

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

3
4 **Authors:** Dominique Gravel^{1,2,*}, Benjamin Baiser³, Jennifer A. Dunne⁴, Jens-Peter
5 Kopelke⁵, Neo D. Martinez⁶, Tommi Nyman⁷, Timothée Poisot^{2,8}, Daniel B. Stouffer⁹,
6 Jason M. Tylianakis^{9,10}, Spencer A. Wood¹¹, Tomas Roslin¹²

7
8 **1:** Canada Research Chair in Integrative Ecology. Département de biologie, Univer-
9 sité de Sherbrooke, 2500 Boulevard l'Université, Sherbrooke (Québec). J1K 2R1

10 **2:** Québec Centre for Biodiversity Science

11 **3:** University of Florida, Department of Wildlife Ecology and Conservation, 110 Newins-
12 Ziegler Hall, PO Box 110430, Gainesville, Fl. 32611-0430

13 **4:** Santa Fe Institute, 1399 Hyde Park Road, Santa Fe, NM 87501, USA.

14 **5:** Senckenberg Research Institute and Natural History Museum, Senckenberganlage 25,
15 D-60325 Frankfurt am Main, Germany

16 **6:** Ecology and Evolutionary Biology Department, University of Arizona, P.O. Box
17 210088, Tucson, AZ 85721, USA

18 **7:** University of Eastern Finland, Department of Environmental and Biological Sciences,
19 P.O. Box 111, FI-80101 Joensuu, Finland

20 **8:** Université de Montréal, Département des Sciences Biologiques, 90 Avenue Vincent
21 d'Indy, Montréal, QC H2V3S9, Canada.

22 **9:** Centre for Integrative Ecology, School of Biological Sciences, University of Canterbury,
23 Private Bag 4800, Christchurch, New Zealand

24 **10:** Department of Life Sciences, Imperial College London, Silwood Park Campus, Buck-
25 hurst Road, Ascot, Berkshire SL5 7PY, United Kingdom

26 **11:** University of Washington, School of Environmental and Forest Sciences, Box 352100,
27 Seattle, WA 98195, USA

28 **12:** Department of Ecology, Swedish University of Agricultural Sciences, Box 7044, 750
29 07 Uppsala, Sweden

30
31 **Keywords:** networks, spatial ecology, co-occurrence, probability of interaction

32 **Words in the abstract:** 226

33 **Words in the main text:** 7043

34 **Words in the legends:** 502

35
36 Authorship: DG, BB, JAD, NDM, TP, SAW, DBS and JMT contributed to the
37 conceptual framework; JPK, TN and TR contributed the data; DG developed the prob-
38 abilistic approach and performed the analysis; DG, TP and TR wrote the first draft and
39 all of the authors edited the manuscript.

40 **Abstract**

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

59 Introduction

60 Community ecology is *the study of the interactions that determine the distribution and*
61 *abundance of organisms* (Krebs, 2009). Despite a general consensus on this definition
62 (Scheiner & Willig, 2007), research on variation in community structure has mostly fo-
63 cused on the spatial and temporal turnover of species composition (Anderson *et al.*, 2011),
64 neglecting variation in the way species interact with each other despite accumulating em-
65 pirical evidence that this is a major source of diversity (Poisot *et al.*, 2015b). Given this
66 omission, it is perhaps not surprising that biogeographers are still struggling to establish
67 whether interactions actually impact the distribution of species at large spatial scales
68 (Wisz *et al.*, 2012; Kissling *et al.*, 2012): treating interactions as fixed events neglects a
69 large part of the complexity of empirical communities, and will most likely deliver under-
70 whelming results. Recent attempts at accounting for interactions in species distribution
71 models (Pollock *et al.*, 2014; Pellissier *et al.*, 2013) have brought some methodological
72 advances, but are not sufficient for two reasons. First, these techniques are still based
73 on a ‘species-based’ approach to communities, where interactions are merely treated as
74 fixed covariates affecting distribution. Second, they failed to provide a conceptual step
75 forward, both in their treatment of interactions and in the quality of the predictions they
76 make.

77 Network approaches offer a convenient representation of communities because they
78 simultaneously account for species composition and their interactions. Species are repre-
79 sented as nodes, so that networks already encompass all the information used by current
80 approaches; in addition, interactions are represented by links, so that networks provide
81 additional, higher-order information on community structure. To date, studies of net-
82 work diversity have mostly been concerned with the distribution of interactions within
83 locations, and less so with variation among locations (Dunne, 2006; Bascompte & Jor-
84 dano, 2007; Ings *et al.*, 2009; Kéfi *et al.*, 2012). There is, however, ample evidence that

85 interaction networks vary in space and time (Laliberté & Tylianakis, 2010; Poisot *et al.*,
86 2012; Albouy *et al.*, 2014; Poisot *et al.*, 2016b; Trøjelsgaard *et al.*, 2015), even though
87 there is no common framework with which to generalize these results. Metacommunity
88 theory provides explanations for variation in the distribution of the nodes (Gravel *et al.*,
89 2011; Pillai *et al.*, 2011), but there is no such explanation to the variation of node and
90 link occurrences. Consequently, we urgently need a conceptual framework to formalize
91 these observations, as it is the only way towards fulfilling the goal of community ecol-
92 ogy: providing cogent predictions about, and understanding of, the structure of ecological
93 communities.

94 Given the historically different approaches to modelling the distributions of species vs.
95 interactions, there is a clear need to bring the two together. Here, we offer an integrated
96 approach to do so, adopting the view that community structure is best represented as
97 a network of ecological interactions. Based on this idea, we propose a new description
98 of the basic concept of the ecological niche that integrates the effect of the environment
99 on species distribution and on the ecological interactions among them. Building on this
100 concept, we develop a quantitative theory to explain turnover of interactions in space and
101 time. We first present the conceptual framework, and then formalize it mathematically,
102 using a probabilistic model to represent the sampling of the regional pool of interactions.
103 At the level of species pairs, the statistical approach could be conceived as an interaction
104 distribution model. At the community level, the approach provides a likelihood-based
105 method to compare different hypotheses of network turnover. As an illustrative example,
106 we apply this novel framework to a large data set on host–parasite interactions across
107 Europe and find that two aspects of the environment (temperature and precipitation)
108 exert a strong imprint on species co-occurrence, but not on species interactions. The
109 network structure changes systematically across the latitudinal gradient, with a peak of
110 connectance at intermediate latitudes.

111 **The two dimensions of community structure**

112 The problem of community assembly is often formulated as *how are species sampled from*
113 *a regional pool to constitute a local community (Götzenberger et al., 2012)?* This question
114 could be rewritten to address the problem of network assembly, as *how do samples from*
115 *a regional pool of interactions constitute a local interaction network?* An illustration of
116 this problem for a food web is provided in Fig. 1. The regional pool of interactions,
117 the *metaweb*, represents potential interactions among all species that could be found in a
118 given area. In this particular case, there are 275 nodes, and 1173 links among the plants
119 (52 nodes), herbivores (96 nodes), and parasitoids (127 nodes) from Northern Europe. An
120 instance of a local community is also illustrated, with 45 nodes and 93 interactions. Only
121 55.0% of all potential interactions (plant-herbivore or herbivore parasitoid combinations)
122 are realized locally, revealing the stochastic nature of ecological interactions. Our objec-
123 tive here is to provide a conceptual framework to explain the sampling of the regional
124 pool of interactions, along with a quantitative method to predict it. The problem could
125 be formalized sequentially by understanding first why only a fraction of the species co-
126 occur locally and second why these species do or do not interact.

127 There are multiple causes of spatial turnover of species co-occurrence. The first and
128 most-studied driver is the effect of variation in the abiotic environment on species perfor-
129 mance. Combined with specific responses in demography, it generates variation among
130 sites by selecting the locally fittest species (Leibold *et al.*, 2004). Stochasticity plays
131 an additional role, either because colonization and extinction events (Hanski, 1999) are
132 inherently unpredictable or because strong non-linear feedbacks in community dynamics
133 generate alternative transients and equilibria (Chase, 2007; Vellend *et al.*, 2014). Analyses
134 of community turnover are usually performed with data represented in a table with rows
135 corresponding to sites (or measurements) and columns to species. Metrics of beta diver-
136 sity quantify the variance of this community data (Legendre *et al.*, 2005). Traditional

137 approaches rely on measures of dissimilarity among communities, such as the Jaccard or
138 Bray–Curtis indices. More recent approaches decompose total variation of the commu-
139 nity data into species and site contributions to beta diversity (Legendre & De Cáceres,
140 2013), and further partition it into dissimilarity due to changes in species richness and
141 dissimilarity due to actual species turnover (Baselga, 2010). Even though these methods
142 compare whole lists of species among sites or measurements, they remain fundamentally
143 ”species-based”, since they report variation within columns. None of them explicitly
144 considers variation of associations (i.e., of pairs or higher-order motifs – Stouffer *et al.*
145 2007).

146 Similarly, we are now getting a better understanding of interaction turnover. As men-
147 tioned above, in the network approach to community structure, species and interactions
148 are represented by nodes and links, respectively. Associations can also be represented
149 by matrices in which entries represent the occurrence or intensity of interactions among
150 species (rows and columns). Network complexity is then computed as the number of
151 interactions (in the case of binary networks) or interaction diversity (in the case of quan-
152 titative networks, Bersier *et al.* 2002). Variability in community structure consequently
153 arises from the turnover of species composition, along with turnover of interactions among
154 pairs of species. The occurrence and intensity of interactions could vary because of the
155 environment, species abundance, and higher-order interactions (Poisot *et al.*, 2015b).
156 Variation in community composition can be independent of variation of ecological inter-
157 actions, suggesting that species and interaction distribution may well respond to different
158 drivers (Poisot *et al.*, 2012).

159 The ”niche” is by far the dominant concept invoked to explain species distributions
160 and community assembly, from the local to the global scale. Following Hutchinson 1957,
161 the niche is viewed as the set of environmental conditions allowing a population to estab-
162 lish and persist (see also Holt 2009). Community turnover arises as a result of successive

163 replacement of species along an environmental gradient, in agreement with the Gleasonian
164 view of communities (Gleason, 1926). The concept is straightforward to put into prac-
165 tice with species distribution models, as it maps naturally on available distributional and
166 environmental data. Consequently, a vast array of statistical tools have been developed
167 to implement it (e.g. BIOMOD Thuiller 2003, MaxEnt Phillips *et al.* 2006). It is how-
168 ever much harder to account for ecological interactions within this approach (Townsend
169 *et al.*, 2011). As such, these interactions are often viewed as externalities constraining
170 or expanding the range of environmental conditions required for a species to maintain a
171 viable population (Pulliam, 2000; Soberón, 2007).

172 Interestingly, the ecological network literature also has its own "niche model" to po-
173 sition a species in a community (Williams & Martinez, 2000). The niche of a species in
174 this context represents the multidimensional space of all of its interactions. Each species
175 is characterized by a niche position, an optimum and a range over three to five different
176 niche axes (Williams & Martinez, 2000; Eklöf *et al.*, 2013). The niche model of food
177 web structure and its variants have successfully explained the complexity of a variety
178 of networks, from food webs to plant–pollinator systems (Allesina *et al.*, 2008; Williams
179 *et al.*, 2010; Eklöf *et al.*, 2013). This conceptual framework is, however, limited to local
180 communities, and does not provide any explanation for the turnover of network structure
181 along environmental gradients.

182 **The integrated niche**

183 Despite several attempts to update the concept of the ecological niche, ecologists have not
184 moved far beyond the "n-dimensional hypervolume" defined by Hutchinson. Despite its
185 intuitive interpretation and easy translation into species distribution models (Boulangeat
186 *et al.*, 2012; Blonder *et al.*, 2014), the concept has been frequently criticized (Hardin, 1960;

187 Peters, 1991; Silvertown, 2004), and several attempts have been made to expand and
188 improve it (Pulliam, 2000; Chase & Leibold, 2003; Soberón, 2007; Holt, 2009; McNerny
189 & Etienne, 2012b).

190 Part of the problem surrounding the niche concept has been clarified with the dis-
191 tinction between Eltonian and Grinnellian definitions (Chase & Leibold, 2003). The
192 Grinnellian dimension of the niche is the set of environmental conditions required for a
193 species to maintain a population in a location. The Grinnellian niche is intuitive to ap-
194 ply, and constitutes the conceptual backbone of species distribution models. The Eltonian
195 niche, on the other hand, is the effect of a species on its environment. While this aspect
196 of the niche is well known by community ecologists, it is trickier to turn into predictive
197 models. Nonetheless, the development of the niche model of food web structure (Williams
198 & Martinez, 2000) and its parameterization using functional traits (Gravel *et al.*, 2013;
199 Bartomeus *et al.*, 2016) made it more operational.

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

213 els that also consider ecological interactions. Thus, a more integrative description of the
214 niche will be key to understand spatial and temporal turnover in community structure.

215 We propose to integrate the two perspectives of the niche using a visual representa-
216 tion of both components (Fig. 2). The underlying rationale is that, in addition to the
217 environmental constraints on demographic performance (Fig.2 top panel), any organism
218 requires resources to meet its metabolic demands and to sustain reproduction (Fig. 2
219 nodes in network of bottom panel). Abiotic environmental axes are any non-consumable
220 factors affecting the demographic performance of an organism. Alternatively, the resource
221 axes are traits of the resources that allow interactions with the consumer. The niche can
222 therefore be viewed as the set of abiotic environmental variables (the Grinnellian compo-
223 nent) along with the set of traits (the Eltonian component) that allow a population to
224 establish and to persist at a location. Accordingly, each species can be characterized by
225 an optimal position along both the environmental (x-axis) and the trait (y-axis) plane.
226 The integrated niche is then the hypervolume where interactions can occur and sustain
227 a population.

228 This approach radically changes the representation of the niche, putting species dis-
229 tributions and ecological interactions into the same formalism. Moreover, it allows the
230 limits of the niche axes to be independent of each other (as in the example in Fig. 2),
231 or to interact. For instance, the optimal prey size for predatory fishes could decline with
232 increasing temperature (Gibert & DeLong, 2014), which would make diet boundaries
233 functions of the environment. Alternatively, we could also consider that the growth rate
234 of the predator changes with the size of its prey items, thereby altering the environmental
235 boundaries.

236 A probabilistic representation of interaction networks 237 in space

238 We now formalize the integrated niche with a probabilistic approach to interactions and
239 distributions. In particular, we seek to represent the probability that an interaction
240 between species i and j occurs at location y . We define L_{ijy} as a stochastic variable, and
241 focus on the probability that this event occurs, $P(L_{ijy})$. We note that the occurrence of
242 an interaction is dependent on the co-occurrence of species i and j . This argument might
243 seem trivial at first, but the explicit consideration of this condition in the probabilistic
244 representation of ecological interactions will prove instrumental to understanding their
245 variation. We define X_{iy} as a stochastic variable representing the occurrence of species
246 i at location y . The quantity we seek to understand is the probability of a joint event,
247 conditional on the set of environmental conditions E_y :

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

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

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

251 The first term on the right-hand side of the equation is the probability of observing
252 the two species co-occurring at location y . It corresponds to the Grinnellian dimension of
253 the niche. The second term represents the probability that an interaction occurs between
254 species i and j , given that they are co-occurring. This predicate can be refined using
255 information on trait distribution and trait matching rules ((Bartomeus *et al.*, 2016)).

256 Above, we referred to this entity as the "metaweb" and it corresponds to the Eltonian
257 dimension of the niche. Below, we will see how this formalism can be directly fit to
258 empirical data. But before turning to an application, we will discuss the interpretation
259 of different variants of these two terms.

260 Variants of co-occurrence

261 There are several variants to the co-occurrence probability, representing different hy-
262 potheses concerning temporal and spatial variation in network structure (see the explicit
263 formulations in Table 1). The simplest model relates the probability of co-occurrence
264 directly to the environment, $P(X_{iy}, X_{jy}|E_y)$. In this situation, there are no underlying
265 assumptions about the ecological processes responsible for co-occurrence. It could arise
266 because interactions constrain distribution (Pollock *et al.*, 2014; Cazelles *et al.*, 2016) or,
267 alternatively, because of environmental requirements shared between i and j . In the for-
268 mer case, species are not independent of each other and the conditional occurrence must
269 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 latter
270 case, species are independent, and only the marginal occurrence must be accounted for,
271 $P(X_{ijy}|E_y) = P(X_{iy}|E_y)P(X_{jy}|E_y)$.

272 The co-occurrence probability itself could depend on ecological interactions. This
273 should be viewed as the realized component of the niche (i.e. the distribution when
274 accounting for species interactions). Direct pairwise interactions such as competition,
275 facilitation, and predation have long been studied for their impact on co-distribution
276 (e.g. Diamond 1975; Connor & Simberloff 1979; Gotelli 2000). Second- and higher-order
277 interactions (e.g. trophic cascades) could also affect co-occurrence. Co-occurrence of
278 multiple species embedded in ecological networks is a topic of its own, however, and is
279 influenced by both network topology and species richness (Cazelles *et al.*, 2016). Not
280 only direct interactions influence co-occurrence, but indirect interactions do as well (e.g.

281 plant species sharing an herbivore, or herbivores sharing parasitoids, could repel each
282 other in space Holt & Lawton 1993). The impact of direct interactions and first-order
283 indirect interactions on co-occurrence tends to vanish with increasing species richness in
284 the community. Further, co-occurrence is also influenced by the covariance of interacting
285 species to an environmental gradient (Cazelles *et al.*, 2015). Because of the complexity
286 of relating co-occurrence to the structure of interaction networks, we will focus here on
287 the variation of interactions and not on their distribution, and leave this specific issue for
288 the Perspectives section and future research.

289 Variants of the metaweb

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

298 Second, ecological interactions could also vary with the environment, so that $P(L_{ijy}|E_y)$.
299 Although it is rare to see a conditional representation of pairwise ecological interactions,
300 experimental studies have frequently revealed interactions to be sensitive to the environ-
301 ment. For instance, (McKinnon *et al.*, 2010) showed that predation risks of shorebirds
302 vary at the continental scale, decreasing from the south to the north. It is also common
303 to see increasing top-down control with temperature (e.g. Shurin *et al.* 2012; Gray *et al.*
304 2015). Effects of the environment on interactions also propagate up the community and
305 influence network structure (Tylianakis *et al.*, 2007; Woodward *et al.*, 2010; Petchey *et al.*,

306 2010).

307 **Application: continental-scale variation of host-parasite** 308 **community structure**

309 We now turn to an illustration of our framework with the analysis of an empirical dataset
310 of host–parasite networks sampled throughout the south–north environmental gradient
311 in continental Europe. The focal system consists of local food webs of willows (genus
312 *Salix*), their galling insects, and the natural enemies (parasitoids and inquilines) of these
313 gallers. Targeting this system, we ask: i) how much does network structure vary across
314 the gradient, and ii) what is the primary driver of network turnover across the gradient?

315 **Data**

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

329 ing (Kopelke, 2003). In terms of associations between the trophic levels, phylogeny-based
330 comparative studies have demonstrated that galls represent "extended phenotypes" of
331 the gallers, meaning that gall form, location, and chemistry is determined mainly by the
332 galling insects and not by their host plants (Nyman & Julkunen-Tiitto, 2000). Because
333 galler parasitoids have to penetrate a protective wall of modified plant tissue in order to
334 gain access to their victims, gall morphology has been inferred to strongly affect the asso-
335 ciations between parasitoids and hosts (Nyman *et al.*, 2007). Thus, the set of parasitoids
336 attacking each host is presumably constrained by the form, size, and thickness of its gall.

337 Local realizations of the willow–galler–parasitoid network were reconstructed from
338 community samples collected between 1982 and 2010. During this period, willow galls
339 were collected at 370 sites across Central and Northern Europe. Sampling was conducted
340 in the summer months of June and/or July, i.e., during the later stages of larval de-
341 velopment. Galler species were identified on the basis of willow host species and gall
342 morphology, as these are distinct for each sawfly species. At each site, galls were ran-
343 domly collected from numerous willow individuals in an area of about 0.1–0.3 km^2 . Some
344 sites were visited more than once, with a total of 641 site visits across the 370 sites.
345 GPS coordinates were recorded for each location; for our analyses, current annual mean
346 temperature and precipitation were obtained from WorldClim using the R package raster
347 (Hijmans, 2015). While other covariates could have also been considered, these two vari-
348 ables are likely representative of the most important axes of the European climate, and are
349 also more easily interpretable than reduced variables obtained, for example, by principal
350 component analysis.

351 The methods used for rearing parasitoids from the galls have been previously de-
352 scribed by Kopelke 2003. In brief, galls were opened to score the presence of galler
353 or parasitoid/inquiline larvae. Parasitoid larvae were classified to preliminary morphos-
354 pecies, and the identity of each morphospecies was determined by connecting them to

355 adults emerging after hibernation. The galls were reared by storing single galls in small
356 glass tubes (Kopelke, 1985). Hibernation of galls containing parasitoids took place either
357 within the glass tubes or between blotting paper in flowerpots filled with clay granulate or
358 a mixture of peat dust and sand. These pots were stored over the winter in a roof garden
359 and/or in a climatic chamber. In most cases, the matching of larval morphospecies with
360 adult individuals emerging from the rearings allowed the identification of the parasitoids
361 to the species level. Nonetheless, in some cases, individuals could only be identified to
362 one of the (super)families Braconidae, Ichneumonidae, and Chalcidoidea. This was par-
363 ticularly the case when only remains of faeces, vacant cocoons of parasitoids, and/or
364 dead host larvae were found, as was the case when parasitoids had already emerged from
365 the gall. As a result, the largest taxon in the data set, "Chalcidoidea indeterminate",
366 represents a superfamily of very small parasitoids that are hard to distinguish.

367 In total, 146,622 galls from 52 *Salix* taxa were collected for dissection and rearing.
368 These galls represented 96 galler species, and yielded 42,133 individually-identified par-
369 asitoids. Of these, 25,170 (60%) could be identified to the species level. Overall, 127
370 parasitoid and inquiline taxa were distinguished in the material. Data on host associa-
371 tions within subsets of this material have been previously reported by (Kopelke, 1999)
372 and (Nyman *et al.*, 2007). The current study represents the first analysis of the full data
373 set from a spatial perspective.

374 **Analysis**

375 Computing the probability of observing an interaction involves fitting a set of binomial
376 models and collecting their estimated probabilities. For the sake of illustration, we con-
377 sidered second-order generalized linear models – although more flexible fitting algorithms
378 (e.g. GAM or Random Forest) could equally well be used, as long as the algorithm
379 can estimate the probability for each observation. The data consist of a simple (albeit

380 large and full of zeros) table with the observation of each species, X_{iy} and X_{jy} , their
381 co-occurrence, X_{ijy} , the observation of an interaction L_{ijy} , and environmental co-variates
382 E_y . Thus, there is one row per pair of species per site. We considered that an absence
383 of a record of an interaction between co-occurring species at a site means a true absence
384 (see below for a discussion on this issue).

385 We compared three models for the co-occurrence probability. The first one directly
386 models the co-occurrence probability conditional on the local environment, $P(X_{iy}, X_{jy}|E_y)$
387 (models are listed at Table 1 and 2). Hence, this model makes no assumptions about the
388 mechanisms driving co- occurrence for any given environment, and instead uses the in-
389 formation directly available in the data. It thereby indirectly accounts for the effect of
390 interactions on co-occurrence, if there are any. The second model considers independent
391 occurrence of species. In this case, we independently fit $P(X_{iy}|E_y)$ and $P(X_{jy}|E_y)$, and
392 we then take their product to derive the probability of co-occurrence. This model should
393 be viewed as a null hypothesis with respect to the first model, since a comparison be-
394 tween the respective models will reveal if there is significant spatial association of the two
395 species beyond a joint response to the shared environment (Cazelles *et al.*, 2016). Finally,
396 the third model assumes that the probability of co-occurrence is independent of the en-
397 vironment and thus constant throughout the landscape. In other words, $P(X_{iy}, X_{jy})$ is
398 obtained by simply counting the number of observed co-occurrences divided by the total
399 number of observations. Thus, the comparison between the first and third model allows
400 us to test the hypothesis that co-occurrence is conditional on the environment. Whenever
401 the environment was included as a covariate in the GLM, we considered a second-order
402 polynomial response for both temperature and precipitation in order to account for op-
403 tima in environmental conditions. There are consequently five parameters for the first
404 model when fitting a given pair of species, 10 parameters for the second, and only one for
405 the third model.

406 Following the same logic, we compared three models of the interaction probability. The
407 first model conditions the interaction probability on the local environmental variables,
408 $P(L_{ijy}|X_{iy}, X_{jy}, E_y)$. Consequently, the model was fit to the subset of the data where the
409 two species co-occur. The second model fits the interaction probability independently of
410 the local environmental variables, $P(L_{ijy}|X_{iy}, X_{jy})$. It corresponds to the number of times
411 the two species were observed to interact when co-occurring, divided by the number of
412 times that they co-occurred. The third model is an extreme case performed only to test
413 the hypothesis that if two species are found to interact at least once, then they should
414 interact whenever they co-occur, $P(L_{ijy}|X_{iy}, X_{jy}) = 1$. While not necessarily realistic,
415 this model tests an assumption commonly invoked in the representation of local networks
416 from the knowledge of a deterministic metaweb. There are consequently five parameters
417 for the first model, a single parameter for the second model and no parameter to evaluate
418 for the third model (where the interaction probability is fixed by the hypothesis).

419 We fit the different models to each pair of species and recorded the predicted prob-
420 abilities. The joint probability $P(L_{ijy}, X_{iy}, X_{jy})$ was then computed from Eq. 2, and
421 the likelihood of each observation was computed as $\mathcal{L}(\theta_{ijy}|D_{ijy}) = P(L_{ijy}, X_{iy}, X_{jy})$ if an
422 interaction was observed, and as $\mathcal{L}(\theta_{ijy}|D_{ijy}) = 1 - P(L_{ijy}, X_{iy}, X_{jy})$ if no interaction
423 was observed. The log-likelihood was summed over the entire dataset to compare the
424 different models by AIC. Not surprisingly, there was a very large number of species pairs
425 for which this model could not be computed, as they simply never co-occurred. For these
426 pairs, we have no information of the interaction probability, and they were consequently
427 removed from the analysis. The log-likelihood reported across the entire dataset was
428 summed over all pairs of species observed to co- occur at least once. Interactions between
429 the first (Salix) and second (gallers) trophic layers and those between the second and
430 third (parasitoids) were considered separately. Finally, we used the full model (in which
431 both co-occurrence and the interaction are conditional on the environment) to interpolate

432 species distributions and interaction probabilities across the entire European continent.
433 We reconstructed the expected network for each location in a 1 X 1 km grid and computed
434 the probabilistic connectance following (Poisot *et al.*, 2016a).

435 All of the data are openly available in the database *mangal* (Poisot *et al.*, 2015a) and
436 all R scripts for running the analysis, are provided in the Supplementary Material.

437 Results

438 Despite the extensive sampling, many pairs of species were observed to co- occur only a
439 few times. This made it difficult to evaluate interaction probabilities with any reasonable
440 confidence interval. Thus, we start with an example of a single pair of species selected
441 because of its high number of co- occurrences ($N_{ij} = 38$): the leaf folder *Phyllocolpa*
442 *prussica* and the parasitoid *Chrysocharis elongata*. These two fairly abundant species
443 were observed $N_i = 49$ and $N_j = 121$ times, respectively, across the 370 sites, and
444 they were found to interact with a marginal probability $P(L_{ij}) = 0.55$, which means
445 they interacted at 21 different locations. Here, a comparison of model fit (Table 1)
446 reveals that conditioning the interaction probability on local environmental conditions
447 adds no explanatory power beyond a model assuming the same probability of interaction
448 anywhere in space (Model 1 vs Model 2). Moreover, when the two species co-occur, the
449 occurrence of the interaction was insensitive to the environment (Model 2 vs Model 3).
450 Alternatively, climatic variables significantly impacted co-occurrence (Model 3 vs Model
451 4). The neutral model performed worse than the non-random co- occurrence model
452 (Model 3 vs Model 6). The full model revealed that the greatest interaction probability
453 occurred at intermediate temperature and precipitation, simply because this is where
454 the two species most frequently co-occur (Fig. 3). The probabilities of co-occurrence
455 and interaction can be represented in space, where we found that the highest interaction
456 probability occurred in Central Europe (Fig. 4).

457 We evaluated each model for all pairs of species in order to better understand the large-
458 scale drivers of network turnover. The results were highly consistent among trophic layers
459 (Salix–gallers and gallers–parasitoids; Table 2). Across all pairs of species, the conditional
460 representation of interactions performed better than the marginal one (Model 1 vs Model
461 2); that is, interactions did not occur systematically whenever the two species were found
462 co-occurring. Hence, in addition to species turnover, the stochastic nature of interactions
463 contributes to network variability. In total, we recorded 1,173 pairs of interactions, only
464 290 of which occurred more than five times. Out of these 290 interactions, 143 were
465 systematically detected whenever the two species co-occurred. In the instances when
466 species co- occurred, the two environmental variables considered proved relatively poor
467 predictors of their interactions (Model 2 vs Model 3). Not surprisingly, for both types of
468 interactions (Salix–galler and galler–parasitoid), the log- likelihood increased when the
469 environment was considered. However, the extra number of parameters exceeded the gain
470 in log-likelihood and inflated AIC. Therefore, the most parsimonious model excluded the
471 effect of the environment. On the basis of log-likelihood only, co-occurrence was non-
472 neutral for both Salix–galler and galler–parasitoid interactions. Thus, according to AIC,
473 the best model was the one of non-random co-occurrence (Model 3 vs Model 6) for both
474 types of interactions.

475 The approach we present not only has implications for understanding the biogeogra-
476 phy of pairwise interactions and interaction networks, but also for evaluating the quality
477 of metawebs. We investigated the reliability of the estimated metaweb across the entire
478 dataset with summary statistics of species co-occurrence. As mentioned above, across
479 the 17,184 potential pairs of species, only 1,173 pairs interacted in at least a single lo-
480 cation, yielding a connectance of 0.068. However, only 4,459 pairs of species were found
481 co- occurring at least once across all locations. There are consequently 12,725 gaps of
482 information in the metaweb (74.1% - see Fig. 5). As we cannot know whether the non-

483 co-occurring species would indeed interact if found together, a more appropriate estimate
484 of connectance would be $C = 1173/4459 = 0.263$. This result reveals that the evaluation
485 of the sampling quality of ecological networks is a problem on its own and well worth
486 further attention.

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

497 Interpretation

498 We have proposed that the representation of community structure and its variation in
499 space and time is best captured by the formalism of ecological networks, as both the
500 distribution of species and their interspecific interactions can then be accounted for. We
501 consequently revised the niche concept in order to integrate its abiotic and biotic com-
502 ponents that vary over time and space. This integrated niche was represented visually
503 with an ordination of species into an environmental space and a trait space. The fun-
504 damental niche of a species is represented as the set of environmental conditions and
505 resources that allow a species to establish in a location, thereby integrating the Eltonian
506 and the Grinnellian components of the niche. We then translated the concept mathe-

507 matically by investigating the probability of the joint occurrences of species and their
508 interaction, which should be interpreted as an interaction distribution model. We used
509 this approach to characterize the turnover of the structure of ecological interactions in a
510 species-rich tri-trophic network across Western Europe, finding that the primary driver
511 of network variation is the turnover in species composition. To our knowledge, this is the
512 first continental-wide analysis of the drivers of network structure from empirical data on
513 the occurrence of interactions (see Baiser *et al.* 2012; Albouy *et al.* 2014; Poisot *et al.*
514 2016b).

515 Applying the framework to our large data set on host–parasite interactions across
516 Europe revealed key features in the interaction between *Salix* taxa, their herbivores,
517 and the natural enemies of these herbivores. Consistent with a general increase in the
518 diversity of *Salix* towards boreal areas (Cronk *et al.*, 2015), overall species richness of
519 the networks increased towards the north. The distribution of *Salix* species richness
520 largely matched those of gallers and parasitoids. These observations within Europe are
521 also matched by the ones found at a global scale for *Salix* (Argus, 1997; Cronk *et al.*,
522 2015; Wu *et al.*, 2015) and sawflies (Kouki *et al.*, 1994; Kouki, 1999). Species richness
523 in a common group of parasitic wasps, the Ichneumonidae, was originally presumed to
524 show a similar ”reversed latitudinal gradient”, but this observation has been recently
525 challenged by findings of rather high ichneumonid diversity in the tropics (Veijalainen
526 *et al.*, 2013). Nevertheless, ichneumonid subfamilies specifically associated with sawflies
527 (Ctenopelmatinae, Tryphoninae) are clearly less diverse in the south.

528 Exactly what processes are responsible for the distribution of species richness at dif-
529 ferent trophic levels is yet to be established (but see e.g. Roininen *et al.* 2005; Nyman
530 *et al.* 2010; Leppänen *et al.* 2014), but as a net outcome of different latitudinal trends
531 across trophic levels, the distribution of co- occurrence and therefore of potential in-
532 teractions differed between the first and second layers of feeding links. The correlation

533 between expected *Salix* and gallers richness was 0.73, while it was 0.58 between gallers
534 and their parasitoids. Therefore, the ratio of herbivores to *Salix* species is essentially
535 constant across Europe, while each herbivore species is potentially attacked by a and a
536 lower trophic level at the same site was clearly affected by the richer enemy community
537 at higher latitudes. Consequently, overall connectance peaks in Northern Europe (Fig.
538 6).

539 In terms of species interacting with each other, our analysis suggests that the environ-
540 ment leaves a detectable imprint on species co-occurrence, but only a slight mark on the
541 occurrence of realized links among species in a specific place: the probability of finding
542 a given combination of species at a higher and a lower trophic level at the same site was
543 clearly affected by the environment, whereas the probability of observing an interaction
544 between the two was not detectably so. This applies to the example species *Phyllocolpa*
545 *prussica* and *Chrysocharis elongata* (Figs 2 and 3), but also to all species pairs more
546 generally. For the example species pair, the full model revealed that the interaction prob-
547 ability peaks at intermediate temperature and precipitation, simply because this is where
548 the two species co-occur most often. This does not imply that species will always interact
549 when they meet – although this is a basic assumption in most documented metawebs to
550 date (e.g. Havens 1992; Wood *et al.* 2015). Rather, an interaction is a stochastic process
551 whose probability is also influenced by the probability with which species co-occur. What
552 we cannot reliably know is how this stochasticity splits into two sampling processes – i.e.,
553 the extent to which a species at the higher trophic level runs into a species at the lower
554 level co-occurring at the site, and the extent to which this interaction is detected by an
555 observer collecting a finite sample. Future work will be required to document the relative
556 importance of these two sources of uncertainty in the occurrence of interactions.

557 Perspectives

558 Evidence that the structure of ecological networks does vary across habitats (e.g. Tylianakis
559 *et al.* 2007), over environmental gradients Lurgi *et al.* 2012 and in time (Trøjelsgaard *et al.*,
560 2015) is accumulating rapidly. It is not clear, however, to what extent the turnover of
561 network structure is driven by a systematic change in species composition or of pair-
562 wise interactions (Poisot *et al.*, 2012, 2015b). Our model comparison of host-parasite
563 interactions revealed that most of the turnover is driven by species-specific responses to
564 the environment, impacting species richness, and that co-occurrence was mostly neutral.
565 Further, the occurrence of interactions among host and parasite is highly stochastic even
566 when both are present, and not predictable by the variables considered by us. We know
567 that interactions vary with the environment in other systems, for instance, herbivory
568 (Shurin *et al.*, 2012) and predation (McKinnon *et al.*, 2010; Legagneux *et al.*, 2014) are
569 often found to increase with temperature, resulting in spatial variation of trophic cascades
570 (Gray *et al.*, 2015). What remains unclear, however, is the extent to which such variation
571 is driven by a turnover of species composition along gradients, or a turnover of the in-
572 teractions. Here we found that interactions vary substantially but non-predictably along
573 the annual temperature and the precipitation gradient. Clearly, the lack of detectable
574 signal may be due to our choice of covariates. Indeed, a previous study on a similar
575 system identified habitat characteristics as the primary drivers of interactions (Nyman
576 *et al.*, 2015). New investigations with other systems will thus be required to challenge this
577 result. Under all circumstances, documenting the relationship between the environment
578 and the occurrence of interactions at continental scales is critical for understanding how
579 large-scale variation of trophic regulation influences community dynamics and ecosystem
580 functioning (Harfoot *et al.*, 2014).

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

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

600 At present, there is only indirect support for the hypothesis that interacting species are
601 conditionally distributed but this possibility should be the topic of more specific hypothe-
602 sis testing. The impact of ecological interactions on the distribution of co-occurrence has
603 been the topic of many publications since Diamond 1975 seminal study on competition
604 and "checkerboard" distribution, but pairwise approaches have only recently received at-
605 tention (Veech, 2013). Whether two interacting species are more closely associated in
606 space remains unclear, since most approaches based on null models consider community-
607 level metrics (e.g. Gotelli 2000), such as the C-score, thereby making it hard to evaluate
608 if specific interactions do indeed affect co-occurrence. The expansion of the framework

609 we describe to account for the difference between the realized and the fundamental niche
610 will therefore require further investigation of the impact of interactions on co-occurrence.

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

629 Conclusion

630 Our representation of spatial variation of community structure presents a new approach
631 for the study of the biogeography of ecological networks. We see the following key chal-
632 lenges and opportunities ahead in this exciting area of research:

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

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

645 **3. From interaction probabilities to a distribution of network properties.**
646 Metrics are available to analyze the structure of probabilistic networks (Poisot *et al.*,
647 2016a). These metrics are useful as first approximation, but they assume independence
648 among interactions. This might not be the case in nature because of the role of co-
649 occurrence and shared environmental requirements. We also need to better understand
650 the distribution of network properties arising from probabilistic interactions.

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

655 **5. Effects of ecological interactions on co-occurrence.** We have intentionally
656 omitted the feedback of ecological interactions on co-occurrence in this framework. As
657 abundance can impact the occurrence of interactions, and conversely since interactions
658 impact abundance (Canard *et al.*, 2014), we could reasonably expect that interactions

659 will also influence co-occurrence. Theory in this regard does exist for simple three-species
660 modules (Cazelles *et al.*, 2015), but its extension to entire co-occurrence networks will
661 prove critical in the future, especially given the interest in using co- occurrence to infer
662 ecological interactions (Morales-Castilla *et al.*, 2015; Morueta-Holme *et al.*, 2016).

663 **Acknowledgements**

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

670 References

- 671 Albouy, C., Velez, L., Coll, M., Colloca, F., Le Loc'h, F., Mouillot, D. & Gravel, D. (2014)
672 From projected species distribution to food-web structure under climate change. *Global*
673 *Change Biology* **20**, 730–741.
- 674 Allesina, S., Alonso, D. & Pascual, M. (2008) A general model for food web structure.
675 *Science* **320**, 658–61.
- 676 Anderson, M.J., Crist, T.O., Chase, J.M., Vellend, M., Inouye, B.D., Freestone, A.L.,
677 Sanders, N.J., Cornell, H.V., Comita, L.S., Davies, K.F., Harrison, S.P., Kraft, N.J.B.,
678 Stegen, J.C. & Swenson, N.G. (2011) Navigating the multiple meanings of beta-
679 diversity: a roadmap for the practicing ecologist. *Ecology Letters* **14**, 19–28.
- 680 Argus, G. (1997) Infrageneric classification of *Salix* (Salicaceae) in the New World. *Sys-*
681 *tematic Botany Monographs* **52**, 1–121.
- 682 Baiser, B., Gotelli, N.J., Buckley, H.L., Miller, T.E. & Ellison, A.M. (2012) Geographic
683 variation in network structure of a nearctic aquatic food web. *Global Ecology and Bio-*
684 *geography* **21**, 579–591.
- 685 Bartomeus, I., Gravel, D., Tylianakis, J., Aizen, M., Dickie, I. & Bernard-Verdier, M.
686 (2016) A common framework for identifying linkage rules across different types of in-
687 teractions. *Functional Ecology* pp. n/a–n/a.
- 688 Bascompte, J. & Jordano, P. (2007) Plant-Animal Mutualistic Networks: The Architec-
689 ture of Biodiversity. *Annual Review of Ecology, Evolution, and Systematics* **38**, 567–
690 593.
- 691 Baselga, A. (2010) Partitioning the turnover and nestedness components of beta diversity.
692 *Global Ecology and Biogeography* **19**, 134–143.

- 693 Bersier, L.F., Banašek-Richter, C. & Cattin, M.F. (2002) Quantitative descriptors of
694 food-web matrices. *Ecology* **83**, 2394–2407.
- 695 Blonder, B., Lamanna, C., Violle, C. & Enquist, B.J. (2014) The n-dimensional hyper-
696 volume. *Global Ecology and Biogeography* **23**, 595–609.
- 697 Boulangeat, I., Gravel, D. & Thuiller, W. (2012) Accounting for dispersal and biotic
698 interactions to disentangle the drivers of species distributions and their abundances.
699 *Ecology Letters* **15**, 584–593.
- 700 Canard, E., Mouquet, N., Mouillot, D., Stanko, M., Miklisova, D. & Gravel, D. (2014)
701 Empirical evaluation of neutral interactions in host-parasite networks. *The American*
702 *naturalist* **183**, 468–79.
- 703 Cazelles, K., Araújo, M.B., Mouquet, N. & Gravel, D. (2016) A theory for species co-
704 occurrence in interaction networks. *Theoretical Ecology* **9**, 39–48.
- 705 Cazelles, K., Mouquet, N., Mouillot, D. & Gravel, D. (2015) On the integration of biotic
706 interaction and environmental constraints at the biogeographical scale. *Ecography* pp.
707 n/a–n/a.
- 708 Chase, J. & Leibold, M. (2003) *Ecological niches*. Chicago University Press, Chicago.
- 709 Chase, J.M. (2007) Drought mediates the importance of stochastic community assembly.
710 *Proceedings of the National Academy of Sciences of the United States of America* **104**,
711 17430–4.
- 712 Connor, E. & Simberloff, D. (1979) The assembly of species communities: chance or
713 competition? *Ecology* **60**, 1132–1140.

- 714 Cronk, Q., Ruzzier, E., Belyaeva, I. & Percy, D. (2015) Salix transect of Europe: lati-
715 tudinal patterns in willow diversity from Greece to arctic Norway. *Biodiversity Data*
716 *Journal* **3**.
- 717 Diamond, J. (1975) Assembly of species communities. *Ecology and evolution of communi-*
718 *ties* (eds. M. Cody & J. Diamond), pp. 342–444, Harvard University Press, Cambridge.
- 719 Dunne, J.A. (2006) The network structure of food webs. *Ecological networks: Linking*
720 *structure and dynamics* (eds. M. Pascual & J.A. Dunne), pp. 27–86, Oxford University
721 Press, Oxford.
- 722 Eklöf, A., Jacob, U., Kopp, J., Bosch, J., Castro-Urgal, R., Chacoff, N.P., Dalsgaard, B.,
723 de Sassi, C., Galetti, M., Guimarães, P.R., Lomáscolo, S.B., Martín González, A.M.,
724 Pizo, M.A., Rader, R., Rodrigo, A., Tylianakis, J.M., Vázquez, D.P. & Allesina, S.
725 (2013) The dimensionality of ecological networks. *Ecology letters* **16**, 577–583.
- 726 Gibert, J.P. & DeLong, J.P. (2014) Temperature alters food web body-size structure.
727 *Biology Letters* **10**.
- 728 Gleason, H. (1926) The individualistic concept of the plant association. *Bulletin of the*
729 *Torrey Botanical Club* **53**, 7–26.
- 730 Gotelli, N.J. (2000) Null Model Analysis of Species Co-Occurrence Patterns. *Ecology* **81**,
731 2606.
- 732 Götzenberger, L., de Bello, F., Bråthen, K.A., Davison, J., Dubuis, A., Guisan, A., Lepš,
733 J., Lindborg, R., Moora, M., Pärtel, M., Pellissier, L., Pottier, J., Vittoz, P., Zobel,
734 K. & Zobel, M. (2012) Ecological assembly rules in plant communities—approaches,
735 patterns and prospects. *Biological reviews of the Cambridge Philosophical Society* **87**,
736 111–27.

- 737 Gravel, D., Massol, F., Canard, E., Mouillot, D. & Mouquet, N. (2011) Trophic theory
738 of island biogeography. *Ecology letters* **14**, 1010–6.
- 739 Gravel, D., Poisot, T., Albouy, C., Velez, L. & Mouillot, D. (2013) Inferring food web
740 structure from predator-prey body size relationships. *Methods in Ecology and Evolution*
741 **4**, 1083–1090.
- 742 Gray, S.M., Poisot, T., Harvey, E., Mouquet, N., Miller, T.E. & Gravel, D. (2015) Temper-
743 ature and trophic structure are driving microbial productivity along a biogeographical
744 gradient. *Ecography* pp. n/a–n/a.
- 745 Hanski, I. (1999) *Metapopulation Ecology*. Oxford University Press, Oxford.
- 746 Hardin, G. (1960) The competitive exclusion principle. *Science* **131**, 1292–1297.
- 747 Harfoot, M.B.J., Newbold, T., Tittensor, D.P., Emmott, S., Hutton, J., Lyutsarev, V.,
748 Smith, M.J., Scharlemann, J.P.W. & Purves, D.W. (2014) Emergent global patterns of
749 ecosystem structure and function from a mechanistic general ecosystem model. *PLoS*
750 *Biol* **12**, 1–24.
- 751 Havens, K. (1992) Scale and structure in natural food webs. *Science* **257**, 1107–9.
- 752 Hijmans, R.J. (2015) *raster: Geographic Data Analysis and Modeling*. R package version
753 2.5-2.
- 754 Holt, R.D. (2009) Bringing the Hutchinsonian niche into the 21st century : Ecological
755 and evolutionary perspectives. *Proceedings of the National Academy of Sciences* **106**,
756 19659–19665.
- 757 Holt, R.D. & Lawton, J.H. (1993) Apparent competition and enemy-free space in insect
758 host-parasitoid communities. *The American Naturalist* **142**, 623–645, PMID: 19425963.

- 759 Hutchinson, G. (1957) Concluding remarks. *Cold Spring Harbour Symposium* **22**, 415–
760 427.
- 761 Ings, T.C., Montoya, J.M., Bascompte, J., Blüthgen, N., Brown, L., Dormann, C.F.,
762 Edwards, F., Figueroa, D., Jacob, U., Jones, J.I., Lauridsen, R.B., Ledger, M.E.,
763 Lewis, H.M., Olesen, J.M., van Veen, F.J.F., Warren, P.H. & Woodward, G. (2009)
764 Ecological networks—beyond food webs. *The Journal of animal ecology* **78**, 253–69.
- 765 Kéfi, S., Berlow, E.L., Wieters, E.a., Navarrete, S.a., Petchey, O.L., Wood, S.a., Boit, A.,
766 Joppa, L.N., Lafferty, K.D., Williams, R.J., Martinez, N.D., Menge, B.a., Blanchette,
767 C.a., Iles, A.C. & Brose, U. (2012) More than a meal... integrating non-feeding inter-
768 actions into food webs. *Ecology letters* pp. 291–300.
- 769 Kissling, W.D., Dormann, C.F., Groeneveld, J., Hickler, T., Kühn, I., McNerny, G.J.,
770 Montoya, J.M., Römermann, C., Schiffers, K., Schurr, F.M., Singer, A., Svenning,
771 J.C., Zimmermann, N.E. & O’Hara, R.B. (2012) Towards novel approaches to mod-
772 elling biotic interactions in multispecies assemblages at large spatial extents. *Journal*
773 *of Biogeography* **39**, 2163–2178.
- 774 Kopelke, J.P. (1985) Über die Biologie und Parasiten der gallenbildenden Blattwespe-
775 naften *Pontania dolichura* (Thoms, 1871), *P. vesicator* (Bremi 1849) und *P. viminalis* (L.
776 1758) (Hymenoptera: Tenthredinidae). *Faunistisch-Ökologische Mitteilungen* **5**, 331–
777 344.
- 778 Kopelke, J.P. (1999) Gallenerzeugende Blattwespen Europas - Taxonomische Grundlagen,
779 Biologie und Ökologie (Tenthredinidae: Nematinae: Euura, Phyllocolpa, Pontania).
780 *Courier Forschungsinstitut Senckenberg* **212**, 1–183.
- 781 Kopelke, J.P. (2003) Natural enemies of gall-forming sawflies on willows (*Salix* spp.).
782 *Entomologia Generalis* **26**, 277–312.

- 783 Kouki, J. (1999) Latitudinal gradients in species richness in northern areas: some excep-
784 tional patterns. *Ecological Bulletins* **47**, 30–37.
- 785 Kouki, J., Niemelä, P. & Viitasaari, M. (1994)) Reversed latitudinal gradient in species
786 richness of sawflies (Hymenoptera, Symphyta). *Annales Zoologici Fennici* **31**, 83–88.
- 787 Krebs, C. (2009) *Ecology: The Experimental Analysis of Distribution and Abundance*. 6th
788 ed. Pearson, San Francisco.
- 789 Laliberté, E. & Tylianakis, J.M. (2010) Deforestation homogenizes tropical parasitoid-
790 host networks. *Ecology* **91**, 1740–1747.
- 791 Legagneux, P., Gauthier, G., Lecomte, N., Schmidt, N.M., Reid, D., Cadieux, M.C.,
792 Berteaux, D., Bêty, J., Krebs, C.J., Ims, R.a., Yoccoz, N.G., Morrison, R.I.G., Leroux,
793 S.J., Loreau, M. & Gravel, D. (2014) Arctic ecosystem structure and functioning shaped
794 by climate and herbivore body size. *Nature Climate Change* **E2168**, 1–5.
- 795 Legendre, P., Borcard, D. & Peres-Neto, P.R. (2005) Analyzing beta diversity: partition-
796 ing the spatial variation of composition data. *Ecological Monographs* **75**, 435–450.
- 797 Legendre, P. & De Cáceres, M. (2013) Beta diversity as the variance of community data:
798 dissimilarity coefficients and partitioning. *Ecology Letters* **16**, 951–963.
- 799 Leibold, M., Holyoak, M., Mouquet, N., Amarasekare, P., Chase, J.M., Hoopes, M.,
800 Holt, R.D., Shurin, J.B., Law, R., Tilman, D., Loreau, M. & Gonzalez, A. (2004) The
801 metacommunity concept : a framework for multi-scale community ecology. *Ecology*
802 *Letters* **7**, 601–613.
- 803 Leppänen, S., Malm, T., Värri, K. & Nyman, T. (2014) A comparative analysis of genetic
804 differentiation across six shared willow host species in leaf- and bud-galling sawflies.
805 *PLoS ONE* **9**, e116286.

- 806 Lurgi, M., López, B.C. & Montoya, J.M. (2012) Novel communities from climate change.
807 *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*
808 **367**, 2913–22.
- 809 Martinez, N.D., Hawkins, B.a., Dawah, H.A. & Feifarek, B.P. (1999) Effects of Sampling
810 Effort on Characterization of Food-Web Structure. *Ecology* **80**, 1044–1055.
- 811 McNerny, G.J. & Etienne, R.S. (2012a) Ditch the niche - is the niche a useful concept in
812 ecology or species distribution modelling? *Journal of Biogeography* **39**, 2096–2102.
- 813 McNerny, G.J. & Etienne, R.S. (2012b) Stitch the niche - a practical philosophy and
814 visual schematic for the niche concept. *Journal of Biogeography* **39**, 2103–2111.
- 815 McKinnon, L., Smith, P.A., Nol, E., Martin, J.L., Doyle, F.I., Abraham, K.F., Gilchrist,
816 H.G., Morrison, R.I.G. & Bêty, J. (2010) Lower predation risk for migratory birds at
817 high latitudes. *Science* **327**, 326–327.
- 818 Morales-Castilla, I., Matias, M.G., Gravel, D. & Araújo, M.B. (2015) Inferring biotic
819 interactions from proxies. *Trends in Ecology & Evolution* **30**, 347–356.
- 820 Morueta-Holme, N., Blonder, B., Sandel, B., McGill, B.J., Peet, R.K., Ott, J.E., Violle,
821 C., Enquist, B.J., Jørgensen, P.M. & Svenning, J.C. (2016) A network approach for
822 inferring species associations from co-occurrence data. *Ecography* pp. n/a–n/a.
- 823 Nyman, T., Bokma, F. & Kopelke, J.P. (2007) Reciprocal diversification in a complex
824 plant-herbivore-parasitoid food web. *BMC Biology* **5**, 49.
- 825 Nyman, T. & Julkunen-Tiitto, R. (2000) Manipulation of the phenolic chemistry of wil-
826 lows by gall-inducing sawflies. *Proceedings of the National Academy of Sciences* **97**,
827 13184–13187.

- 828 Nyman, T., Leppänen, S.A., Várkonyi, G., Shaw, M.R., Koivisto, R., Barstad, T.E.,
829 Vikberg, V. & Roininen, H. (2015) Determinants of parasitoid communities of willow-
830 galling sawflies: habitat overrides physiology, host plant and space. *Molecular Ecology*
831 **24**, 5059–5074.
- 832 Nyman, T., Vikberg, V., Smith, D. & BoevéJ-L (2010) How common is ecological speci-
833 ation in plant-feeding insects? A “Higher” Nematinae perspective. *BMC Evolutionary*
834 *Biology* **10**, e266.
- 835 Pellissier, L., Rohr, R., Ndiribe, C., Pradervand, J.N., Salamin, N., Guisan, A. & Wisz, M.
836 (2013) Combining food web and species distribution models for improved community
837 projections. *Ecology and Evolution* **3**, 4572–4583.
- 838 Petchey, O.L., Brose, U. & Rall, B.C. (2010) Predicting the effects of temperature on food
839 web connectance. *Philosophical transactions of the Royal Society of London. Series B,*
840 *Biological sciences* **365**, 2081–91.
- 841 Peters, R. (1991) *A critique for ecology*. Cambridge University Press, Cambridge.
- 842 Phillips, S.J., Anderson, R.P. & Schapire, R.E. (2006) Maximum entropy modeling of
843 species geographic distributions. *Ecological Modelling* **190**, 231 – 259.
- 844 Pillai, P., Gonzalez, A. & Loreau, M. (2011) Metacommunity theory explains the emer-
845 gence of food web complexity. *Proceedings of the National Academy of Sciences of the*
846 *United States of America* **108**, 19293–19298.
- 847 Poisot, T., Baiser, B., Dunne, J.A., Kéfi, S., Massol, F., Mouquet, N., Romanuk, T.N.,
848 Stouffer, D.B., Wood, S.A. & Gravel, D. (2015a) mangal – making ecological network
849 analysis simple. *Ecography* **37**.
- 850 Poisot, T., Canard, E., Mouillot, D., Mouquet, N. & Gravel, D. (2012) The dissimilarity
851 of species interaction networks. *Ecology letters* **XX**, XX–XX.

- 852 Poisot, T., Cirtwill, A.R., Cazelles, K., Gravel, D., Fortin, M.J. & Stouffer, D.B. (2016a)
853 The structure of probabilistic networks. *Methods in Ecology and Evolution* **7**, 303–312.
- 854 Poisot, T. & Gravel, D. (2014) When is an ecological network complex? connectance
855 drives degree distribution and emerging network properties. *PeerJ* **2**, e251.
- 856 Poisot, T., Gravel, D., Leroux, S., Wood, S.A., Fortin, M.J., Baiser, B., Cirtwill, A.R.,
857 Araújo, M.B. & Stouffer, D.B. (2016b) Synthetic datasets and community tools for the
858 rapid testing of ecological hypotheses. *Ecography* **39**, 402–408.
- 859 Poisot, T., Stouffer, D.B. & Gravel, D. (2015b) Beyond species: why ecological interac-
860 tions vary through space and time. *Oikos* **124**, 243–251.
- 861 Pollock, L.J., Tingley, R., Morris, W.K., Golding, N., O’Hara, R.B., Parris, K.M., Vesk,
862 P.A. & McCarthy, M.A. (2014) Understanding co-occurrence by modelling species si-
863 multaneously with a joint species distribution model (jsdm). *Methods in Ecology and*
864 *Evolution* **5**, 397–406.
- 865 Pulliam, H.R. (2000) On the relationship between niche and distribution. *Ecology Letters*
866 **3**, 349–361.
- 867 Roininen, H., Nyman, T. & Zinovjev, A. (2005) Biology, ecology, and evolution of gall-
868 inducing sawflies (Hymenoptera: Tenthredinidae and Xyelidae). *Biology, Ecology, and*
869 *Evolution of Gall-Inducing Arthropods* (eds. A. Raman, C.W. Schaefer & T.M. With-
870 ers), pp. 467–494, Science Publishers, Inc.
- 871 Scheiner, S.M. & Willig, M.R. (2007) A general theory of ecology. *Vital And Health*
872 *Statistics. Series 20 Data From The National Vitalstatistics System Vital Health Stat*
873 *20 Data Natl Vital Sta .*
- 874 Shurin, J.B., Clasen, J.L., Greig, H.S., Kratina, P. & Thompson, P.L. (2012) Warm-
875 ing shifts top-down and bottom-up control of pond food web structure and function.

- 876 *Philosophical Transactions of the Royal Society of London B: Biological Sciences* **367**,
877 3008–3017.
- 878 Silvertown, J. (2004) Plant coexistence and the niche. *Trends in Ecology and Evolution*
879 **19**, 605–611.
- 880 Skvortsov, A.K. (1999) Willows of Russia and adjacent countries. Taxonomical and geo-
881 graphical revision. *Univ. Joensuu Fac. Math. Natl. Sci. Rep. Ser* **39**, 1–307.
- 882 Soberón, J. (2007) Grinnellian and eltonian niches and geographic distributions of species.
883 *Ecology Letters* **10**, 1115–1123.
- 884 Stouffer, D.B., Camacho, J., Jiang, W. & Amaral, L.a.N. (2007) Evidence for the existence
885 of a robust pattern of prey selection in food webs. *Proceedings of the Royal Society B:*
886 *Biological Sciences* **274**, 1931–40.
- 887 Thuiller, W. (2003) Biomod – optimizing predictions of species distributions and project-
888 ing potential future shifts under global change. *Global Change Biology* **9**, 1353–1362.
- 889 Tilman, D. (1982) *Resource competition and community structure*. Princeton University
890 Press, Princeton.
- 891 Townsend, P.A., Soberón, S., Pearson, R.G., Anderson, R.P., Martínez-Meyer, E., Miguel,
892 N. & Araújo, M.B. (2011) *Ecological Niches and Geographic Distributions*. Princeton
893 University Press, Princeton.
- 894 Trøjelsgaard, K., Jordano, P., Carstensen, D.W. & Olesen, J.M. (2015) Geographical
895 variation in mutualistic networks: similarity, turnover and partner fidelity. *Proceedings*
896 *of the Royal Society of London B: Biological Sciences* **282**, 20142925.
- 897 Tylianakis, J.M., Tscharntke, T. & Lewis, O.T. (2007) Habitat modification alters the
898 structure of tropical host–parasitoid food webs. *Nature* **445**, 202–205.

- 899 Veech, J.A. (2013) A probabilistic model for analysing species co-occurrence. *Global Ecol-*
900 *ogy and Biogeography* **22**, 252–260.
- 901 Veijalainen, A., Sääksjärvi, I., Erwin, T., Gómez, I. & Longino, J. (2013) Subfamily
902 composition of Ichneumonidae (Hymenoptera) from western Amazonia: insights into
903 diversity of tropical parasitoid wasps. *Insect Conservation and Diversity* **6**, 28–37.
- 904 Vellend, M., Srivastava, D.S., Anderson, K.M., Brown, C.D., Jankowski, J.E., Kleyn-
905 hans, E.J., Kraft, N.J.B., Letaw, A.D., Macdonald, A.A.M., Maclean, J.E., Myers-
906 Smith, I.H., Norris, A.R. & Xue, X. (2014) Assessing the relative importance of neutral
907 stochasticity in ecological communities. *Oikos* **123**, 1420–1430.
- 908 Williams, R. & Martinez, N. (2000) Simple rules yield complex food webs. *Nature* **404**,
909 180–183.
- 910 Williams, R.J., Anandanadesan, A. & Purves, D. (2010) The probabilistic niche model
911 reveals the niche structure and role of body size in a complex food web. *PloS One* **5**,
912 e12092.
- 913 Wisz, M.S., Pottier, J., Kissling, W.D., Pellissier, L., Lenoir, J., Damgaard, C.F., Dor-
914 mann, C.F., Forchhammer, M.C., Grytnes, J.A., Guisan, A., Heikkinen, R.K., Høye,
915 T.T., Kühn, I., Luoto, M., Maiorano, L., Nilsson, M.C., Normand, S., Ockinger, E.,
916 Schmidt, N.M., Termansen, M., Timmermann, A., Wardle, D.a., Aastrup, P. & Sven-
917 ning, J.C. (2012) The role of biotic interactions in shaping distributions and realised
918 assemblages of species: implications for species distribution modelling. *Biological Re-*
919 *views* .
- 920 Wood, S.A., Russell, R., Hanson, D., Williams, R.J. & Dunne, J.A. (2015) Effects of
921 spatial scale of sampling on food web structure. *Ecology and Evolution* **5**, 3769–3782.

- 922 Woodward, G., Perkins, D.M. & Brown, L.E. (2010) Climate change and freshwater
923 ecosystems: impacts across multiple levels of organization. *Philosophical transactions*
924 *of the Royal Society of London. Series B, Biological sciences* **365**, 2093–106.
- 925 Wu, J., Nyman, T., Wang, D.C., Argus, G., Yang, Y.P. & Chen, J.H. (2015) Phylogeny
926 of *Salix* subgenus *Salix* s.l. (Salicaceae): delimitation, biogeography, and reticulate
927 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, galls and parasitoids.

Interaction	#	Metaweb model	Co-occurrence model	LL	npars	AIC
Salix-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

928 **Figure legends**

929 **Figure 1**

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

938 **Figure 2**

939 **Visual representation of the integrated niche.** In biogeography, the niche is consid-
940 ered the set of environmental conditions where the intrinsic growth rate r is positive (Holt,
941 2009). The horizontal axis represents an environmental gradient impacting the growth of
942 the focal species (in red). The location of each species along this gradient represents their
943 optimum, and the vertical dotted lines represent the limits of the Grinnellian niche of the
944 focal species. In food web ecology, the Eltonian niche represents the location of a species
945 in the food web, as determined by its niche position (n) and its niche optimum (c). The
946 vertical axis represents a niche gradient, for example a trait such as body size. The loca-
947 tion of each species along this gradient represents their niche position. The focal species
948 will feed only on prey species occupying niche locations within a given interval around the
949 optimum, represented by the horizontal lines. The integrated Grinnellian and Eltonian
950 niche corresponds to the square in the middle where an interaction is possible owing to
951 a match of traits and spatial distribution. According to our probabilistic framework, the

952 central square represents the area where the joint probability of observing co-occurrence
953 and interactions is positive.

954 **Figure 3**

955 **Probabilistic representation of the interaction probability between a leaf folder**
956 **(*Phyllocolpa prussica*) and a parasitoid (*Chrysocharis elongata*) across gradi-**
957 **ents of annual average temperature and annual precipitation.** The representation
958 is based on predictions from Model 3 (see Table 1). In the left panel, open circles repre-
959 sent the absence of both species, whereas closed circles represent co- occurrence and plus
960 signs the occurrence of only one of the two species. In the other two panels, open circles
961 represent co-occurrence but an absence of interaction and closed circles the occurrence of
962 an interaction.

963 **Figure 4**

964 **Probabilistic representation of the interaction probability between a leaf folder**
965 **(*Phyllocolpa prussica*) and a parasitoid (*Chrysocharis elongata*) across Eu-**
966 **rope.** The maps are generated from probabilities predicted by the model illustrated
967 in Fig. 3.

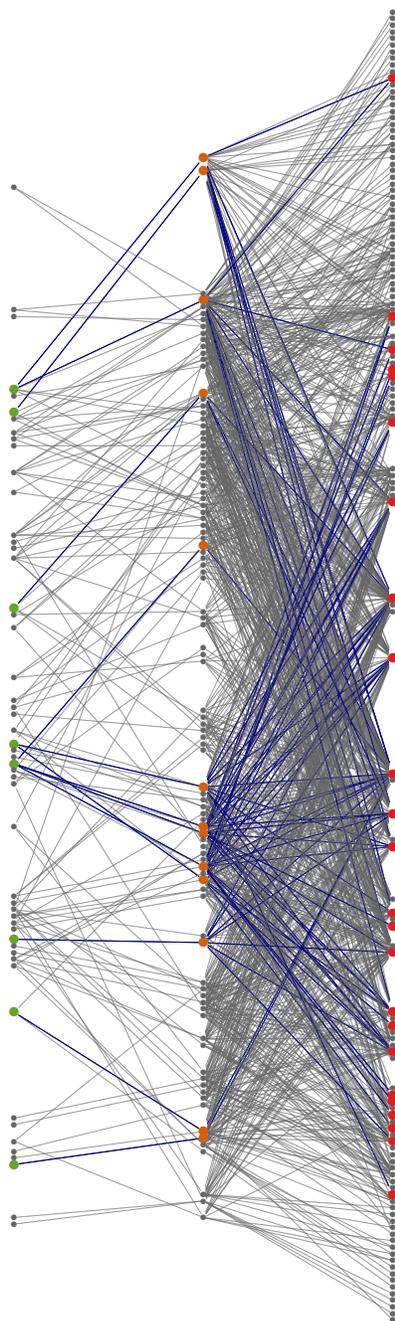
968 **Figure 5**

969 **Representation of the Salix-galler and galler-parasitoid metawebs.** Black cells
970 indicate species pairs for which at least one interaction was recorded, white cells indicate
971 absence of recorded interactions and grey cells show pairs of species never detected at
972 the same site (and hence species pairs for which we have no information on whether they
973 would interact should they co-occur).

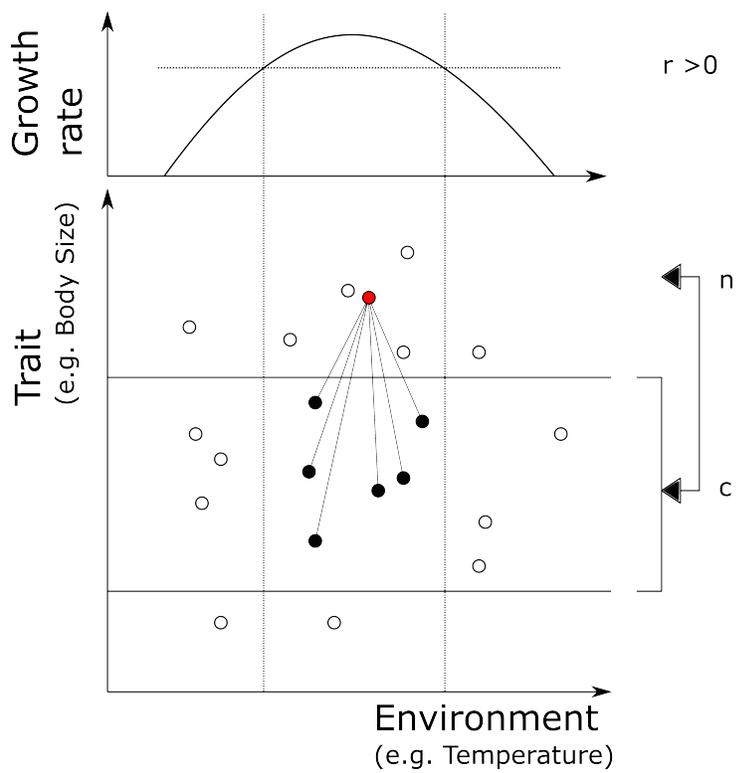
974 **Figure 6**

975 **Mapping the distribution of species richness, the number of links and con-**
976 **nectance across Europe.** The representation is based on predictions from Model 3
977 (see Table 2). Species richness is obtained by summation of individual occurrence prob-
978 abilities, and link density by summation of interaction probabilities.

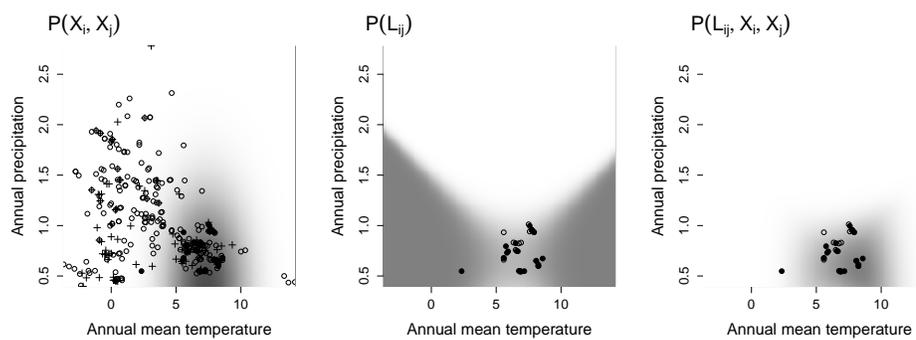
979 **Figure 1**



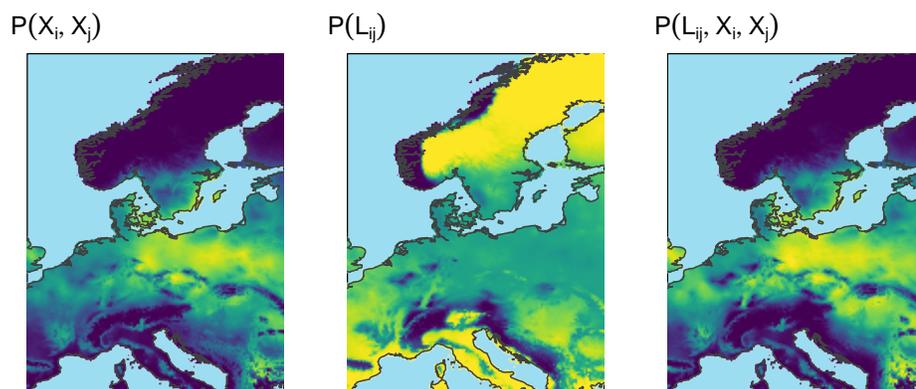
980 **Figure 2**



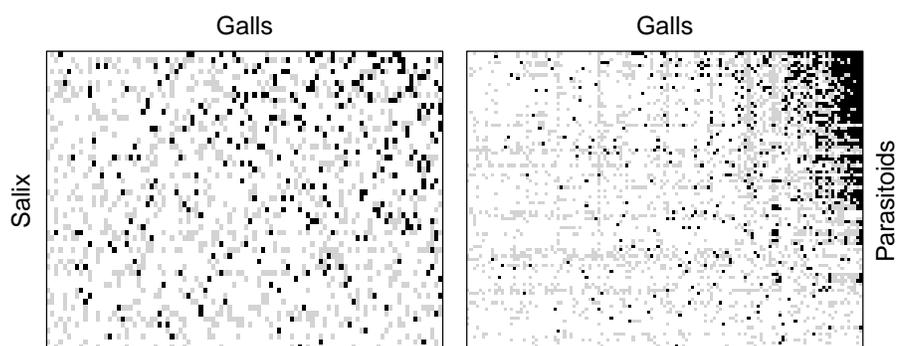
981 **Figure 3**



982 **Figure 4**



983 **Figure 5**



984 **Figure 6**

