Navigating the integration of Biotic Interactions in Biogeography

Wilfried Thuiller\textsuperscript{1}, Irene Calderon-Sanou\textsuperscript{1,2,3}, Loïc Chalmandrier\textsuperscript{4}, Pierre Gaüzere\textsuperscript{1}, Louise O'Connor\textsuperscript{1,5}, Marc Ohlmann\textsuperscript{†}, Giovanni Poggiato\textsuperscript{1} and Tamara Münkemüller\textsuperscript{1}

\textsuperscript{1} Univ. Grenoble Alpes, Univ. Savoie Mont Blanc, CNRS, LECA, F-38000 Grenoble
\textsuperscript{2} German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig, Puschstrasse 4, 04103, Leipzig, Germany
\textsuperscript{3} Inst. of Biology, Martin Luther Univ. Halle Wittenberg, Halle (Saale), Germany
\textsuperscript{4} Univ. Grenoble Alpes, CNRS, LIPhy, F-38000 Grenoble
\textsuperscript{5} Biodiversity, Ecology and Conservation group, International Institute of Applied Systems Analysis, Laxenburg, Austria

\textsuperscript{†}Deceased before submission (17/06/2023).
Abstract

Biotic interactions are widely recognized as the backbone of ecological communities, but how best to study them is a subject of intense debate, especially at macro-ecological scales. While some researchers claim that biotic interactions need to be observed directly, others use proxies and statistical approaches to infer them. Despite this ambiguity, studying and predicting the influence of biotic interactions on biogeographic patterns is a thriving area of research with crucial implications for conservation. Three distinct approaches are currently being explored. The first approach involves empirical observation and measurement of biotic interactions' effects on species demography in laboratory or field settings. While these findings contribute to theory and to understanding species’ demographies, they can be challenging to generalize on a larger scale. The second approach centers on inferring biotic associations from observed co-occurrences in space and time. The goal is to distinguish the environmental and biotic effects on species distributions. The third approach constructs extensive potential interaction networks, known as metanetworks, by leveraging existing knowledge about species ecology and interactions. This approach analyses local realizations of these networks using occurrence data and allows understanding large distributions of multi-taxon assemblages. In this piece, we appraise these three approaches, highlighting their respective strengths and limitations. Instead of seeing them as conflicting, we advocate for their integration to enhance our understanding and expand applications in the emerging field of interaction biogeography. This integration shows promise for ecosystem understanding and management in the Anthropocene era.
Introduction

The study of biotic interactions has entangled researchers in debates surrounding their precise definition and delineation. This term can encompass various phenomena, such as a butterfly feeding on a flower, the direct and indirect effects of one species on the population dynamics of another, or the biotic factors influencing species ranges (Alexander et al., 2015). In the literature, biotic interactions have been observed in the field, measured through controlled experiments, collected from expert knowledge, or inferred from statistical relationships between co-occurring species (Lany et al., 2018; Ohlmann et al., 2018) or from species traits (Caron et al., 2022; Gravel et al., 2013). Consequently, the definition of biotic interactions can vary depending on the research question, the organisms under study, the available data or data collection methods, and the study setup. Furthermore, the study of biotic interactions becomes progressively more challenging as spatial scales, species richness, and the variety of interaction types (e.g., competition, mutualism, predator-prey) increase. Nevertheless, the field of biogeography is now more than ever driven to comprehend the role of biotic interactions in large-scale biodiversity patterns, and to utilize this understanding to construct more robust biodiversity models and scenarios (Pollock et al., 2020; Urban et al., 2016).

In the early 20th century, the study of biotic interactions primarily focused on understanding the temporal dynamics of interacting populations, driven by the emergence of mathematical ecology and findings from experiments and observational studies. The theoretical work of Lotka (1925) and Volterra (1926) provided valuable insights into how the temporal dynamics of predators and prey are interconnected and influence one another. Elton's empirical investigations on the fluctuations of hare and lynx populations further supported these predictions (Elton & Nicholson, 1942). Additionally, experimental studies by Gause (1934a) supported the predictions of Lotka-Volterra competition models, demonstrating the potential of competition to restrict coexistence. Notably, during this period, there was limited consideration
given to how these results generalize to large spatial scales, interactions across trophic levels, and whether models could be used for predictions.

In the 1960s and 1970s, biogeographers, including Diamond (1975), sought to overpass the limitations imposed by small scales (in terms of space, species richness, and interaction types) by employing statistical techniques, particularly null models, to infer biotic interactions from patterns of co-occurring species. They hypothesized that interacting species would co-occur more (e.g., mutualism) or less frequently (e.g., competition) than expected by chance. Nowadays, researchers employ more sophisticated algorithms to infer biotic interactions from observed species occurrences, while accounting for environmental effects that may lead to synchronized species responses unrelated to interactions (Lany et al., 2018; Ovaskainen et al., 2017). Observed data are often spatial in nature, consisting of species abundances or simple presence/absence information. While inference methods can be powerful (see Dormann et al., 2018; Ovaskainen et al., 2017) for more comprehensive reviews), there exists a significant gap between the mathematical definition of biotic interactions at the individual or population level and the inferred processes derived from empirical data, which are typically measured at the species level and do not incorporate temporal dynamics (Blanchet et al., 2020).

In recent decades, there has been a growing interest in a distinct type of biotic interaction data that revolves around the concept of metawebs or metanetworks (Dunne, 2006). Metanetworks integrate information on species interactions from various sources, including observations (e.g., GLOBI), expert knowledge, literature reviews (Maiorano et al., 2020), and phylogenetic or trait inference (Caron et al., 2022; Llewelyn et al., 2023). Metanetwork data thus provide a comprehensive summary of all potential species interactions, without specifying the spatial and temporal variability in the strength and realisation of these interactions (Maiorano et al., 2020). The spatial information enters when local networks are constructed by filtering the metanetwork based on local co-occurrence or co-abundance data (Braga et al.,
Thus, metanetworks offer an intermediate perspective on biotic interactions, bridging the earlier definitions based on field observation and population dynamics, and the broader-scale inferences derived from co-occurrence data.

Rather than lamenting the limitations of individual approaches and concepts of biotic interactions in biogeography, we advocate for harnessing their respective strengths. Integrating and harmonizing these approaches is crucial not only to better understand the effects of biotic interactions on the distribution of biodiversity, but also to account for biotic interactions in biodiversity modelling and scenarios. In this paper, we review and elucidate the foundations of different concepts surrounding biotic interactions, and propose pathways for integrating distinct approaches in the field of interaction biogeography (Figure 1).
Section 1 - From in-situ observations and experiments to mathematical formulations

The study of biotic interactions is facilitated by focusing on smaller spatio-temporal scales, a reduced number of species, and specific interaction types, along with controlled study designs. At small scales, researchers can directly observe the interactions between individuals and witness their observable effects on each other's demography. Some might argue that these direct observations lie at the essence of the definition of biotic interactions. Unsurprisingly, earliest studies on biotic interactions were on experimental and observational approaches.

As early as 1862, Darwin observed and described how the shape and structure of orchids had evolved to attract and exploit specific species of bees, resulting in a mutualistic relationship (Darwin, 1862). Competition as a biotic interaction was defined through Gause's experiments with two species of Paramecium, demonstrating that when multiple species competed for the same resources, one species would eventually out-compete the other and become dominant (Gause, 1934b). Yet, it is here crucial to differentiate between experimental studies and observational studies. In experimental studies, organisms are manipulated to quantify not only the frequency of interactions, but also the impact of biotic interactions on other species’ demography (Figure 1). Theory, mathematics, and experiments are explicitly linked to unravel the dynamics of competitive, mutualistic, or predatory interactions and their influence on demography, and on community dynamics and stability (Kraft et al., 2015; Schupp et al., 2010).

On the other hand, observational studies measure the frequency of specific interactions (e.g., flower visitation rates, fruit removal), and indirectly infer the effect of biotic interactions on other species’ demography (Koch et al., 2023; Vázquez et al., 2005). Interactions are still observed, but their effects are often assumed or estimated through changes in abundance. Measured or estimated effects of biotic interactions on demography can then be formalized in mathematical models of community dynamics to allow for temporal extrapolation (Godoy & Levine, 2014).
But is an extrapolation to biogeographic scales and species rich systems possible as well? Biotic interactions are thought to influence large scale species distributions and diversity patterns (Comita et al., 2014; Gotelli et al., 2010), yet there is still very little empirical evidence to support it (Thuiller et al., 2015). Upscaling empirical studies of biotic interactions to biogeographical scales is a challenging task. The inherent challenge is to parametrise mathematical and mechanistic models describing species interactions and their effects on demography in multi-species communities. The number of pairwise interaction parameters that would need to be directly observed or manipulated increases non-linearly with the number of species (Figure 1). Furthermore, pairwise interactions may change along environmental gradients (for example, interactions among plants may shift from competition to facilitation along stress gradients, (Maestre et al., 2009). Sampling biotic interactions at large spatial scales presents a significant challenge, requiring extensive observational or experimental efforts that span diverse environments and encompass a wide range of species or taxa (i.e., the Eltonian shortfall, Hortal et al., 2015). New technologies might help here, not only allowing to measure different types of interactions more efficiently, but also to move from pairwise interactions to entire networks (Hartig et al., n.d.). For example, advancements in gut content analyses have significantly improved the quantification of resources acquired by organisms (Casey et al., 2019). Additionally, the integration of video and camera-traps with automatic recognition systems holds promise in observing multi-trophic interaction, such as plant-pollinator-herbivore-predator dynamics (Droissart et al., 2021). Furthermore, the increasingly widespread use of acoustic recorders facilitates a comprehensive understanding of species behaviour and how interactions may evolve over time at a large scale (De La Torre Cerro & Holloway, 2021; Schöner et al., 2016). These innovative in situ technologies can be deployed across expansive spatial areas, proving invaluable in measuring biotic interactions. However, one should note that most of these technologies need further development to be readily applicable in the context
of interaction biogeography (i.e., large spatial scales, multiple sensors and multiple types of observations), and they do not overcome the assumptions underlying the association between biotic interactions and demographic rates.

**Section 2 - From spatial patterns to statistical interactions**

Statistical approaches designed to infer biotic interactions from co-occurrence data have evolved considerably since the pioneering work of Diamond (1975) and Stone & Roberts (1990) on check-board scores. With the increasing availability of species distribution datasets and advancements in statistical modelling, the inference of biotic interactions from co-occurrence patterns has emerged as a vibrant research field (Losapio et al., 2021). This approach to interaction biogeography deviates from a strict definition of biotic interactions, as it does not rely on direct observations or ecological expert knowledge. Instead, statistical models are utilized to infer interactions between species based on the signal these interactions may have left in species co-occurrence or co-abundance patterns (Blanchet et al., 2020; Poggiato et al., 2021). The underlying rationale of this approach is that if two species interact, their distributions will depend on each other. The objective is to identify species with non-independent co-distributions, indicating a statistical interaction, which is assumed to be caused by biotic interactions (Figure 1). However, it is important to note that this approach infers marginal dependencies between species pairs (Veech, 2013), which means it does not consider the influence of other species on the joint distribution of species pairs. For instance, two species may co-occur merely because they share the same prey. To address this issue, statistical models employ conditional correlations to highlight conditional dependencies, which are pairwise associations between species while controlling for the effects of other species. The outcome of these models is often a conditional correlation network, representing the inferred associations based on co-occurrence patterns. Considering that species can co-occur due to shared
environmental preferences (i.e., niche similarity), statistical models also incorporate environmental variables as predictors to account for the effect of the environment (Pollock et al., 2014; Warton et al., 2015). Here, we clarify the connections between two recently introduced classes of models: graphical models and joint species distribution models (JSDMs).

Graphical models seek to infer the dependency structure between species using a graph with appropriate sparsity, which can be either directed (Larsen et al., 2012) or undirected (Harris, 2016; Ohlmann et al., 2018). These models are flexible and can be applied to various distribution types or combinations of distributions. Environmental covariates are typically included as additional nodes in the graph. On the other hand, JSDMs were initially developed to predict multi-species distributions based on environmental drivers while accounting for species dependencies (Ovaskainen et al., 2017; Pollock et al., 2014). These dependencies are represented by a residual covariance matrix that captures marginal dependencies, and conditional dependencies between species are obtained by inverting this matrix. Although JSDMs and undirected graphical models have not been directly compared in the literature, they share similarities in describing conditional dependence relationships between species (Momal et al., 2020). Specifically, a JSDM for Gaussian data without environmental covariates corresponds precisely to a Gaussian graphical model without penalization.

It is important to highlight that, regardless of the sophisticated statistical techniques employed, this approach relies solely on static co-occurrence data. As a result, it does not capture the mechanisms that govern the temporal dynamics between species as described by theoretical models. There is thus a significant gap between the statistical definition of a conditional correlation network and the actual effects of biotic interactions on demography and thus species' distributions. This gap arises from the fundamental hypothesis of this approach, which assumes that static co-occurrence patterns provide informative signals about the outcomes of biotic interactions. However, as mentioned in the first section, theoretical model
expectations suggest that the effects of biotic interactions on distribution patterns are dynamic (e.g., Lotka, 1925). For example, with static co-occurrence data of predator-prey systems, it is unclear whether high or low co-occurrence should be expected because predators both reduce prey populations and are more abundant when more prey is available.

Furthermore, even in cases where co-occurrence patterns do contain information about biotic interactions, this signal can be blurred by various factors. These factors include missing environmental covariates or a mismatch in scales (Blanchet et al., 2020; Dormann et al., 2018). Additionally, technical issues related to the models themselves can limit our ability to accurately infer biotic interactions. For example, JSDMs struggle to disentangle the influences of biotic and environmental signals (Poggiato et al., 2021). Empirical studies have also highlighted the poor correspondence between inferred interaction networks and known networks, further supporting these limitations (Freilich et al., 2018; Sander et al., 2017).

Moreover, the statistical models used in this approach typically yield either directed acyclic graphs (e.g., Bayesian networks, structural equation models, hierarchical models) or undirected networks (e.g., JSDMs, partial correlation networks). As a result, it becomes impossible to represent asymmetric interactions or more complex interaction structures such as feedback loops, which are common in ecosystems (Neutel et al., 2002).

Yet, statistical models provide a means to infer associations between species that co-occur more or less than expected by chance in a given environment. These associations can be informative on their own, such as identifying species associated with endangered ones (Han et al., 2020). By examining these associations in relation to species traits, we can investigate whether species with similar traits exhibit negative associations (limiting similarity) or not (hierarchical competition, (Elo et al., 2023). Additionally, these models can quantify the relative contributions of the environment, species statistical associations, and spatial correlation for each species (Norberg et al., 2019) or within each community (Leibold et al., 2022). This could
help partition the roles of different processes in community assembly and explore how they vary with spatial scale or environmental conditions.

**Section 3 - From fundamental knowledge and data integration to metanetworks**

Species interaction data can be acquired not only through new measurements but also from comprehensive databases, empirical knowledge, and various sources of information across different spatial and temporal dimensions (Le Guillarme & Thuiller, 2023). The accumulation of this collective knowledge now enables us to draw generalizations about interactions across diverse environments and taxa. When information is lacking for certain species, interactions can be inferred from known species to those that are functionally or phylogenetically similar (Caron et al., 2022; Llewelyn et al., 2023; Strydom et al., 2022). Through the integration and analysis of these diverse data sources, researchers can then enhance the understanding of interactions on a broader scale, bridging the gap caused by limited direct observations (Compson et al., 2018; Gravel et al., 2013; Maiorano et al., 2020). The set of interactions extracted or inferred from prior knowledge should be regarded as potential interactions, as they may not have been directly measured or observed *in situ*, or only observed in limited locations and time frames. Importantly, potential interactions do not necessarily assume any impact on species’ demography. All known potential interactions can be synthesized in a metanetwork, which represents a specific type of network that encompasses a set of species and all their potential interactions (Dunne, 2006).

The metanetwork concept holds a substantial value in the field of biogeography, as it can be both conceptually and operationally linked to the species pool and filtering framework of community ecology (Figure 1). In this regard, the metanetwork comprises all species from the regional pool of species and, in addition, contains information on all their potential interactions. Comparably to the filtering framework of community ecology, species can be filtered through
dispersal, abiotic and biotic filters into local multi-trophic networks akin to the local communities (Calderón-Sanou et al., 2021; Saravia et al., 2022). In practice, we can thus assume that locally co-occurring species passed all these filters and that they potentially interact as established in the metanetwork.

The metanetwork concept, when coupled with metacommunity theory, sheds light on the processes underlying the assembly of network structures across different spatial scales in biogeography (Figure 1). It allows for the examination of large-scale factors such as geographic area, co-evolution, immigration, and diversification, and their influence on the size and structure of local networks (O’Connor et al., 2020; Pugh & Field, 2022; Saravia et al., 2022). The metanetwork also serves as a baseline for potential interactions, enabling the construction of null models for local interaction networks. Departures from these null models provide valuable insights into the processes driving the formation of network structures from regional to local scales (Morlon et al., 2014). Moreover, the metanetwork concept offers the opportunity to integrate biotic interactions explicitly into biogeography theory. Incorporating biotic interactions into island biogeography models (Gravel et al., 2011; Massol et al., 2017), considering them as facets of diversity (Gaüzère et al., 2022), or incorporating them into scaling biodiversity relationships (Galiana et al., 2021) have advanced our understanding of how ecological communities are assembled across space.

Yet, it makes a fundamental assumption that local interactions are identical to those in the metanetwork, which ignores the variability of biotic interactions in response to different environmental conditions, landscape configuration, and contexts (Bimler et al., 2018; Michalet et al., 2006). Moreover, the combination of data from different sources may introduce biases, inconsistencies, or uncertainties, due to variations in methodologies, data quality, and varying sampling effort. Sensitivity analyses can help identify sources of uncertainty and enhance the accuracy and reliability of the metanetwork and subsequent analyses. By systematically varying
the presence or absence of specific links in the metanetwork, researchers can evaluate the robustness of their analyses and conclusions to the uncertainties in the network structure.

Section 4 - Ways forward in interaction biogeography

In the preceding three sections, we outlined three general classes of approaches to incorporate biotic interactions in biogeographical studies, and highlighted their respective advantages and limitations. Direct observations and experiments are most effective in capturing the temporal demographic dynamics resulting from biotic interactions, but can hardly be scaled up to biogeographic scales, species rich systems and multiple types of interactions. The statistical inference approach can reveal interactions between all species of a system by using data on species co-occurrences that are commonly available, but risks providing biased and potentially erroneous estimates of species interactions. The metanetwork approach is anchored in community assembly theory, is based on integrated expert knowledge and thus less prone to use unrealistic estimates of species interactions, but it only describes potential interactions, ignores interaction variability across environments, and does not allow modelling temporal demographic dynamics. Acknowledging the synergistic potential they hold, we conclude this outlook by presenting three avenues to strengthen our comprehension and prediction of large-scale biotic interactions, facilitating their seamless integration into conservation studies and risk assessments.

In the first section, we discussed how measurements of biotic interactions from observational and experimental studies have long been used to parametrize mathematical models of community dynamics. However, this approach faces challenges due to the non-linear increase in the interaction matrix with the number of species (Figure 1, limitation in taxonomic and spatial scales). To address this challenge and reduce the parameter space, a proposed solution is to consider that pairwise interactions are not independent but shaped by common
ecological drivers. Modeling these relationships can then be used to infer the pairwise interactions (Chalmandrier et al., 2022; Kissling et al., 2012). For instance, various approaches have been suggested for dimension reduction, such as incorporating metabolic theory and allometric scaling (Boit et al., 2012; Hudson & Reuman, 2013), considering species' functional traits (or phylogeny) and their role in demographic rates (Blonder et al., 2017; Chalmandrier et al., 2021). These dimension reduction techniques offer two key advantages: they can significantly reduce the parameter space, and they can provide ecological insights into the drivers of interactions. However, it was only recently that the combination of demographic Lotka-Volterra models, dimension reduction through trait relationships, and statistical models allowed the application of these ideas to species-rich systems (Figure 1, Chalmandrier et al., 2021). This approach allows estimating biotic interactions among numerous species with only a few parameters linking known traits to unknown biotic interaction coefficients. It can be used in conjunction with the metanetwork to estimate the strength of links among species using relevant functional traits (e.g., Danet et al., 2021; Ibanez et al., 2013). However, using mechanistic models to incorporate biotic interactions in biogeographical analyses has its limitations. Many models, like Lotka-Volterra models, are designed to mimic community dynamics at small spatial scales, neglecting how steep environmental gradients can alter the nature of interactions across biogeographical scales. To address this, more research is required to develop mechanistic models that account for changes in biotic interactions along environmental gradients, such as temperature (Armitage & Jones, 2020) or nutrient (Koffel et al., 2021).

In the second section, we explored the challenges of inferring biotic interactions from co-occurrence patterns due to numerous confounding factors (Blanchet et al., 2020; Poggiato et al., 2021). Nevertheless, the wealth of knowledge compiled in metanetworks offers valuable insights (Figure 1). Leveraging the regional-scale interaction network should allow using it as input for modelling, providing a deeper understanding of the role of biotic interactions in
shaping species communities and, ultimately, incorporating this information into species
distribution predictions (Staniczenko et al., 2017; Wisz et al., 2012). For instance, Ohlmann et
al. (2023) developed a model, based on Markov random fields (see also N. J. Clark et al., 2018),
to concurrently analyse the presence and absence of all species within a given area in function
of environmental covariates and the topological structure of the available metanetwork. This
model allows to capture the spatial variation of the influence of biotic and abiotic factors on
communities, and provides a first means to integrate network ecology into joint species
community modelling (Ohlmann et al., 2023). Another possibility is to explicitly account for
known prey (or predators) from the metanetwork to model the niches of predators and improve
their predictions (or prey, Poggiato et al., 2022). Such a model considers the direction of
interactions (top-down vs. bottom-up) and includes species' predators or prey as additional
covariates. Notably, the model aids in filtering potential metanetworks into the realized local
networks by removing statistically insignificant species interactions, thus refining the
metanetwork approach presented in Section 3. By measuring variable importance, the model
determines the relative significance of biotic interactions compared with environmental
filtering for different species, and uncovers the spatial patterns and environmental determinants
of biodiversity distribution (Poggiato et al., 2022).

Yet, the integration approach outlined above still relies on a correlative snapshot of co-
occurrence patterns to gauge the effect of biotic interactions on species distributions. Further
development could involve integrating diverse data sources providing explicit spatial and
temporal information, such as novel technology data (e.g., camera-traps, acoustic recorders,
environmental DNA), in situ and experimental data, and citizen science (Section 1, Hartig et
al., n.d.). Incorporating temporal data would allow models that account for biotic interactions
(Ohlmann et al., 2023; Poggiato et al., 2022) to be temporally explicit while utilizing known
information derived from metanetworks. Techniques like multivariate autoregressive models
and convergent cross-mapping could help extract the effect of biotic interactions from dynamic
data, akin to causal time series inference (Abrego et al., 2021; A. T. Clark et al., 2015; J. S. Clark et al., 2020; Lany et al., 2018; Thorson et al., 2016).

Integrating various data sources would also benefit from theoretical and experimental knowledge on the effect of biotic interactions on species dynamics, as described in the first section. Incorporating this knowledge into the modeling process would better constrain statistical predictions to align with theoretical and experimental expectations. Modelling frameworks that accommodate diverse data sources or even experiments, and prior information would be particularly valuable in achieving this task (Isaac et al., 2020; Talluto et al., 2016). While theory and models can be developed and tested using limited data or simulated scenarios, the refinement and applicability of these models depend on real-world, large-scale data, hence the need for filling gaps in species interactions data on a broad taxonomic and spatial scale (i.e., the Eltonian shortfall). Ultimately, a comprehensive and integrative approach holds tremendous potential for advancing our understanding of biotic interactions and their implications for species distributions and ecological dynamics (Hallam & Harris, 2023).

Acknowledgements: This collaborative endeavour is dedicated to the memory of Marc Ohlmann, whose untimely departure deeply touched us all. His spirit continues to inspire and guide us in our pursuit of understanding and studying ecological dependencies. We thank Laure Gallien for fruitful discussions about the different ways of measuring biotic interactions. This work was funded by the French National Research Agency (ANR) (EcoNet: ANR-18-CE02-0010-01, GAMBAS: ANR-18-CE02-0025, TransAlp: ANR-20-CE02-0021). WT, GP, and LOC also acknowledge the French Biodiversity Office and the HorizonEurope NaturaConnect project (No: 101060429) and LC the H2020 DRYvER project (No. 869226). P.G. was supported by Marie Curie Actions of the European Horizon 2020 under REA grant agreement (No.101026394, INDEBT).
References:


Global Change Biology, geb.16866. https://doi.org/10.1111/geb.16866


Figure caption:

Figure 1: Integrating three different approaches to advance interaction biogeography. 1) Field observations and experiments provide data on species interactions, to test and develop theory and mathematical formulations (green arrows). 2) Analysis of spatial co-occurrence data reveals statistical dependencies between species (light blue arrows). 3) From fundamental knowledge and data integration to metanetworks and local network realisations (purple arrows). 4) Based on the first three approaches, different integration pathways to advance the field of interaction biogeography. Through various icons, we highlighted the strengths and limitations of the different approaches concerning their ability to account for large taxonomic (i.e., species rich systems) and spatial scales and generality (potentially along steep environmental gradients), whether they rely or not on strong theory and concept, and whether they account for temporal dynamics.