-
603 pothesis testing. The impact of ecological interactions on the distribution of co-occurrence
604 has been the topic of many publications since Diamond (1975) seminal study on competi-
605 tion and "checkerboard" distribution, but pairwise approaches have only recently received
606 attention (Veech, 2013). Whether two interacting species are more closely associated in
607 space remains unclear, since most approaches based on null models consider community-
608 level metrics (e.g. Gotelli 2000), such as the C-score, thereby making it hard to evaluate
609 if specific interactions do indeed affect co-occurrence. The expansion of the framework
610 we describe to account for the difference between the realized and the fundamental niche
611 will therefore require further investigation of the impact of interactions on co-occurrence.

612 Ecological networks are known to be extremely sparse, *i.e.* they have far more absences
613 than presences of interactions. Absences of interactions, however, can come from different
614 sources. The fact that unequal sampling at the local scale can affect our understanding
615 of network structure is well documented (Martinez *et al.*, 1999). In a spatial context,
616 however, some interactions may be undocumented simply because the species involved
617 have never been observed to co-occur. Although these cases are reported as a lack of
618 interactions, in actuality we cannot make any reliable inference from them: since the
619 species have never been observed together, it remains possible that they would interact
620 if they did. A fundamentally different category of absences of interactions are then those
621 reported after multiple observations of species co-occurrence. Thus, to gain confidence
622 that the probability of an interaction is low, extensive sampling (that is, several records
623 of co- occurrence) is needed. Generally, our confidence that the interaction is indeed

impossible will increase with the number of observations of the species pair. Seeing that this is essentially a Bernoulli process (the probability that the species will interact given their presence), the breadth of the confidence interval is expected to saturate after a fixed number of observations, which can be set as a threshold above which a species pair has finally been observed "often enough". For instance, if two species are observed together at $k = 10$ locations but are never found interacting ($n = 0$), this means that the maximum likelihood estimate for the probability of this binomial distribution is $p = 0$. Using the Clopper-Pearson method to compute the 95% confidence interval, we find that the real value of the probability of the interaction occurring is somewhere between 0 and 0.31. This interval is rather large, indicating that a significant amount of sampling is required to have credible interaction networks. Fortunately, there are Bayesian methods which could be used to complement available information and reduce this uncertainty (Cirtwill *et al.*, 2018).

Conclusion

Our representation of spatial variation of community structure offers a new approach for 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.

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

653 **3. From interaction probabilities to a distribution of network properties.**
654 Metrics are available to analyse the structure of probabilistic networks (Poisot *et al.*,
655 2016b). These metrics are useful as first approximation, but they assume independence
656 among interactions. This might not be the case in nature because of the role of co-
657 occurrence and shared environmental requirements. We also need to better understand
658 the distribution of network properties arising from probabilistic interactions.

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

663 **5. Effects of ecological interactions on co-occurrence.** We have intentionally
664 omitted the feedback of ecological interactions on co-occurrence in this framework. As
665 abundance can impact the occurrence of interactions, and, conversely since interactions
666 impact abundance (Canard *et al.*, 2014), we could reasonably expect that interactions
667 will also influence co-occurrence. Theory in this regard does exist for simple three-species
668 modules (Cazelles *et al.*, 2015), but its extension to entire co-occurrence networks will
669 prove critical in the future, especially given the interest in using co- occurrence to infer
670 ecological interactions (Morales-Castilla *et al.*, 2015; Morueta-Holme *et al.*, 2016).

671 Acknowledgements

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

678 **References**

- 679 Albouy, C., Velez, L., Coll, M., Colloca, F., Le Loc'h, F., Mouillot, D. & Gravel, D. (2014)
680 From projected species distribution to food-web structure under climate change. *Global
681 Change Biology* **20**, 730–741.
- 682 Allesina, S., Alonso, D. & Pascual, M. (2008) A general model for food web structure.
683 *Science* **320**, 658–61.
- 684 Anderson, M.J., Crist, T.O., Chase, J.M., Vellend, M., Inouye, B.D., Freestone, A.L.,
685 Sanders, N.J., Cornell, H.V., Comita, L.S., Davies, K.F., Harrison, S.P., Kraft, N.J.B.,
686 Stegen, J.C. & Swenson, N.G. (2011) Navigating the multiple meanings of beta-
687 diversity: a roadmap for the practicing ecologist. *Ecology Letters* **14**, 19–28.
- 688 Argus, G. (1997) Infrageneric classification of Salix (Salicaceae) in the New World. *Sys-
689 tematic Botany Monographs* **52**, 1–121.
- 690 Bartomeus, I., Gravel, D., Tylianakis, J., Aizen, M., Dickie, I. & Bernard-Verdier, M.
691 (2016) A common framework for identifying linkage rules across different types of in-
692 teractions. *Functional Ecology* **30**, 1894–1903.
- 693 Bascompte, J. & Jordano, P. (2007) Plant-Animal Mutualistic Networks: The Architec-
694 ture of Biodiversity. *Annual Review of Ecology, Evolution, and Systematics* **38**, 567–
695 593.
- 696 Baselga, A. (2010) Partitioning the turnover and nestedness components of beta diversity.
697 *Global Ecology and Biogeography* **19**, 134–143.
- 698 Baskett, C.A., Schemske, D.W. & Novotny, V. (2018) Latitudinal patterns of herbivore
699 pressure in a temperate herb support the biotic interactions hypothesis. *Ecology Letters*
700 **21**, 578–587.

- 701 Bersier, L.F., Banašek-Richter, C. & Cattin, M.F. (2002) Quantitative descriptors of
702 food-web matrices. *Ecology* **83**, 2394–2407.
- 703 Blonder, B., Lamanna, C., Violle, C. & Enquist, B.J. (2014) The n-dimensional hyper-
704 volume. *Global Ecology and Biogeography* **23**, 595–609.
- 705 Boulangeat, I., Gravel, D. & Thuiller, W. (2012) Accounting for dispersal and biotic
706 interactions to disentangle the drivers of species distributions and their abundances.
707 *Ecology Letters* **15**, 584–593.
- 708 Canard, E., Mouquet, N., Mouillot, D., Stanko, M., Miklisova, D. & Gravel, D. (2014)
709 Empirical evaluation of neutral interactions in host-parasite networks. *The American
710 naturalist* **183**, 468–79.
- 711 Cazelles, K. (2016) *Influence des interactions biotiques sur la répartition géographique des
712 espèces*. PhD. Dissertation. Université du Québec à Rimouski, Rimouski.
- 713 Cazelles, K., Araújo, M.B., Mouquet, N. & Gravel, D. (2016) A theory for species co-
714 occurrence in interaction networks. *Theoretical Ecology* **9**, 39–48.
- 715 Cazelles, K., Mouquet, N., Mouillot, D. & Gravel, D. (2015) On the integration of biotic
716 interaction and environmental constraints at the biogeographical scale. *Ecography* **39**,
717 921–931.
- 718 Chase, J. & Leibold, M. (2003) *Ecological niches*. Chicago University Press, Chicago.
- 719 Chase, J.M. (2007) Drought mediates the importance of stochastic community assembly.
720 *Proceedings of the National Academy of Sciences of the United States of America* **104**,
721 17430–4.
- 722 Cirtwill, A., Eklöf, A., Roslin, T., Wootton, K. & Gravel, D. (2018) A quantitative
723 framework for investigating the reliability of network construction. *bioRxiv* .

- 724 Clark, J.S., Dietze, M., Chakraborty, S., Agarwal, P.K., Ibanez, I., LaDeau, S. & Wolosin,
725 M. (2007) Resolving the biodiversity paradox. *Ecology Letters* **10**, 647–659.
- 726 Connor, E. & Simberloff, D. (1979) The assembly of species communities: chance or
727 competition? *Ecology* **60**, 1132–1140.
- 728 Cronk, Q., Ruzzier, E., Belyaeva, I. & Percy, D. (2015) Salix transect of Europe: lati-
729 tudinal patterns in willow diversity from Greece to arctic Norway. *Biodiversity Data*
730 *Journal* **3**, e6258.
- 731 Diamond, J. (1975) Assembly of species communities. *Ecology and evolution of communi-*
732 *ties* (eds. M. Cody & J. Diamond), pp. 342–444, Harvard University Press, Cambridge.
- 733 Dunne, J.A. (2006) The network structure of food webs. *Ecological networks: Linking*
734 *structure and dynamics* (eds. M. Pascual & J.A. Dunne), pp. 27–86, Oxford University
735 Press, Oxford.
- 736 Eklöf, A., Jacob, U., Kopp, J., Bosch, J., Castro-Urgal, R., Chacoff, N.P., Dalsgaard, B.,
737 de Sassi, C., Galetti, M., Guimarães, P.R., Lomáscolo, S.B., Martín González, A.M.,
738 Pizo, M.A., Rader, R., Rodrigo, A., Tylianakis, J.M., Vázquez, D.P. & Allesina, S.
739 (2013) The dimensionality of ecological networks. *Ecology Letters* **16**, 577–583.
- 740 Gibert, J.P. & DeLong, J.P. (2014) Temperature alters food web body-size structure.
741 *Biology Letters* **10**.
- 742 Gleason, H. (1926) The individualistic concept of the plant association. *Bulletin of the*
743 *Torrey Botanical Club* **53**, 7–26.
- 744 Godsoe, W., Jankowski, J., Holt, R.D. & Gravel, D. (2017) Integrating biogeography with
745 contemporary niche theory. *Trends in Ecology and Evolution* **32**, 488–499.

- 746 González-Salazar, C., Stephens, C.R. & Marquet, P.A. (2013) Comparing the relative con-
747 tributions of biotic and abiotic factors as mediators of species' distributions. *Ecological*
748 *Modelling* **248**, 57 – 70.
- 749 Gotelli, N.J. (2000) Null Model Analysis of Species Co-Occurrence Patterns. *Ecology* **81**,
750 2606.
- 751 Götzenberger, L., de Bello, F., Bråthen, K.A., Davison, J., Dubuis, A., Guisan, A., Lepš,
752 J., Lindborg, R., Moora, M., Pärtel, M., Pellissier, L., Pottier, J., Vittoz, P., Zobel,
753 K. & Zobel, M. (2012) Ecological assembly rules in plant communities—approaches,
754 patterns and prospects. *Biological reviews of the Cambridge Philosophical Society* **87**,
755 111–27.
- 756 Gravel, D., Albouy, C. & Thuiller, W. (2016) The meaning of functional trait composition
757 of food webs for ecosystem functioning. *Philosophical Transactions of the Royal Society*
758 *of London B: Biological Sciences* **371**.
- 759 Gravel, D., Massol, F., Canard, E., Mouillot, D. & Mouquet, N. (2011) Trophic theory
760 of island biogeography. *Ecology Letters* **14**, 1010–6.
- 761 Gravel, D., Poisot, T., Albouy, C., Velez, L. & Mouillot, D. (2013) Inferring food web
762 structure from predator-prey body size relationships. *Methods in Ecology and Evolution*
763 **4**, 1083–1090.
- 764 Gray, S.M., Poisot, T., Harvey, E., Mouquet, N., Miller, T.E. & Gravel, D. (2016) Temper-
765 ature and trophic structure are driving microbial productivity along a biogeographical
766 gradient. *Ecography* **39**, 981–996.
- 767 Hanski, I. (1999) *Metapopulation Ecology*. Oxford University Press, Oxford.
- 768 Hardin, G. (1960) The competitive exclusion principle. *Science* **131**, 1292–1297.

- 769 Harfoot, M.B.J., Newbold, T., Tittensor, D.P., Emmott, S., Hutton, J., Lyutsarev, V.,
770 Smith, M.J., Scharlemann, J.P.W. & Purves, D.W. (2014) Emergent global patterns of
771 ecosystem structure and function from a mechanistic general ecosystem model. *PLoS*
772 *Biol* **12**, 1–24.
- 773 Harris, D.J. (2016) Inferring species interactions from co-occurrence data with markov
774 networks. *Ecology* **97**, 3308–3314.
- 775 Havens, K. (1992) Scale and structure in natural food webs. *Science* **257**, 1107–1109.
- 776 Hijmans, R.J. (2015) *raster: Geographic Data Analysis and Modeling*. R package version
777 2.5-2.
- 778 Holt, R.D. (2009) Bringing the Hutchinsonian niche into the 21st century : Ecological
779 and evolutionary perspectives. *Proceedings of the National Academy of Sciences* **106**,
780 19659–19665.
- 781 Holt, R.D. & Lawton, J.H. (1993) Apparent competition and enemy-free space in insect
782 host-parasitoid communities. *The American Naturalist* **142**, 623–645.
- 783 Hutchinson, G. (1957) Concluding remarks. *Cold Spring Harbour Symposium* **22**, 415–
784 427.
- 785 Ings, T.C., Montoya, J.M., Bascompte, J., Blüthgen, N., Brown, L., Dormann, C.F.,
786 Edwards, F., Figueroa, D., Jacob, U., Jones, J.I., Lauridsen, R.B., Ledger, M.E.,
787 Lewis, H.M., Olesen, J.M., van Veen, F.J.F., Warren, P.H. & Woodward, G. (2009)
788 Ecological networks—beyond food webs. *The Journal of animal ecology* **78**, 253–269.
- 789 Kéfi, S., Berlow, E.L., Wieters, E.a., Navarrete, S.a., Petchey, O.L., Wood, S.a., Boit, A.,
790 Joppa, L.N., Lafferty, K.D., Williams, R.J., Martinez, N.D., Menge, B.a., Blanchette,
791 C.a., Iles, A.C. & Brose, U. (2012) More than a meal... integrating non-feeding inter-
792 actions into food webs. *Ecology Letters* pp. 291–300.

- 793 Kissling, W.D., Dormann, C.F., Groeneveld, J., Hickler, T., Kühn, I., McInerny, G.J.,
794 Montoya, J.M., Römermann, C., Schiffers, K., Schurr, F.M., Singer, A., Svenning,
795 J.C., Zimmermann, N.E. & O'Hara, R.B. (2012) Towards novel approaches to mod-
796 elling biotic interactions in multispecies assemblages at large spatial extents. *Journal*
797 *of Biogeography* **39**, 2163–2178.
- 798 Kopelke, J., Nyman, T., Cazelles, K., Gravel, D., Vissault, S. & Roslin, T. (2017) Food-
799 web structure of willow-galling sawflies and their natural enemies across europe. *Ecology*
800 **98**, 1730–1730.
- 801 Kopelke, J.P. (1999) Gallenerzeugende Blattwespen Europas - Taxonomische Grundlagen,
802 Biologie und Ökologie (Tenthredinidae: Nematinae: Euura, Phyllocolpa, Pontania).
803 *Courier Forschungsinstitut Senckenberg* **212**, 1–183.
- 804 Kopelke, J.P. (2003) Natural enemies of gall-forming sawflies on willows (*Salix* spp.).
805 *Entomologia Generalis* **26**, 277–312.
- 806 Kouki, J. (1999) Latitudinal gradients in species richness in northern areas: some excep-
807 tional patterns. *Ecological Bulletins* **47**, 30–37.
- 808 Kouki, J., Niemelä, P. & Viitasaari, M. (1994)) Reversed latitudinal gradient in species
809 richness of sawflies (Hymenoptera, Symphyta). *Annales Zoologici Fennici* **31**, 83–88.
- 810 Krebs, C. (2009) *Ecology: The Experimental Analysis of Distribution and Abundance*. 6th
811 ed. Pearson, San Francisco.
- 812 Laigle, I., Aubin, I., Digel, C., Brose, U., Boulangeat, I. & Gravel, D. (2018) Species
813 traits as drivers of food web structure. *Oikos* **127**, 316–326.
- 814 Laliberté, E. & Tylianakis, J.M. (2010) Deforestation homogenizes tropical parasitoid-
815 host networks. *Ecology* **91**, 1740–1747.

- 816 Legagneux, P., Gauthier, G., Lecomte, N., Schmidt, N.M., Reid, D., Cadieux, M.C.,
817 Berteaux, D., Béty, J., Krebs, C.J., Ims, R.a., Yoccoz, N.G., Morrison, R.I.G., Leroux,
818 S.J., Loreau, M. & Gravel, D. (2014) Arctic ecosystem structure and functioning shaped
819 by climate and herbivore body size. *Nature Climate Change* **E2168**, 1–5.
- 820 Legendre, P., Borcard, D. & Peres-Neto, P.R. (2005) Analyzing beta diversity: partition-
821 ing the spatial variation of composition data. *Ecological Monographs* **75**, 435–450.
- 822 Legendre, P. & De Cáceres, M. (2013) Beta diversity as the variance of community data:
823 dissimilarity coefficients and partitioning. *Ecology Letters* **16**, 951–963.
- 824 Leibold, M., Holyoak, M., Mouquet, N., Amarasekare, P., Chase, J.M., Hoopes, M.,
825 Holt, R.D., Shurin, J.B., Law, R., Tilman, D., Loreau, M. & Gonzalez, A. (2004) The
826 metacommunity concept : a framework for multi-scale community ecology. *Ecology*
827 *Letters* **7**, 601–613.
- 828 Leppänen, S., Malm, T., Värri, K. & Nyman, T. (2014) A comparative analysis of genetic
829 differentiation across six shared willow host species in leaf- and bud-galling sawflies.
830 *PLoS ONE* **9**, e116286.
- 831 Lurgi, M., López, B.C. & Montoya, J.M. (2012) Novel communities from climate change.
832 *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*
833 **367**, 2913–22.
- 834 Martinez, N.D., Hawkins, B.a., Dawah, H.A. & Feifarek, B.P. (1999) Effects of Sampling
835 Effort on Characterization of Food-Web Structure. *Ecology* **80**, 1044–1055.
- 836 McInerny, G.J. & Etienne, R.S. (2012a) Ditch the niche - is the niche a useful concept in
837 ecology or species distribution modelling? *Journal of Biogeography* **39**, 2096–2102.
- 838 McInerny, G.J. & Etienne, R.S. (2012b) Stitch the niche - a practical philosophy and
839 visual schematic for the niche concept. *Journal of Biogeography* **39**, 2103–2111.

- 840 McKinnon, L., Smith, P.A., Nol, E., Martin, J.L., Doyle, F.I., Abraham, K.F., Gilchrist,
841 H.G., Morrison, R.I.G. & Béty, J. (2010) Lower predation risk for migratory birds at
842 high latitudes. *Science* **327**, 326–327.
- 843 Morales-Castilla, I., Matias, M.G., Gravel, D. & Araújo, M.B. (2015) Inferring biotic
844 interactions from proxies. *Trends in Ecology & Evolution* **30**, 347–356.
- 845 Morueta-Holme, N., Blonder, B., Sandel, B., McGill, B.J., Peet, R.K., Ott, J.E., Violette,
846 C., Enquist, B.J., Jørgensen, P.M. & Svenning, J.C. (2016) A network approach for
847 inferring species associations from co-occurrence data. *Ecography* **39**, 1139–1150.
- 848 Nyman, T. & Julkunen-Tiitto, R. (2000) Manipulation of the phenolic chemistry of wil-
849 lows by gall-inducing sawflies. *Proceedings of the National Academy of Sciences* **97**,
850 13184–13187.
- 851 Nyman, T., Leppänen, S.A., Várkonyi, G., Shaw, M.R., Koivisto, R., Barstad, T.E.,
852 Vikberg, V. & Roininen, H. (2015) Determinants of parasitoid communities of willow-
853 galling sawflies: habitat overrides physiology, host plant and space. *Molecular Ecology*
854 **24**, 5059–5074.
- 855 Nyman, T., Vikberg, V., Smith, D. & BoevéJ-L (2010) How common is ecological speci-
856 ation in plant-feeding insects? A “Higher” Nematinae perspective. *BMC Evolutionary
857 Biology* **10**, e266.
- 858 Ovaskainen, O., Tikhonov, G., Norberg, A., Blanchet, F.G., Duan, L., Dunson, D., Roslin,
859 T., Abrego, N. & Chave, J. (2017) How to make more out of community data? a
860 conceptual framework and its implementation as models and software. *Ecology Letters*
861 **20**, 561–576.
- 862 Pellissier, L., Rohr, R., Ndiribe, C., Pradervand, J.N., Salamin, N., Guisan, A. & Wisz, M.

- 863 (2013) Combining food web and species distribution models for improved community
864 projections. *Ecology and Evolution* **3**, 4572–4583.
- 865 Petchey, O.L., Brose, U. & Rall, B.C. (2010) Predicting the effects of temperature on food
866 web connectance. *Philosophical transactions of the Royal Society of London. Series B,*
867 *Biological sciences* **365**, 2081–91.
- 868 Peters, R. (1991) *A critique for ecology*. Cambridge University Press, Cambridge.
- 869 Phillips, S.J., Anderson, R.P. & Schapire, R.E. (2006) Maximum entropy modeling of
870 species geographic distributions. *Ecological Modelling* **190**, 231 – 259.
- 871 Pillai, P., Gonzalez, A. & Loreau, M. (2011) Metacommunity theory explains the emer-
872 gence of food web complexity. *Proceedings of the National Academy of Sciences of the*
873 *United States of America* **108**, 19293–19298.
- 874 Poisot, T., Baiser, B., Dunne, J.A., Kéfi, S., Massol, F., Mouquet, N., Romanuk, T.N.,
875 Stouffer, D.B., Wood, S.A. & Gravel, D. (2016a) mangal – making ecological network
876 analysis simple. *Ecography* **39**, 384–390.
- 877 Poisot, T., Canard, E., Mouillot, D., Mouquet, N. & Gravel, D. (2012) The dissimilarity
878 of species interaction networks. *Ecology Letters* **15**, 1353–1361.
- 879 Poisot, T., Cirtwill, A.R., Cazelles, K., Gravel, D., Fortin, M.J. & Stouffer, D.B. (2016b)
880 The structure of probabilistic networks. *Methods in Ecology and Evolution* **7**, 303–312.
- 881 Poisot, T. & Gravel, D. (2014) When is an ecological network complex? connectance
882 drives degree distribution and emerging network properties. *PeerJ* **2**, e251.
- 883 Poisot, T., Gravel, D., Leroux, S., Wood, S.A., Fortin, M.J., Baiser, B., Cirtwill, A.R.,
884 Araújo, M.B. & Stouffer, D.B. (2016c) Synthetic datasets and community tools for the
885 rapid testing of ecological hypotheses. *Ecography* **39**, 402–408.

- 886 Poisot, T., Guéveneux-Julieu, C., Fortin, M.J., Gravel, D. & Legendre, P. (2018) Hosts,
887 parasites and their interactions respond to different climatic variables. *Global Ecology*
888 and Biogeography
- 889 Poisot, T., Stouffer, D.B. & Gravel, D. (2015) Beyond species: why ecological interactions
890 vary through space and time. *Oikos* **124**, 243–251.
- 891 Pollock, L.J., Tingley, R., Morris, W.K., Golding, N., O'Hara, R.B., Parris, K.M., Vesk,
892 P.A. & McCarthy, M.A. (2014) Understanding co-occurrence by modelling species si-
893 multaneously with a joint species distribution model (jsdm). *Methods in Ecology and*
894 *Evolution* **5**, 397–406.
- 895 Pulliam, H.R. (2000) On the relationship between niche and distribution. *Ecology Letters*
896 **3**, 349–361.
- 897 Roininen, H., Nyman, T. & Zinovjev, A. (2005) Biology, ecology, and evolution of gall-
898 inducing sawflies (Hymenoptera: Tenthredinidae and Xyelidae). *Biology, Ecology, and*
899 *Evolution of Gall-Inducing Arthropods* (eds. A. Raman, C.W. Schaefer & T.M. With-
900 ers), pp. 467–494, Science Publishers, Inc.
- 901 Scheiner, S.M. & Willig, M.R. (2007) A general theory of ecology. *Theoretical Ecology* **1**,
902 21–28.
- 903 Schleuning, M., Fründ, J., Klein, A.M., Abrahamczyk, S., Alarcón, R., Albrecht, M., An-
904 dersson, G., Bazarian, S., Böhning-Gaese, K., Bommarco, R., Dalsgaard, B., Dehling,
905 D., Gotlieb, A., Hagen, M., Hickler, T., Holzschuh, A., Kaiser-Bunbury, C., Kreft,
906 H., Morris, R., Sandel, B., Sutherland, W., Svenning, J.C., Tscharntke, T., Watts,
907 S., Weiner, C., Werner, M., Williams, N., Winqvist, C., Dormann, C. & Blüthgen,
908 N. (2012) Specialization of mutualistic interaction networks decreases toward tropical
909 latitudes. *Current Biology* **22**, 1925 – 1931.

- 910 Shurin, J.B., Clasen, J.L., Greig, H.S., Kratina, P. & Thompson, P.L. (2012) Warming
911 shifts top-down and bottom-up control of pond food web structure and function.
912 *Philosophical Transactions of the Royal Society of London B: Biological Sciences* **367**,
913 3008–3017.
- 914 Silvertown, J. (2004) Plant coexistence and the niche. *Trends in Ecology and Evolution*
915 **19**, 605–611.
- 916 Simanonok, M.P. & Burkle, L.A. (2014) Partitioning interaction turnover among alpine
917 pollination networks: spatial, temporal, and environmental patterns. *Ecosphere* **5**,
918 art149.
- 919 Skvortsov, A.K. (1999) Willows of Russia and adjacent countries. Taxonomical and geo-
920 graphical revision. *Univ. Joensuu Fac. Math. Natl. Sci. Rep. Ser* **39**, 1–307.
- 921 Soberón, J. (2007) Grinnellian and eltonian niches and geographic distributions of species.
922 *Ecology Letters* **10**, 1115–1123.
- 923 Staniczenko, P.P., Sivasubramaniam, P., Suttle, K.B., Pearson, R.G. & Lawler, J. (2017)
924 Linking macroecology and community ecology: refining predictions of species distribu-
925 tions using biotic interaction networks. *Ecology Letters* **20**, 693–707.
- 926 Stephens, C. & Heau, J. (2009) Using biotic interaction networks for prediction in biodi-
927 versity and emerging diseases. *PloS one* **4**, e5725.
- 928 Stouffer, D.B., Camacho, J., Jiang, W. & Amaral, L.a.N. (2007) Evidence for the existence
929 of a robust pattern of prey selection in food webs. *Proceedings of the Royal Society B:*
930 *Biological Sciences* **274**, 1931–1940.
- 931 Thuiller, W. (2003) Biomod – optimizing predictions of species distributions and project-
932 ing potential future shifts under global change. *Global Change Biology* **9**, 1353–1362.

- 933 Tilman, D. (1982) *Resource competition and community structure*. Princeton University
934 Press, Princeton.
- 935 Townsend, P.A., Soberón, S., Pearson, R.G., Anderson, R.P., Martínez-Meyer, E., Miguel,
936 N. & Araújo, M.B. (2011) *Ecological Niches and Geographic Distributions*. Princeton
937 University Press, Princeton.
- 938 Trøjelsgaard, K., Jordano, P., Carstensen, D.W. & Olesen, J.M. (2015) Geographical
939 variation in mutualistic networks: similarity, turnover and partner fidelity. *Proceedings*
940 *of the Royal Society of London B: Biological Sciences* **282**, 20142925.
- 941 Tylianakis, J.M., Tscharntke, T. & Lewis, O.T. (2007) Habitat modification alters the
942 structure of tropical host–parasitoid food webs. *Nature* **445**, 202–205.
- 943 Veech, J.A. (2013) A probabilistic model for analysing species co-occurrence. *Global Ecol-*
944 *ogy and Biogeography* **22**, 252–260.
- 945 Veijalainen, A., Sääksjärvi, I., Erwin, T., Gómez, I. & Longino, J. (2013) Subfamily
946 composition of Ichneumonidae (Hymenoptera) from western Amazonia: insights into
947 diversity of tropical parasitoid wasps. *Insect Conservation and Diversity* **6**, 28–37.
- 948 Vellend, M., Srivastava, D.S., Anderson, K.M., Brown, C.D., Jankowski, J.E., Kleyn-
949 hans, E.J., Kraft, N.J.B., Letaw, A.D., Macdonald, A.A.M., Maclean, J.E., Myers-
950 Smith, I.H., Norris, A.R. & Xue, X. (2014) Assessing the relative importance of neutral
951 stochasticity in ecological communities. *Oikos* **123**, 1420–1430.
- 952 Williams, R. & Martinez, N. (2000) Simple rules yield complex food webs. *Nature* **404**,
953 180–183.
- 954 Williams, R.J., Anandanadesan, A. & Purves, D. (2010) The probabilistic niche model
955 reveals the niche structure and role of body size in a complex food web. *PloS One* **5**,
956 e12092.

- 957 Wisz, M.S., Pottier, J., Kissling, W.D., Pellissier, L., Lenoir, J., Damgaard, C.F., Dor-
958 mann, C.F., Forchhammer, M.C., Grytnes, J.A., Guisan, A., Heikkinen, R.K., Høye,
959 T.T., Kühn, I., Luoto, M., Maiorano, L., Nilsson, M.C., Normand, S., Ockinger, E.,
960 Schmidt, N.M., Termansen, M., Timmermann, A., Wardle, D.a., Aastrup, P. & Sven-
961 ning, J.C. (2013) The role of biotic interactions in shaping distributions and realised
962 assemblages of species: implications for species distribution modelling. *Biological Re-*
963 *views* **88**, 15–30.
- 964 Wood, S.A., Russell, R., Hanson, D., Williams, R.J. & Dunne, J.A. (2015) Effects of
965 spatial scale of sampling on food web structure. *Ecology and Evolution* **5**, 3769–3782.
- 966 Woodward, G., Perkins, D.M. & Brown, L.E. (2010) Climate change and freshwater
967 ecosystems: impacts across multiple levels of organization. *Philosophical transactions*
968 *of the Royal Society of London. Series B, Biological sciences* **365**, 2093–106.
- 969 Wu, J., Nyman, T., Wang, D.C., Argus, G., Yang, Y.P. & Chen, J.H. (2015) Phylogeny
970 of Salix subgenus Salix s.l. (Salicaceae): delimitation, biogeography, and reticulate
971 evolution. *BMC Evolutionary Biology* **15**, e31.

Table 1: Example of data used to perform the analysis. Traditional species distribution models represent the occurrence of single species as a function of the environment, with models such as $X_i \text{ Temp}$ performed one species after the other. Some also tried to represent both the abiotic and biotic niche, with models such as $X_i \text{ Temp} + X_j$ (e.g. González-Salazar *et al.* 2013), while the new joint species distribution models (Ovaskainen *et al.*, 2017) consider the entire community at once ($\mathbf{X} \text{ Temp}$) and aims at representing co-distribution (X_{ij}). Here the focus is on the occurrence of interactions, L_{ij} . The full model (see model 3, Tables 1 and 2) considers the result of two submodels, one representing the occurrence of interactions given co-occurrence, $L_{ij} \text{ Temp}$ (evaluated only on the subset of data where $X_{ij} = 1$), and the other representing the co-occurrence, $X_{ij} \text{ Temp}$.

Location	Sp_i	Sp_j	X_i	X_j	X_{ij}	L_{ij}	Av. temp
Kühkopf	Phyllocolpa prussica	Chrysocharis elongata	1	1	1	1	8.25
Kühkopf	Phyllocolpa leucapsis	Chrysocharis elongata	0	1	0	0	8.25
Ötztal	Phyllocolpa prussica	Chrysocharis elongata	0	0	0	0	-1.50
Ötztal	Phyllocolpa leucapsis	Chrysocharis elongata	0	0	0	0	-1.50
Kilpisjärvi	Phyllocolpa prussica	Chrysocharis elongata	0	1	0	0	-1.92
Kilpisjärvi	Phyllocolpa leucapsis	Chrysocharis elongata	0	1	0	0	-1.92
Wüstensachsen	Phyllocolpa prussica	Chrysocharis elongata	1	1	1	0	5.58
Wüstensachsen	Phyllocolpa leucapsis	Chrysocharis elongata	1	1	1	1	5.58
Steinamoen	Phyllocolpa prussica	Chrysocharis elongata	0	1	0	0	3.17
Steinamoen	Phyllocolpa leucapsis	Chrysocharis elongata	0	1	0	0	3.17

Table 2: Summary of model comparison for the interaction between the leaf folder *Phyllocolpa prussica*) and the parasitoid *Chrysocharis elongata*

. Note that E is a short notation including both temperature and precipitation.

#	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 3: 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	$P(L_{ijy} X_{iy}, X_{jy}, E_y)$	$P(X_{iy})P(X_{jy})$	-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

972 **Figure legends**

973 **Figure 1**

974 **Non-random sampling of the metaweb.** Network assembly can be viewed as a sam-
975 pling process of the regional pool of potential interactions. Species (indicated by colored
976 nodes) are sampled first, and among the species found in the local network, only some in-
977 teractions (indicated by blue links) occur. We characterize these sampling processes with
978 the quantitative framework proposed in this paper. As a concrete illustration of metaweb
979 sampling, we here show a local interaction network among *Salix* (bottom/green), gallers
980 (center/blue), and parasitoids (top/red). The metaweb was constructed by aggregat-
981 ing interactions observed across 370 local networks. Nodes were distributed to minimize
982 crossing of interactions.

983 **Figure 2**

984 **Visual representation of the integrated niche.** In biogeography, the niche is con-
985 sidered the set of environmental conditions where the intrinsic growth rate r is positive
986 (Holt, 2009). The horizontal axis represents an environmental gradient impacting the
987 growth of a focal species (red point). The location of each species (grey points) along
988 this gradient represents their optimum, and the vertical dotted lines represent the limits
989 of the Grinnellian niche of the focal species. In food web ecology, the Eltonian niche
990 represents the location of a species in the food web, as determined by its niche position
991 (n) and its niche optimum (c). The vertical axis represents a niche gradient, for example
992 a trait such as body size. The location of each species along this gradient represents their
993 niche position. The focal species will feed only on host/prey species occupying niche lo-
994 cations within a given interval around the optimum, represented by the horizontal lines.
995 The integrated Grinnellian and Eltonian niche corresponds to the square in the middle

996 where an interaction is possible owing to a match of traits and spatial distribution. The
997 central square represents the area where the joint probability of observing co-occurrence
998 and interactions is positive.

999 **Figure 3**

1000 **Probabilistic representation of the interaction probability between a leaf folder**
1001 (*Phyllocolpa prussica*) and a parasitoid (*Chrysocharis elongata*) across gradi-
1002 ents of annual average temperature and annual precipitation. The colour gra-
1003 dient represents the probability of observing the event, from 0 (white) to 1 (black). The
1004 representation is based on predictions from Model 3 (see Table 1). In the left panel,
1005 open circles represent the absence of both species, whereas closed circles represent co-
1006 occurrence and plus signs the occurrence of only one of the two species. In the other
1007 two panels, open circles represent co-occurrence but an absence of interaction and closed
1008 circles the occurrence of an interaction.

1009 **Figure 4**

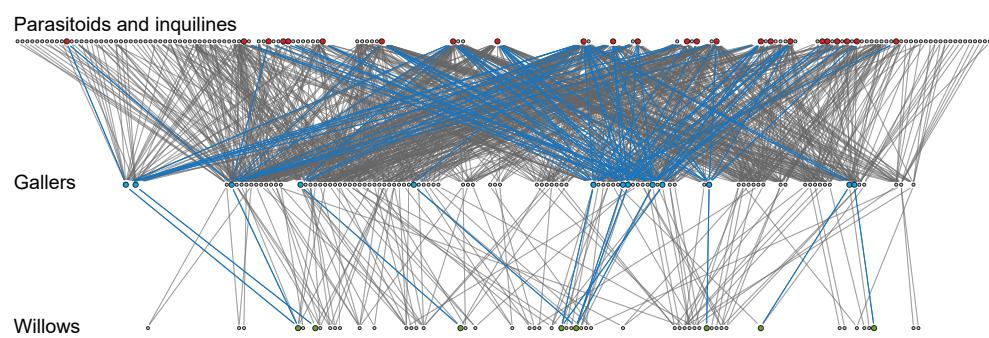
1010 **Probabilistic representation of the interaction probability between a leaf folder**
1011 (*Phyllocolpa prussica*) and a parasitoid (*Chrysocharis elongata*) across Eu-
1012 rope. Climate data are extracted from WorldClim over a 1km x 1km grid and predic-
1013 tions made according to the model illustrated at Fig. 3. Colour gradient ranges from low
1014 values (dark blue) to high values (yellow).

1015 **Figure 5**

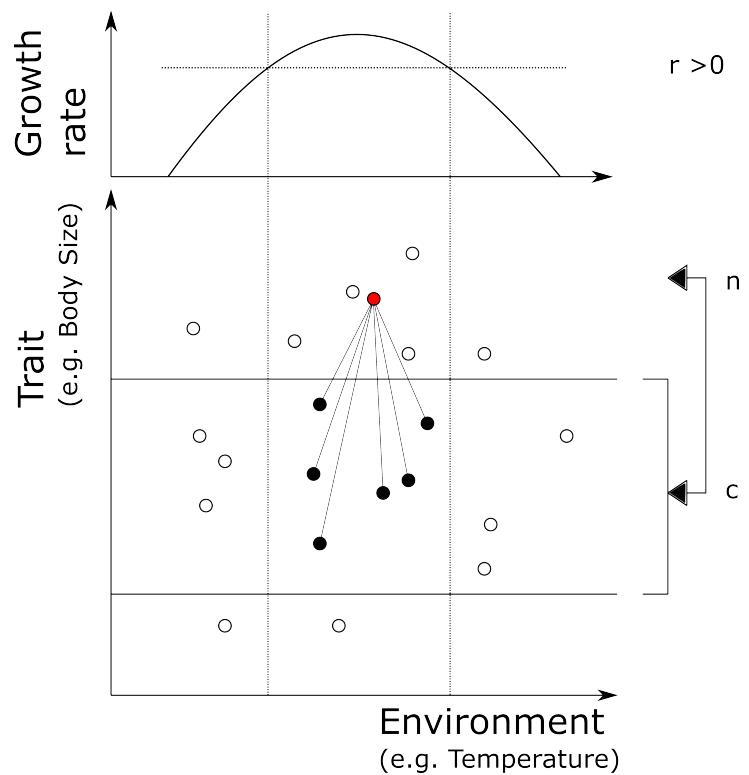
1016 **Mapping the distribution of species richness, the number of links and con-**
1017 **nnectance across Europe.** The representation is based on predictions from Model 3 (see
1018 Table 2) using climate data extracted from WorldClim over a 1km x 1km grid. Species

richness is obtained by summation of individual occurrence probabilities, and link density by summation of interaction probabilities. Color gradient ranges from low values (dark blue) to high values (yellow).

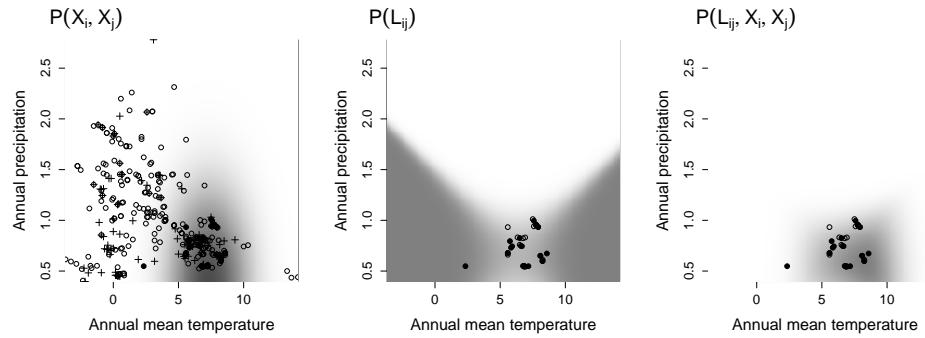
1022 **Figure 1**



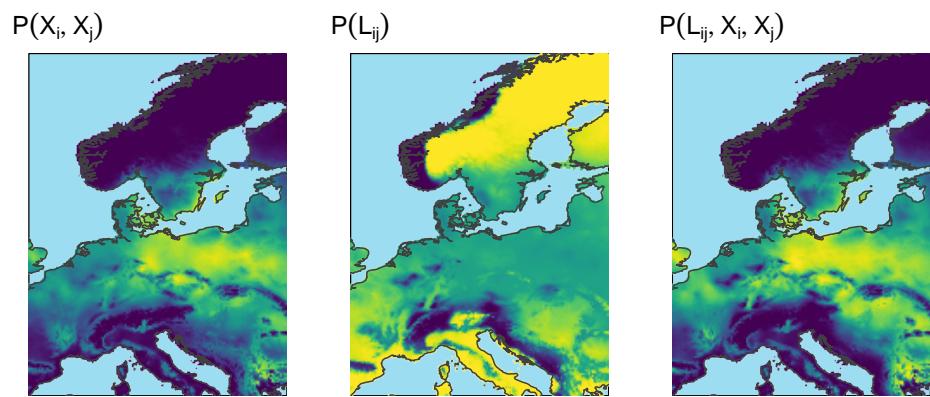
₁₀₂₃ Figure 2



1024 **Figure 3**



₁₀₂₅ **Figure 4**



₁₀₂₆ **Figure 5**

