Biotic interactions are fundamental drivers governing biodiversity locally, yet their effects on geographical variation in community composition (i.e. incidence-based) and community structure (i.e. abundance-based) at regional scales remain controversial. Ecologists have only recently started to integrate different types of biotic interactions into community assembly in a spatial context, a theme that merits further empirical quantification. Here, we applied partial correlation networks to infer the strength of spatial dependencies between pairs of organismal groups and mapped the imprints of biotic interactions on the assembly of pond metacommunities. To do this, we used a comprehensive empirical dataset from Mediterranean landscapes and adopted the perspective that community assembly is best represented as a network of interacting organismal groups. Our results revealed that the co-variation among the beta diversities of multiple organismal groups is primarily driven by biotic interactions and, to a lesser extent, by the abiotic environment. These results suggest that ignoring biotic interactions may undermine our understanding of assembly mechanisms in spatially extensive areas and decrease the accuracy and performance of predictive models. We further found strong spatial dependencies in our analyses which can be interpreted as functional relationships among several pairs of organismal groups (e.g. macrophytes–macroinvertebrates, fish–zooplankton). Perhaps more importantly, our results support the notion that biotic interactions make crucial contributions to the species sorting paradigm of metacommunity theory and raise the question of whether these biologically-driven signals have been equally underappreciated in other aquatic and terrestrial ecosystems. Although more research is still required to empirically capture the importance of biotic interactions across ecosystems and at different spatial resolutions and extents, our findings may allow decision makers to better foresee the main consequences of human-driven impacts on inland waters, particularly those associated with the addition or removal of key species.

Keywords: beta diversity, environmental filtering, metacommunity ecology, network analysis, species sorting, trophic guild
Introduction

A long-standing and overarching question in ecology is what governs the geographical variation of biodiversity on Earth (Logue et al. 2011). Interest in this topic grew in the 1970s when Diamond (1975) and Connor and Simberloff (1979) debated the assembly rules shaping the co-occurrence of certain species. Although modern ecology has largely moved away from these ideas (Gravel et al. 2019), scholars still widely embrace the central role of environmental determinism, biotic interactions, stochasticity and historical contingencies on community assembly (Heino et al. 2015). In particular, the study of beta diversity (i.e. spatial variation in community composition and community structure, Anderson et al. 2011) has become a foundational concept used to shed light on these assembly processes governing distributional patterns in the different realms of life (Chase and Myers 2011). As a consequence, recent decades have witnessed a remarkable increase in the number of studies examining beta diversity across multiple geographical and environmental gradients (see Mori et al. 2018 for a review). Central to this perspective is metacommunity theory (Logue et al. 2011), which has been a particularly useful framework for integrating the abiotic environment, biotic interactions and dispersal events as drivers of biological diversity across spatial scales (Heino et al. 2015).

One major weakness of beta diversity and metacommunity studies is that they often empirically ignore biotic interactions (e.g. competition, predation and facilitation), despite extensive theoretical evidence that such interactions produce clear spatial dependencies among different organismal groups (Poisot et al. 2016a, Ohlmann et al. 2018, Silknetter et al. 2020). Given this omission, it is perhaps not surprising that ecologists are still struggling to disentangle whether biotic interactions actually affect the geographical distributions of species (Wisz et al. 2013), resulting in fundamental gaps in our understanding of the assembly mechanisms structuring communities of interacting organismal groups (Gravel et al. 2019), particularly for taxa other than terrestrial plants (Pringle et al. 2016). This gap is particularly conspicuous for lentic ecosystems, where ecologists have focused far more on the abiotic control of community assembly than on complex interaction networks among locations (see Heino et al. 2015 for a comprehensive review), which is also reflected in current bioassessment and management policies (Heino 2013a). Additionally, accumulating evidence seems to support the importance of species sorting signals along geographical and environmental gradients in lentic systems (Heino 2013b, Alahuhta et al. 2018), suggesting that the structure of local communities is primarily determined by differences in species’ niches (Soininen 2014), and thus highlighting that the inclusion of biotic interactions may improve generalisation and statistical accuracy in predictive models of community assembly. This is because it is still unclear which part of the species sorting signal corresponds to species filtered by abiotic-mediated versus biotic-mediated constraints (Leibold et al. 2004). There is thus a need to expand ecological research into more realistic scenarios of community assembly by integrating a plethora of potential biotic interactions at the regional scale and under a range of environmental conditions among sites (Jabot and Bascompte 2012).

Recently, there has been a surge of interest in using network approaches to understand how biotic interactions among organismal groups drive the structure and distribution of biological communities (Zarnetske et al. 2017, Lee et al. 2019). However, such studies remain rare and have mostly been concerned with the distribution of multi-trophic interactions within locations (Holomuzki et al. 2010, Zhao et al. 2019), and less so with the geographical variation among locations in spatially extensive systems (but see Ohlmann et al. 2018, Gravel et al. 2019). Attempts to capture these interactions at the metacommunity level have often been restricted by both the difficulties in compiling comprehensive multi-trophic inventories and the expertise of individual scholars (Lee et al. 2019). Given that recent empirical and theoretical studies emphasise the spatial component of biotic couplings (Wisz et al. 2013), there is a need to bridge the gap between the beta diversity (i.e. metacommunity) perspective and the network approach. A relatively new probabilistic graphical model, the Graphical Lasso (Friedman et al. 2007, Mazumder and Hastie 2012), shows promise for addressing this challenge. For instance, this quantitative approach has also been used to model species interactions (Harris 2016) and can be applied to infer the strength of spatial dependencies between pairs of organismal groups without any a priori assumption on the overall structure of the interaction network (Ohlmann et al. 2018), thereby potentially expanding ecological research into more realistic scenarios of community assembly.

Here, we applied a type of Markov networks (i.e. partial correlation networks inferred using the Graphical Lasso) to a comprehensive empirical dataset of multiple organismal groups (i.e. aquatic macrophytes, phytoplankton, zooplankton, macroinvertebrates and fish, Table 1) and mapped the imprints of biotic interactions on the assembly of pond biotas. We used species occurrences and abundances to reflect variation in community composition and structure, respectively, and adopted the perspective that community assembly is best represented as a network of ecological interactions between pairs of organismal groups. Specifically, we hypothesised (H1) that the abiotic environment would be more important than biotic interactions for variation in community composition (D’Am en et al. 2018), supporting Grinnellian ideas (Chase and Leibold 2003) that species occurrences are primarily explained in terms of the resistance of biotas to prevailing abiotic environmental conditions. On the other hand, we predicted (H2) that variation in community structure would be more strongly shaped by the effects of biotic interactions among organismal groups (Rael et al. 2018), following the classical Eltonian view for the spatial variation of species abundances (Chase and Leibold 2003). In this context and given the structuring role of aquatic macrophytes and their potential to operate as foundation species and ecosystem engineers in ponds (Fernández-Aláez et al. 2018), we
expected (H3) that most animal communities would be bottom–up regulated by plant communities. However, we also predicted (H4) that predation pressure by top predators, such as fish and dragonfly larvae, would affect variation in community structure via cascading effects on non-predatory macroinvertebrates, zooplankton and microalgae (Carpenter et al. 2007). For simplicity, macroinvertebrate taxa were separated into three functional feeding categories following Tachet et al. (2002) and Schmidt-Kloiber and Hering (2015). Larvae and adults were considered separately when assigning functional feeding groups. We grouped fish species into two groups (i.e. small fish < 10 cm and big fish > 10 cm) according to their adult size class (Schmidt-Kloiber and Hering 2015). See Supplementary material Appendix 2 for more details.

**Material and methods**

**Study area and field data collection**

The study area, dataset characteristics and field data collection have been described previously (Trigal et al. 2014, García-Girón et al. 2019a), and we thus only highlight the specific geographical and ecological details in the main text. The study region is located within a heterogeneous area of approximately 50 000 km² in north-western Spain (Supplementary material Appendix 1) and has a Continental-Mediterranean climate with strong seasonal variations in temperature and precipitation (see García-Girón et al. 2019a for details). Pond types vary from small (area < 0.5 ha) agricultural waterbodies with high nutrient concentration and mineralisation to relatively big (area > 10 ha) forest-bordered ponds with low conductivity and nutrient content. Although the majority of sites displayed considerable variability in environmental conditions (Supplementary material Appendix 1), all the ponds we studied were shallow (mean depth < 2 m), experienced a strong reduction in water volume during the summer and suffered from various anthropogenic pressures, including water abstraction, nutrient enrichment and alien invasive species.

Between June and July of either 2004 or 2005, we sampled physico-chemical parameters and pond organismal groups (i.e. macrophytes, phytoplankton, zooplankton, macroinvertebrates and fish) in a total of 25 ponds (see Supplementary material Appendix 1–2 for further details). Physico-chemical parameters included mean depth (cm), Secchi depth (cm), pH, oxygen (mg l⁻¹), conductivity (µS cm⁻¹), turbidity (FTU), total nitrogen (mg l⁻¹), nitrate (mg l⁻¹), ammonium (µg l⁻¹), total phosphorus (µg l⁻¹), soluble reactive phosphorus (µg l⁻¹), total suspended solids (mg l⁻¹), volatile suspended solids (mg l⁻¹), dissolved organic carbon (mg l⁻¹) and chlorophyll ‘a’ (mg l⁻¹). We also measured pond area (ha) and catchment land cover (%) on images available at SIGPAC (Spanish Geographical Information System for Agricultural Parcels, <www.sigpac.jcyl.es/visor/>). Climate variables, including mean annual temperature (°C), annual temperature range (°C) and annual precipitation (mm), were obtained from the WorldClim 2.0 (Fick and Hijmans 2017) for each study pond based on a 30-yr average (1 km² resolution data).

**Calculating environmental distances among sites**

All environmental variables (except pH) were transformed to improve normality (logarithmically or logit transformed) and reduce skewness. The environmental distances among sites were then estimated with standardised Euclidean distances from the first two (composite and uncorrelated) axes of a principal component analysis (PCA) performed on all transformed environmental variables. Importantly, the first PCA axis (PCA1) was closely related with water chemistry.
(i.e. high positive loadings for nutrient content and chlorophyll 'a' and high negative loadings for Secchi depth and conductivity), whereas the second PCA axis (PCA2) roughly represented the catchment land cover (i.e. high positive loadings for cropland and high negative loadings for woodland, see Supplementary material Appendix 3 for more details). The above-mentioned environmental variables overshadowed climate gradients in the first two PCA axes.

**Measuring beta diversities of predefined organismal groups**

We calculated pairwise beta diversities (i.e. variation in community composition and community structure) for each predefined organismal group following Baselga (2010, 2013). To do this, we used the Sørensen (incidence-based) and Bray–Curtis (which is an abundance-based extension of the Sørensen index, Legendre and Legendre 2012) dissimilarity indices as incorporated in the R package betapart (Baselga et al. 2018).

**Accounting for spatial structures in the covariation of major organismal groups**

We ran Mantel correlograms using the R package vegan (Oksanen et al. 2019) to examine spatial structures in detail, i.e. to test if pairwise beta diversities are spatially autocorrelated within each distance class. The distance classes were determined by Sturge’s rule (Legendre and Legendre 2012) and p-values were based on 199 permutations with Holm correction for multiple testing (Holm 1979).

**Using the Graphical Lasso to unravel the joint spatial variation of metacommunities**

To represent a network that parsimoniously reflected the partial correlations between variation in community composition and community structure of multiple organismal groups, we applied the script of Ohlmann et al. (2018). In brief, Ohlmann et al. (2018) proposed a framework for inferring and plotting the strength of conditional (spatial) dependencies between pairs of organismal groups using a blockwise coordinate descent procedure for the Lasso (least absolute shrinkage and selection operator) regularisation approach (Tibshirani 1996), the Graphical Lasso (Friedman et al. 2007, Mazumder and Hastie 2012). While the Lasso method was originally developed to produce a suitable description of a parsimonious set of variables (Tibshirani 1996), the Graphical Lasso approach takes the advantage of the properties of Gaussian graphical models to efficiently infer a sparse network, allowing the representation of the conditional dependencies among multiple random variables in this network (here, the beta diversity patterns of multiple organismal groups and the environmental distances, but see Friedman et al. 2007). In other words, this method builds on observed beta diversity patterns to disentangle potential biotic and/or environmental effects on the spatial variation of metacommunities (see Fig. 1 for an illustrative example). To do this, the Graphical Lasso computes an empirical variance-covariance matrix $S$ and estimates a partial correlation matrix that quantifies the degree of relationship between pairs of variables conditional to the other variables (here, a $n \times n$ beta diversity or environmental distance matrix, $n$ being the number of ponds). The variance–covariance matrix $S$ is inverted to compute the precision matrix $P$ ($P=S^{-1}$). Importantly, the Graphical Lasso uses a penalty term in the likelihood (modulated by the $\lambda$ coefficient) to ensure the sparsity of the precision matrix $P$ (refer to Friedman et al. 2007 for computational and mathematical details). Following Foygel and Drton (2010), we used the extended Bayesian information criteria (BIC$_p$) to select an optimal $\lambda$ coefficient. The partial correlation matrix was subsequently calculated from the precision matrix $P$ following Eq. 1 in the R package qgraph (Epskamp et al. 2019):

$$
cor(x_i, x_j | x_{\{i,j\}}) = -\frac{P_{i,j}}{P_{i,i}P_{j,j}}
$$

where $cor(x_i, x_j | x_{\{i,j\}})$ is the partial correlation between the components $i$ and $j$ of a random variable $X$ given all the other components, and $P_{i,i}$, $P_{j,j}$ and $P_{i,j}$ are the elements of the precision matrix $P$, i.e. the partial correlations between the beta diversity of the predefined organismal groups and the standardised environmental (Euclidean) distances.

Following Ohlmann et al. (2018), we expected marginal correlations (i.e. Pearson correlations) to be less informative than partial correlations, not least because marginal correlations usually show confounding effects and spurious values. As the precision matrix $P$ was inverted with a penalty modulated by the $\lambda$ coefficient to ensure sparsity, the partial correlation matrix $cor(x_i, x_j | x_{\{i,j\}})$ was also sparse, allowing the representation of the relationships between the beta diversity of each species group and environmental distances in a Markov network (see Friedman et al. 2007, Mazumder and Hastie 2012 for details).

We represented conditional dependencies between pairs of organismal groups using partial correlation networks and calculated the unweighted and weighted degrees of each node (i.e. the predefined organismal groups and environmental distances) in each network (here, one network for variation in community composition and one network for variation in community structure). The unweighted degree of a node in a particular network represents its number of direct networks, whereas the weighted degree is the total sum of partial correlations between a given node and the other nodes that are directly connected to this group (Murphy 2012). Hence, the higher the sum of the weighted degree, the greater the interdependencies with the beta diversity of other organismal groups, i.e. the more connected a species group (or an environmental distance) is, the more influence it has on variation in community composition and community structure of other groups. Conversely, if beta diversities of two organismal groups are conditionally independent (i.e. has a partial correlation coefficient equals to zero),
they cannot affect each other (Murphy 2012). Note, however, that the Graphical Lasso does not directly infer the interaction network at the species level, and the conditional dependencies at the group level inferred here do not necessarily imply causality (Ohlmann et al. 2018).

Since the Graphical Lasso is expected to be sensitive to the effect of a missing predictor (Mazumder and Hastie 2012), we tested to what extent the addition of a new (initially missing) environmental component affected the structure of the empirical partial correlation networks (here, based on variation in community composition and Bray–Curtis dissimilarity coefficients). Then, we computed the variance–covariance and precision matrices and estimated partial correlations (c) that quantify the degree of relationship between pairs of variables conditional to the other variables. Finally, we represented conditional dependencies using partial correlation networks (d) and calculated the weighted degrees (e). Partial correlation networks and weighted degrees show that water chemistry had strong impacts on the community composition of the organismal groups, whereas aquatic hydrophytes were the most influential organismal group conditioning community structure. Circles’ sizes (c) are proportional to the strength of spatial dependencies between pairs of organismal groups and the environmental distance. Linkage strength (d) is proportional to partial correlation coefficients as represented in the heatmap chart (c). Null partial correlation coefficients are coloured in grey and partial correlations above the median value of the non-null partial correlation coefficients are coloured in blue. Silhouettes follow Fig. 2.

Figure 1. Schematic flow chart of the statistical routines used to unravel the joint spatial variation of metacommunities. In this example, we built a set of data by constructing a regional trophic web and one environmental (Euclidean) distance, i.e. water chemistry (a). The regional trophic web was assumed to have three different organismal groups (i.e. aquatic hydrophytes, filter-feeding zooplankton and predatory macroinvertebrates) and follow the first two hypotheses of the main text. We calculated pairwise beta diversities (here, a \( n \times n \) beta diversity or environmental distance matrix, \( n \) being the number of ponds, \( p_i \)) for each predefined organismal group using the Sørensen (i.e. variation in community composition) and Bray–Curtis (i.e. variation in community structure) dissimilarity coefficients (b). Then, we computed the variance–covariance and precision matrices and estimated partial correlations (c) that quantify the degree of relationship between pairs of variables conditional to the other variables. Finally, we represented conditional dependencies using partial correlation networks (d) and calculated the weighted degrees (e). Partial correlation networks and weighted degrees show that water chemistry had strong impacts on the community composition of the organismal groups, whereas aquatic hydrophytes were the most influential organismal group conditioning community structure. Circles’ sizes (c) are proportional to the strength of spatial dependencies between pairs of organismal groups and the environmental distance. Linkage strength (d) is proportional to partial correlation coefficients as represented in the heatmap chart (c). Null partial correlation coefficients are coloured in grey and partial correlations above the median value of the non-null partial correlation coefficients are coloured in blue. Silhouettes follow Fig. 2.
Results

Synthesising the structure of the empirical networks

Nearly all non-null partial correlations estimated between beta diversities (i.e. variation in community composition and community structure) of each predefined organismal group and environmental distances were positive (Supplementary material Appendix 5). The estimated partial correlation network for variation in community composition (Fig. 2a) had a connectance of 0.13 and was composed of 12 undirected edges (i.e. partial correlation coefficients > 0) out of 91 possible edges and 14 nodes (12 organismal groups and 2 environmental distances). The overall network structure for variation in community structure (Fig. 2b) had the same number of nodes (14 nodes), was defined by 11 undirected edges (out of 91 possible edges) and had a connectance of 0.12. The maximum values of the non-null partial correlation coefficients for both community composition and community structure networks were 0.38 and 0.45, whereas the mean values were 0.16 and 0.18, respectively.

Community assembly is primarily driven by biotic interactions

Filter-feeding zooplankton (unweighted degree = 4, weighted degree = 0.64) and big fish (unweighted degree = 3, weighted degree = 0.64) were the most influential organismal groups structuring the community composition of other groups (i.e. highest weighted degree values, Fig. 3a). Water chemistry (first PCA axis, unweighted degree = 2, weighted degree = 0.49) and aquatic hydrophytes (unweighted degree = 4, weighted degree = 0.45) also had strong effects on the community composition of the remaining organismal groups, as did predatory macroinvertebrates (unweighted degree = 3, weighted degree = 0.41) and small fish (unweighted degree = 1, weighted degree = 0.38). Specifically, variation in the community composition of aquatic hydrophytes led to a change in predatory macroinvertebrate communities, whereas the beta diversity of fish was strongly correlated with that of filter-feeding zooplankton. Predatory macroinvertebrates were further associated with compositional variation of detritivorous macroinvertebrates, and the two fish groups were also correlated.

Figure 2. Undirected partial correlation networks inferred using the Graphical Lasso between variation in community composition (a) and community structure (b) of the organismal groups and environmental distances. Each node represents the beta diversity of a species group or an environmental distance. Colours indicate partial correlation coefficients as represented in Fig. 1 and Supplementary material Appendix 5. Linkage strength is proportional to the value of the partial correlation coefficient.
Conversely, water chemistry variation only affected aquatic hydrophytes and filter-feeding zooplankton (Fig. 2a).

Predatory macroinvertebrates (unweighted degree = 3, weighted degree = 0.6) and helophytes (unweighted degree = 3, weighted degree = 0.48) were the most relevant organismal groups conditioning community structure in the study ponds (Fig. 3b). Scraping macroinvertebrates (unweighted degree = 1, weighted degree = 0.45) and detritivorous macroinvertebrates (unweighted degree = 1, weighted degree = 0.45), as well as filter-feeding zooplankton (unweighted degree = 4, weighted degree = 0.38) and hydrophytes (unweighted degree = 2, weighted degree = 0.33), were also likely to structure the beta diversity of the remaining organismal groups. Conversely, environmental distances had a small impact on variation in the community structure of the organismal groups (unweighted degree of 1 for both water chemistry – PCA1 and land use – PCA2, whereas the mean unweighted degree of the partial correlation network was 1.6; weighted degrees of 0.18 and 0.19 for the first and second PCA axes, respectively, whereas the mean value was 0.26). Importantly, we found strong direct associations between vegetation structure and abundance variations of both predatory macroinvertebrates and filter-feeding zooplankton. Similarly, the beta diversity of scraping macroinvertebrates was strongly related to variation in the community structure of detritivorous macroinvertebrates, whereas the correlation between the two fish groups remained significant. On the other hand, the abiotic environment was important for the community structure of filter-feeding zooplankton and big fish communities (Fig. 2b).

The probability of observing a non-null partial correlation between both community composition and community structure and the environmental features was 0.08, whereas the probabilities of observing a link between the beta diversity of any two organismal groups were 0.15 and 0.14, respectively. Since the variables associated with disconnected nodes were conditionally independent, these results re-emphasise the weaker influence of the abiotic environment on metacommunity organisation.

Adding a new (initially missing) environmental distance (PCA3) did not change the edges between different organismal groups (Supplementary material Appendix 6). The partial correlation networks built using two or three environmental variables had βWN values between 0.04 (for community composition, βWN = 0.17 if only the edges which the weights are above the median value of partial correlation coefficients are considered) and 0.05 (for community structure, βWN = 0.4 if only the edges which the weights are above the median value of partial correlation coefficients are considered), whereas the dissimilarity arising from edges’ turnover in the shared parts of the two networks (βOS) equalled to 0 for both community composition and community structure (βOS = 0.09 if only the edges which the weights are above the median value of partial correlation coefficients are considered). This means that the overall dissimilarity in both networks was essentially due to the edges between the node representing the Euclidean distance matrix from the third environmental PCA axis. Perhaps more importantly, neither randomly resampling the sites (Supplementary material Appendix 4), nor adding new environmental distances to the statistical routine (Supplementary

![Figure 3](image-url). Properties of the inferred networks for variation in community composition (a) and community structure (b) of the organismal groups. The unweighted degree is the number of neighbours of nodes in a plot (here, the undirected correlation networks in Fig. 2). It measures the number of variables that are conditionally dependent on the variable associated with this node. The weighted degree is the total sum of partial correlations between a given node and the other nodes that are directly connected to this group. Dashed lines represent the mean values of unweighted and weighted degrees.
material Appendix 6), compromised our core finding that biotic interactions had a greater influence than the abiotic environment on metacommunity organisation in our study ponds. Finally, we detected only weak, yet sometimes significant spatial autocorrelation in the variation of community composition and community structure of major organisinal groups (Supplementary material Appendix 7).

Discussion

Ecologists have only recently started to integrate different types of biotic interactions within the metacommunity framework (Borthagaray et al. 2014, Ohlmann et al. 2018), a theme that urgently deserves further empirical attention and quantification. Network and multiscale approaches are helping us to fill this gap of knowledge by determining the spatial scale at which biotic interactions are important (e.g. local versus regional), and whether such interactions structure metacommunities (Poisot et al. 2016a). Here, we continued on this path and applied the Graphical Lasso for dissecting the joint spatial structure of multiple organismal groups and the abiotic environment in metacommunity organisation. To the best of our knowledge, the present study represents the first analysis of a multi-trophic dataset in pond ecosystems from both network and metacommunity perspectives. Our results indicated that pairwise environmental distances display low correlations with variation in both community composition and community structure (as measured with the few non-zero partial correlations). This finding refuted our first hypothesis (H1) that species occurrences are primarily explained in terms of the resistance of biotas to prevailing abiotic environmental conditions and, at the same time, confirmed our second expectation (H2) that variation in community structure follows the classical Eltonian view for the spatial variation of species abundances. Similarly, our results revealed that variation in the composition and structure of aquatic macrophyte communities led to a change in both zooplankton and macroinvertebrate communities, partially supporting our third hypothesis (H3) that most animal groups would be bottom-up regulated by macrophytes. Unexpectedly (H4), we found little evidence for cascading effects on phytoplankton. Perhaps most importantly, our integrative framework supported the notion that biotic constraints make crucial contributions to metacommunity organisation in ponds.

Moving beyond the abiotic environment in the species sorting paradigm

Although past studies have provided a comprehensive overview of the abiotic factors driving metacommunity structure in lentic ecosystems (see Heino et al. 2015 for a synthesis), these have failed to empirically test potential biotic interactions, thereby entirely missing a critical aspect of community assembly. Ideally, both the abiotic environment and biotic interactions should have been included in a single study, but their ability to assess which part of the niche-based signal corresponds to species sorting by environmental versus biotic constraints was limited by the statistical methods available at that time. This apparent omission partly reflects the limitations of the variation partitioning models that have hitherto been applied in modern ecology (Borcard et al. 1992), as the fraction characterising the individual contribution of biotic interactions is typically missing from this analytical approach (Soininen 2014).

Alternatively, the Graphical Lasso has been designed to address these limitations. We showed here how it accounts for conditional dependencies among multiple organismal groups and their responses to environmental variation in pond biotas, emphasising the need to go beyond the traditional view of understanding the abiotic environment as the main force of the species sorting paradigm of metacommunity theory. Importantly, our results are broadly in line with several recent studies encompassing a variety of organisms in aquatic ecosystems. For example, Zhao et al. (2019) found that biotic attributes (potentially reflecting biotic interactions) were pivotal in driving the local diversity of other aquatic organisms along a depth gradient in a Chinese lake, whereas Law et al. (2019) suggested that macrophyte richness was an effective surrogate for molluscan, beetle and dragonfly richness in British lentic ecosystems. However, these studies only used biotic predictors as surrogates for local biotic constraints and did not explicitly incorporate network interactions in a spatial context in their analyses. More recently, Pecuchet et al. (2020) also highlighted the importance of integrating functional community dynamics across multiple trophic levels in marine ecosystems. Today, no general consensus has been reached on whether the species sorting signal in metacommunities primarily corresponds to environmental filtering or biotic interactions (Borthagaray et al. 2014). Our results seem to support the central role of biotic constraints for the species sorting paradigm and raise the question of whether species interactions have been equally underappreciated in classical community-based assessments of aquatic and terrestrial ecosystems. We therefore strongly believe that ignoring biotic interactions, even if they cannot be explicitly examined using conventional statistical tools, may compromise our understanding on the niche-based control of metacommunities and decrease the accuracy and performance of predictive models. New investigations including both network and metacommunity analyses will be required to empirically capture these interactions across ecosystems and at different spatial resolutions and extents.

The Graphical Lasso is expected to be sensitive to the effect of a missing predictor since the structure of the network may be affected by the addition of a new environmental variable (Ohlmann et al. 2018). Our results showed that the addition of an initially missing environmental predictor had a minor effect on network structure, thereby guaranteeing the robustness of the analyses. Other missing environmental variables cannot be fully ignored, and we acknowledge that some missing factors might explain the lower than expected predictive power of environmental distances (e.g. seasonal phenology and drought, Chase 2007, Fernández-Aláez et al. 2009).
2020). However, we consider missing environmental variables a minor problem, as we surveyed a large set of water chemistry, land use and climate predictors previously found to be important correlates of freshwater assemblages in general (Heino et al. 2015, Lindholm et al. 2020). By contrast, our findings (Supplementary material Appendix 7) suggest that covariation among pairs of some organismal groups (e.g. macrophytes, filter-feeding zooplankton, predatory macroinvertebrates) was somehow influenced by the spatial autocorrelation in community data. Since spatial patterns in the distribution of species may be related to dispersal influences (Dray et al. 2012), it is hence possible that the source of any pure spatial structure detected here may stem from the effects of dispersal on pond metacommunities, as has been suggested by recent studies in Mediterranean environments (García-Girón et al. 2019a, b).

Unravelling the empirical biotic controls of pond metacommunities

The number of interactions among organismal groups, measured as network connectance, was similar to the only comparable information from lakes and streams in the Iberian Plateau (Sánchez-Hernández et al. 2015), but somewhat higher than that for alpine lakes in America (Harper-Smith et al. 2005), highlighting that the network structure of high-mountain lakes may be relatively simple in comparison with shallow, lowland ponds. Perhaps more importantly, we found strong spatial dependencies in our statistical models which can be interpreted as functional relationships among several pairs of organismal groups.

The Graphical Lasso detected that variation in the composition and structure of macrophyte communities resulted in a change in both filter-feeding zooplankton and predatory macroinvertebrate assemblages. These links most likely reflect the strong direct associations between these invertebrate groups and plant architecture, since submerged hydrophytes provide a refuge from predators and a variety of food sources and habitats (Lacoul and Freedman 2006, Heino 2008). Indeed, beetles can benefit from habitat heterogeneity for egg-laying and through increased prey availability (Law et al. 2019), whereas dragonfly larvae use hydrophytes for foraging and shelter (Goertzen and Suhling 2013). Moreover, the relationships we uncovered between helophytes and predatory macroinvertebrates may be attributed to the fact that adult odonates use helophyte leaves for emergence, perching and egg-laying (Le Gall et al. 2018). We also highlighted that the beta diversity of fish was strongly correlated with that of filter-feeding zooplankton, showing that fish predation may play an important role in structuring daphniid assemblages. This observation is consistent with a large body of literature emphasising the impact of predation in aquatic ecosystems (Carpenter et al. 2001) but, in contrast to temperate and boreal lakes (Jeppesen et al. 2003), these trophic effects did not cascade down to phytoplankton (Fig. 2). This lack of a cascading effect on phytoplankton may be partly explained by the fact that small rotifers and nauplii, which feed on small bacteria, ciliates and flagellates, dominated the regional zooplankton metacommunity (Trigal et al. 2014). Interestingly, we also found evidence for relationships among water chemistry, filter-feeding zooplankton and submerged hydrophytes, which reflects the well-known environmental filtering processes structuring lentic metacommunities in these landscapes (García-Girón et al. 2018, 2019a). Finally, our analyses also showed strong partial correlations between the beta diversities of different groups of fishes (i.e. small fish with big fish) and macroinvertebrates (i.e. detritivores with scrapers), suggesting that these organismal groups may experience competition and/or predation. For fishes, this link might be explained by a combination of competition between size or age classes and asymmetric intraguild predation (Steinmetz et al. 2008), whereas the strong direct association between scraping and detritivorous macroinvertebrates could be attributed to spatially strong exploitative effects usually contributed by body size and location ‘ownership’ status (Holomuzki et al. 2010).

This study is highly unique because the assessment of the geographical variation of biotic interactions and their influence on community assembly is almost entirely neglected in freshwater ecosystems. Importantly, our results emphasise the importance of the spatial dependencies between pairs of organismal groups and highlight some important functional associations (e.g. macrophytes–macroinvertebrates, fish–zooplankton) for the assembly of pond metacommunities. Further, our findings empirically illustrate the information gained from incorporating network models in metacommunity analyses, suggesting that integrating explicit biotic interactions and abiotic factors is necessary to accurately understand metacommunity organisation at the regional scale. We therefore anticipate more research to consolidate knowledge on the relationships of several organismal groups (e.g. feeding guilds) across communities and their associated feedbacks in different ecosystems where comprehensive biodiversity assessments are becoming available. This will be especially important to foresee the main consequences of human-driven impacts on natural ecosystems, and particularly those associated to the addition or removal of key species (Poisot et al. 2012). In combination, these studies should be able to quantify the couplings between biotic and abiotic drivers in landscapes where ecological knowledge is still limited, improve possibilities for generalisation and statistical accuracy, and integrate biologically-driven factors for decision making in environmental management and conservation.

Data availability statement

Data are available from the FigShare Digital Repository: <doi.org/10.6084/m9.figshare.11815434> (García-Girón et al. 2020).

Acknowledgements – We would like to thank all people involved in the field sampling and analysis of the samples, especially Margarita Fernández-Aláez.
Author contributions – JGG conceived the ideas with inputs from JH and JA. JGG carried out all analyses and lead the writing of the manuscript. JH and JA contributed to writing by commenting on the manuscript. FGC and CFA provided the data and valuable background information, which enabled conducting this study. All authors contributed critically to the final draft and gave final approval for publication.

Conflicts of interest – The authors declare no competing financial interests.

References

Alahuhta, J. et al. 2018. Global patterns in the metacommunity structuring of lake macrophytes: regional variations and driving factors. – Oecologia 188: 1167–1182.

Anderson, M. J. et al. 2011. Navigating the multiple meanings of β diversity: a roadmap for the practicing ecologist. – Ecol. Lett. 14: 19–28.

Barnett, A. J. et al. 2007. Functional diversity of crustacean zooplankton communities: towards a trait-based classification. – Freshwater Biol. 52: 796–813.

Baselga, A. 2010. Partitioning the turnover and nestedness components of beta diversity. – Global Ecol. Biogeogr. 19: 134–143.

Baselga, A. 2013. Separating the two components of abundance-based dissimilarity: balanced changes in abundance vs. abundance gradients. – Methods Ecol. Evol. 4: 552–557.

Baselga, A. et al. 2018. betapart package. – R package ver. 1.5.1, <https://cran.r-project.org/web/packages/betapart/index.html>.

Borcard, D. et al. 1992. Partialling out the spatial component of ecological variation. – Ecology 73: 1045–1055.

Borthagaray, A. I. et al. 2014. Inferring species roles in metacommunity structure from species co-occurrence networks. – Proc. R. Soc. B 281: 20141425.

Carpenter, S. R. et al. 2001. Trophic cascades, nutrients and lake productivity: whole-lake experiments. – Ecol. Monogr. 71: 163–186.

Chase, J. M. 2007. Drought mediates the importance of stochastic community assembly. – Proc. Natl Acad. Sci. USA 104: 17430–17434.

Chase, J. M. and Leibold, M. A. 2003. Ecological niches: linking classical and contemporary approaches. – Univ. Chicago Press.

Chase, J. M. and Myers, J. A. 2011. Disentangling the importance of ecological niches from stochastic processes across scales. – Phil. Trans. R. Soc. B 366: 2351–2363.

Cirujano, S. et al. 2014. Flora acuática española. Hidrófitos vasculares. – Real Jardín Botánico.

Connor, E. F. and Simberloff, D. 1979. The assembly of species communities: chance or competition? – Ecology 60: 1132–1140.

D’Amen, M. et al. 2018. Disentangling biotic interactions, environmental filters and dispersal limitation as drivers of species co-occurrence. – Ecography 41: 1233–1244.
Heino, J. et al. 2015. Metacommunity organisation, spatial extent and dispersal in aquatic systems: patterns, processes and prospects. – Freshwater Biol. 60: 845–869.
Holm, S. 1979. A simple sequentially rejective multiple test procedure. – Scand. J. Stat. 6: 65–70.
Holomuzki, J. R. et al. 2010. Biotic interactions in freshwater benthic habitats. – J. N. Am. Benthol. Soc. 29: 220–244.
Jabot, F. and Bascompte, J. 2012. Bitrophic interactions shape biodiversity in space. – Proc. Natl Acad. Sci. USA 109: 4521–4526.
Jeppesen, E. et al. 2003. Impacts of climate warming on lake fish community structure and potential effects on ecosystem function. – Hydrobiologia 646: 73–90.
Lacoul, B. and Freedman, B. 2006. Environmental influences on aquatic plants in freshwater ecosystems. – Environ. Rev. 14: 89–136.
Law, A. et al. 2019. The effectiveness of aquatic plants as surrogates for wider biodiversity in standing fresh waters. – Freshwater Biol. 64: 1664–1575.
Le Gall, M. et al. 2018. Determinant landscape-scale factors on pond odonate assemblages. – Freshwater Biol. 63: 306–317.
Lee, C. K. et al. 2019. Biotic interactions are an unexpected yet critical control on the complexity of an abiotically driven polar ecosystem. – Commun. Biol. 2: 62.
Legendre, P and Legendre, L. 2012. Numerical Ecology. – Elsevier.
Leibold, M. A. et al. 2004. The metacommunity concept: a framework for multi-scale community ecology. – Ecol. Lett. 7: 601–613.
Lindholm, M. et al. 2020. No biotic homogenisation across decades but consistent effects of landscape position and pH on macrophyte communities in boreal lakes. – Ecography 43: 294–305.
Logue, J. B. et al. 2011. Empirical approaches to metacommunities: a review and comparison with theory. – Trends Ecol. Evol. 26: 482–491.
Mazumder, R. and Hastie, T. 2012. The graphical lasso: new insights and alternatives. – Electron. J. Stat. 6: 2125–2149.
Mori, A. S. et al. 2018. β-diversity, community assembly and ecosystem functioning. – Trends Ecol. Evol. 33: 549–564.
Murphy, K. P. 2012. Machine learning: a probabilistic perspective. – The MIT Press.
Ohlmann, M. et al. 2018. Mapping the imprint of biotic interactions on β-diversity. – Ecol. Lett. 21: 1660–1669.
Oksanen, J. et al. 2019. Vegan: community ecology package. – R package ver. 2.5.6, <https://cran.rproject.org/web/packages/vegan/index.html>.
Pechuet, L. et al. 2020. Spatio-temporal dynamics of multi-trophic communities reveal ecosystem-wide functional reorganization. – Ecography 43: 197–208.

Supplementary material (available online as Appendix ecog-05032 at <www.ecography.org/appendix/ecog-05032>). Appendix 1–7.