NICHE AND DISPERSAL EFFECTS ON BIODIVERSITY, COMMUNITY COMPOSITION, AND SPECIES CO-OCCURRENCE IN STREAM ECOSYSTEMS

by

Joseph Leonardo Mruzek

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy at
The University of Texas at Arlington
August 2022

Arlington, Texas

Supervising Committee:

Sophia I. Passy, Supervising Professor
Luke O. Frishkoff
James P. Grover
Chad A. Larson
Robert F. McMahon

Acknowledgments

First and foremost, I want to thank my supervisor, Sophia Passy, for her guidance, advice, patience, and for seeing the potential in me as a scientist. I am proud of the work we have done, and as while we are both preparing to move on from UTA, I hope you continue to involve me, as your ideas, while not always “quick”, are always exciting. I also am thankful to my committee members, past and present: John Campbell, Luke Frishkoff, James Grover, Chad Larson, and Robert McMahon, who have reviewed my work and provided supervision over the past five years. I also want to thank my collaborator, Thibault Leboucher, and undergraduate mentee Dinh Khan Luc, whose willingness to assist with coding or to do tedious data curation I am indebted. Thank you to Laura Mydlarz and her lab, especially to Brad Dimos, and Nick MacKnight, for providing me a break from freshwater. Thank you to all my friends who have supported me over the last 5 years, both in the department, Cristian Hernandez, Kaitlyn Howell, Justin Jacobs, Thor Larson, and Alex Murray, and outside it, Justin Benzino and Gunnlaugur ‘Jay’ Júlíusson for keeping me sane while coding and writing. A special thanks to my lab mates, Katrina Pound, Will Budnick and Mikaela Hanson, whose encouragement, patience, and help kept me going. Finally, I thank my past advisors, Jeff Johansen, James Watling, Rick Relyea, Aaron Stoler, and Kayla Coldsnow, I truly stand on the shoulders of giants.

Abstract

NICHE AND DISPERSAL EFFECTS ON BIODIVERSITY, COMMUNITY COMPOSITION, AND SPECIES CO-OCCURRENCE IN STREAM ECOSYSTEMS

Joseph L. Mruzek

The University of Texas at Arlington

Supervising Professor: Sophia I. Passy, PhD

Niche and dispersal processes are major ecological drivers of freshwater communities but their relative roles across latitudinal gradients, scales, and organismal groups are not well understood. Thus, we performed three studies with sub-continental datasets of stream algae, insects, and fish. In chapter 2 we predicted that: i) niche processes would dominate at high latitudes due to increased climatic stress, consistent with the physiological tolerance hypothesis and the Dobzhansky-MacArthur hypothesis and ii) dispersal limitation would prevail at low latitudes due to narrower niches and smaller range sizes, consistent with the dispersal-ecological specialization tradeoff hypothesis, the latitude-niche breadth hypothesis, and Rapoport's rule. Chapter 3 had two goals: i) compare the relative roles of niche and dispersal processes on metacommunity composition vs. topology of co-occurrence networks and ii) determine if the taxonomic relatedness between co-occurring taxa could be explained by shared niche preference or dispersal history. Chapter 4 investigated the role of spatial scale and dispersal ability on the relationship of local species richness (LSR) with regional species richness (RSR). A saturated LSR-RSR relationship indicates local control of local richness, whereas an unsaturated relationship indicates regional control through dispersal limitation. We predicted i) saturation at small scales due to habitat filtering, but a lack of saturation at large scales due to dispersal limitation and ii)

that the less dispersive fish would shift from a saturated to unsaturated LSR-RSR relationship at a smaller scale than diatoms or insects. In summary, our findings contribute to community ecology by advancing our knowledge of the relative roles of niche and dispersal processes in controlling biodiversity, community composition, and network topology along environmental and spatial gradients.

Table of Contents

List of Tables

List of Figures

Chapter 1

General introduction

Niche and dispersal processes are the main ecological drivers of biodiversity, community composition, and species co-occurrences (Weiher & Keddy, 2004). However, their relative roles are not well understood along latitudinal gradients, on co-occurrence networks, and across scales. Therefore, we present here three original research chapters: First examining latitudinal dependency of these two processes, second, the relative importance of these two processes on co-occurrence network topology, and third examining scale through the lens of the regional species richness – local species richness relationship. In all three chapters we utilize a subcontinental dataset of stream algae, insects, and fish, which are ecologically important groups, and present a gradient in body size and therefore dispersal ability.

In the first chapter, we address the question of whether the effect of niche and dispersal processes on biodiversity varies with latitude. Climatic stress has a positive latitudinal dependence, with harsh winters and high seasonality at the high latitudes posing as a strong environmental filter (Hillebrand, 2004). Multiple climate-based ecological theories, including the Dobzhansky-MacArthur hypothesis (Brown, 2014), physiological tolerance hypothesis (PTH) (Currie *et al.*, 2004), have predicted stronger environmental effects on species composition at high latitudes due to climatic stress. Dispersal limitation also has a latitudinal dependency, as described in the dispersal-ecological specialization tradeoff hypothesis (Jocque *et al.*, 2010), the latitude-niche breadth hypothesis (Vázquez & Stevens, 2004), and Rapoport's rule (Stevens, 1989), which predicts not only will filtering be stronger at high latitudes but also that stronger dispersal effects at low latitudes will be present due to ecological specialization and smaller range sizes. In this chapter we present a test of these hypotheses by assessing the relative importance of niche versus dispersal processes in three stream communities, algae, insects, and fish, along a latitudinal gradient, predicting strong habitat filtering in higher latitudes, and strong dispersal limitation in lower latitudes.

In the second chapter we compare the drivers of metacommunity composition to those of co-occurrence network topology. Co-occurrence networks are collections of nodes, which represent taxa, and edges, which represent the significant pairwise correlations between taxa (Karimi *et al.*, 2017). Co-occurrences are driven mainly by niche processes but also dispersal limitation, as species with similar environmental preferences likely co-exist in the same habitat, or they can co-occur due to shared dispersal history (Bar-Massada *et al.*, 2018). In aquatic organisms, dispersal ability is inversely related to body size (De Bie *et al.*, 2012; Padial *et al.*, 2014), which we predicted would impact metacommunity structure and network topology similarly. Thus, we predicted that dispersal limitation would have a stronger impact on both the overall metacommunity composition, and network topology of relatively larger taxa (i.e., fish), while niche processes would be more impactful for metacommunity composition, and network topology, of smaller organisms (i.e., diatoms and insects). As we expect most co-occurrences to be explained by either niche or dispersal overlap, we asked if these processes would lead to taxonomic similarity due to phylogenetic niche conservatism (Wiens *et al.*, 2010; HilleRisLambers *et al.*, 2012; Lentendu & Dunthorn, 2021). We predicted that taxa whose co-occurrences were found to be explainable through niche or dispersal overlap, would be more closely related that those taxa whose co-occurrences we could not explain, consistent with phylogenetic niche conservation. This would demonstrate a link between species traits, and the mechanism driving co-occurrence.

Finally, the spatial scale of a study can affect the relative importance of niche and dispersal processes, which we analyzed through the lens of the regional species richness (RSR) - local species richness (LSR) relationship. Classic theory posits that a curvilinear RSR-LSR relationship represents a saturated local richness, moderated by local processes, while a linear relationship, is indicative of stronger regional effects, specifically dispersal limitation (Terborgh & Faaborg, 1980; Ricklefs, 1987). The RSR-LSR relationship also depends on spatial scale of the region and species dispersal capacity. We varied regional size, predicting a saturated LSR-RSR relationship at small regional scales and in poor dispersers, while a linear relationship is expected at larger regional scales and in good dispersers (Hillebrand & Blenckner, 2002; Soininen *et al.*, 2009). However, how scale and species group jointly control the RSR-LSR relationship is not well understood. We therefore hypothesized that as regional scale increases, the RSR-LSR relationship would shift from saturated (curvilinear) to unsaturated (linear), due to an increasing roll of regional richness and a decreasing role of environment in constraining local richness. Second, we hypothesized that fish, which are relatively dispersal limited, would shift from saturated to unsaturated at a smaller scale than diatoms and insects.

References

Bar-Massada, A., Yang, Q., Shen, G. & Wang, X. (2018) Tree species co-occurrence patterns change across grains: Insights from a subtropical forest. *Ecosphere*, **9**, 1–12.

De Bie, T., De Meester, L., Brendonck, L., Martens, K., Goddeeris, B., Ercken, D., Hampel, H., Denys, L., Vanhecke, L., Van der Gucht, K., Van Wichelen, J., Vyverman, W. & Declerck, S.A.J. (2012) Body size and dispersal mode as key traits determining metacommunity structure of aquatic organisms. *Ecology Letters*, **15**, 740–747.

Brown, J.H. (2014) Why are there so many species in the tropics? *Journal of Biogeography*, **41**, 8–22.

Currie, D.J., Mittelbach, G.G., Cornell, H. V., Field, R., Guégan, J.F., Hawkins, B.A., Kaufman, D.M., Kerr, J.T., Oberdorff, T., O'Brien, E. & Turner, J.R.G. (2004) Predictions and tests of climate-based hypotheses of broad-scale variation in taxonomic richness. *Ecology Letters*, **7**, 1121–1134.

Hillebrand, H. (2004) On the generality of the latitudinal diversity gradient. *The American Naturalist*, **163**, 192–211.

Hillebrand, H. & Blenckner, T. (2002) Regional and local impact on species diversity - from pattern to processes. *Oecologia*, **132**, 479–491.

HilleRisLambers, J., Adler, P.B., Harpole, W.S., Levine, J.M. & Mayfield, M.M. (2012) Rethinking community assembly through the lens of coexistence theory. *Annual Review of Ecology, Evolution, and Systematics*, **43**, 227–248.

Jocque, M., Field, R., Brendonck, L. & De Meester, L. (2010) Climatic control of dispersal-ecological specialization trade-offs: A metacommunity process at the heart of the latitudinal diversity gradient? *Global Ecology and Biogeography*, **19**, 244–252.

Karimi, B., Maron, P.A., Chemidlin-Prevost Boure, N., Bernard, N., Gilbert, D. & Ranjard, L. (2017) Microbial diversity and ecological networks as indicators of environmental quality. *Environmental Chemistry Letters*, **15**, 265–281.

Lentendu, G. & Dunthorn, M. (2021) Phylogenetic relatedness drives protist assembly in marine and terrestrial environments. *Global Ecology and Biogeography*, **30**, 1532–1544.

Padial, A.A., Ceschin, F., Declerck, S.A.J., De Meester, L., Bonecker, C.C., Lansac-Tôha, F.A., Rodrigues, L., Rodrigues, L.C., Train, S., Velho, L.F.M. & Bini, L.M. (2014) Dispersal ability determines the role of environmental, spatial and temporal drivers of metacommunity structure. *PLoS ONE*, **9**, 1–8.

Ricklefs, R.E. (1987) Community diversity: relative roles of local and regional processes. *Science*, **235**, 167–171.

Soininen, J., Heino, J., Kokocinski, M. & Muotka, T. (2009) Local-regional diversity relationship varies with spatial scale in lotic diatoms. *Journal of Biogeography*, **36**, 720–

727.

Stevens, G.C. (1989) The latitudinal gradient in geographical range: How so many species coexist in the tropics. *The American Naturalist*, **133**, 240–256.

Terborgh, J.W. & Faaborg, J. (1980) Saturation of bird communities. *The American Naturalist*, **116**, 178–195.

Vázquez, D.P. & Stevens, R.D. (2004) The latitudinal gradient in niche breadth: Concepts and evidence. *The American Naturalist*, **164**.

Weiher, E. & Keddy, P.A. eds. (2004) *Ecological Assembly Rules Perspectives, advances, retreats*, Cambridge University Press, Cambridge, UK.

Wiens, J.J., Ackerly, D.D., Allen, A.P., Anacker, B.L., Buckley, L.B., Cornell, H. V., Damschen, E.I., Jonathan Davies, T., Grytnes, J.A., Harrison, S.P., Hawkins, B.A., Holt, R.D., McCain, C.M. & Stephens, P.R. (2010) Niche conservatism as an emerging principle in ecology and conservation biology. *Ecology Letters*, **13**, 1310–1324.

Chapter 2

Stronger niche than dispersal effects on alpha- and beta-diversity of stream algae, insects, and fish across latitudes in the US

Joseph L. Mruzek[1], William R. Budnick[1,2], Chad Larson[3], Dinh K. Luc[1], and Sophia I. Passy[1]*

[1]Department of Biology, University of Texas at Arlington, 501 S. Nedderman Dr., Arlington Texas, 76019, USA

[2]Department of Fisheries and Wildlife, Michigan State University, 480 Wilson Road, East Lansing, Michigan, 48824, USA

[3]Environmental Assessment Program, Washington State Department of Ecology, Lacey, WA, 98503, USA

Introduction

Latitudinal gradients in biodiversity and community composition have been studied for decades, and we continue to improve our understanding of how niche and dispersal processes combine to shape biodiversity (Pianka, 1966; Hillebrand, 2004; Benito *et al.*, 2018). Niche and dispersal processes have been historically viewed as underlying the nearly universal decline of biodiversity toward the poles (Weiher & Keddy, 2004). Latitudinal biodiversity patterns may emerge from stronger environmental control at high vs. low latitudes, filtering out intolerant species. The physiological tolerance hypothesis (PTH) posits that richness is determined by the number of species that can physiologically tolerate the local conditions (Currie *et al.*, 2004). As higher latitudes encompass fewer species able to endure the harsh climatic conditions, these latitudes are comparatively species poor. According to the Dobzhansky-MacArthur hypothesis (DMH) (Brown, 2014), species ranges, and by extension community composition, are driven by different forces along the latitudinal gradient—environmental filtering predominates at high latitudes, while species interactions, at low latitudes.

Climatic latitudinal differences also create commensurate differences in species dispersal capacities and range sizes with direct impact on the latitudinal biodiversity patterns. According to the climate-mediated dispersal-ecological specialization tradeoff hypothesis (Jocque *et al.*, 2010), climatic harshness and seasonal variability at higher latitudes promote dispersal by seasonal migration, colonization of empty patches following die-offs, and formation of resting stages. Greater dispersal leads to broad distributional ranges and ecological generalization. In contrast, the favorable and climatically stable environments in the tropics select for greater ecological specialization which, however, limits dispersal. Since dispersal prevents isolation, which is necessary for speciation, regions are more species-poor at higher than lower latitudes

(Jocque *et al.*, 2010). Increasing niche breadth with latitude is also the essence of the so-called latitude-niche breadth hypothesis (Vázquez & Stevens, 2004). This hypothesis has been tested with abiotic and biotic niches and across ecosystems but remains controversial, given that positive, negative and no relationships between niche breadth and latitude have been reported (Vázquez & Stevens, 2004; Cirtwill *et al.*, 2015; Alahuhta *et al.*, 2017b; Granot & Belmaker, 2020). A related idea is the Rapoport's rule, suggesting that at high latitudes, species tend to have larger ranges because they have broad tolerances and better exploit temporal windows with favorable climatic conditions for dispersal (Janzen, 1967; Stevens, 1989). Conversely, at low latitudes, species are less tolerant with narrower ranges and thus more likely to encounter unfavorable conditions when dispersed. In general, Rapoport's rule has received broad support in terrestrial plants (Morueta-Holme *et al.*, 2013) and mammals (Arita *et al.*, 2005), mixed support in fish (Rohde *et al.*, 1993), and no support in aquatic diving beetles (Heino & Alahuhta, 2019), but has not been tested in algal communities (Soininen & Teittinen, 2019).

Combining ideas from the aforementioned hypotheses, we predicted that the role of niche processes in controlling biodiversity would increase with latitude, while dispersal processes would show an opposite trend. We examined the patterns of α-diversity, local species richness, and β-diversity, compositional dissimilarity among sites, as well as their respective environmental and spatial drivers. Species richness has been shown to generally decline with latitude, a trend known formally as the latitudinal diversity gradient (LDG) (Willig *et al.*, 2003; Hillebrand, 2004), although exceptions have also been reported, e.g. in freshwater diatoms (Passy, 2010; Soininen *et al.*, 2016). The latitudinal patterns of β-diversity have been less investigated (Soininen *et al.*, 2018), with studies revealing negative (Qian, 2009; Baselga, 2010; Soininen *et al.*, 2018) or weak trends (Harrison *et al.*, 1992; Alahuhta *et al.*, 2017a).

Whether the latitudinal diversity gradient has a niche or dispersal basis, or some combination of both, is an ongoing area of research. Stronger environmental filtering at high latitudes was observed in desmid communities (Bestová *et al.*, 2018), new world vascular plants (Morueta-Holme *et al.*, 2013), and Japanese forests (Kubota *et al.*, 2018). Conversely, an increased influence of dispersal-related processes at low latitudes due to range restriction resulting from climatic stability was shown in new world frogs (Gutiérrez-Pesquera *et al.*, 2016) and montane insects (Gill *et al.*, 2016). Freshwater systems contain a high level of biodiversity relative to their size, and are under greater extinction risk due to global change than other ecosystems (Dudgeon *et al.*, 2006; Wiens, 2016; Su *et al.*, 2021). Thus, elucidating the drivers of stream biodiversity across latitudes is critical for conservation.

Using major stream groups, including algae, insects, and fish, along a subcontinental latitudinal gradient, we tested the hypothesis that for α- and β-diversity, the importance of niche processes increases with latitude due to climatic stress, while the importance of dispersal is inversely related to latitude due to greater dispersal limitation. Niche processes were evaluated by climatic conditions and local physico-chemistry because air temperature and precipitation influence stream water temperature, pH, and nutrient concentrations (Jiang *et al.*, 2014) and may have an indirect effect on biodiversity. Dispersal processes were assessed by overland spatial distances. These distances have been successfully used to describe dispersal effects in algae and insects (Keck *et al.*, 2018; He *et al.*, 2020). Even in fish, where instream distances are better predictors of dispersal, overland distances also capture a significant fraction of community variation (Landeiro *et al.*, 2011).

Methods

Study area

Our initial data comprised 731 algal (diatoms and soft algae), 863 insect, and 710 fish samples (each sample representing a single stream locality) from the central United States, between the longitudes of −87° and −100° (Fig. 1). We selected this longitudinal range to avoid alpine effects to the west and maritime effects to the east, which can distort the relationship between latitude and climate. In order to assess the latitudinal drivers of biodiversity, we subdivided the study region into seven latitudinal zones. This distinction was based on 2.5°C isotherms in annual mean temperature, which was derived from the WorldClim database (Hijmans *et al.*, 2005). Samples in this dataset generally showed uneven spatial aggregation among latitudinal zones for all taxonomic datasets. To alleviate this problem, we selected 50 sites for each metacommunity in each latitudinal zone, while controlling for average distance between localities, except for the highest latitudinal zone of algae, which had only 49 samples. While the average pairwise distances between our sites were significantly different between latitudinal zones, the magnitude of difference was small (< 35 km). Overall algal samples were on average 314 km apart, insects 317 km apart, and fish 315 km apart. In all three datasets, the minimum distance between the closest sites was 1 km.

Sample collection and identification

Data for all species groups were obtained from the US Geological Survey's National Water-Quality Assessment (NAWQA) Program and the US Environmental Protection Agency's National Rivers and Streams Assessment (NRSA) using standard protocols (Moulton II *et al.,* 2002; US Environmental Protection Agency, 2013). Algae and insects were collected from the richest-targeted habitats, including hard substrates or macrophytes in faster currents. Samples were limited to those taken during the warm months (May-September) between 1993 and 2014 for algae and 1993 and 2018 for insects. Fish were sampled using seines and electrofishing from stream reaches between April and October from 1993 to 2019. Identifications were made to species for algae and fish, but genus for insects. All abundance-based analyses were conducted with density data (cells/cm^2 for algae and individuals/m^2 for insects) or raw counts for fish.

Physicochemical data consisted of water temperature (°C), nitrate (NO_3, µg/L), pH, total phosphorous (TP, µg/L), and specific conductance (µS/cm), and were obtained from the NAWQA and NRSA datasets, respectively. Elevation (m) was obtained from the WorldClim database (Fick & Hijmans, 2017) and used to calculate slope (% grade) with the package 'raster' (Hijmans *et al.*, 2020). The area of the drainage basin for each site (km^2), obtained from the NAWQA and NRSA datasets, was derived from the National Hydrology Dataset (Moore & Dewald, 2016) and included as a proxy for stream size. Climate data for each site were downloaded from the WorldClim database and included annual mean temperature, temperature seasonality, maximum temperature of the warmest month, minimum temperature of the coldest month, annual precipitation, precipitation of the wettest month, precipitation of the driest month, and precipitation seasonality (Hijmans *et al.*, 2005). Climate, elevation, and slope values for each site represent the average of a 5 km buffer around the sites. Latitude and longitude data were available for all locations. All environmental variables (but pH) were ln-transformed if normality was improved, and slope was arcsine square root-transformed. All variables were standardized (mean = 0, standard deviation = 1). Pairwise distances between all sites were calculated in km using the 'distm' function in the R package 'geosphere' (Hijmans *et al.*, 2019).

Species and diversity metrics

For a more accurate assessment of taxa niche breadths and range sizes, we used a nationwide dataset, consisting of 2687 algal, 3809 insect, and 2753 fish samples (Fig. 1). Niche breadths were calculated as distributions along environmental gradients with outlying mean index (OMI) analysis (Dolédec *et al.*, 2000). We carried out PCA of the environmental variables and retained only axes with eigenvalues ≥ 1. We Hellinger-transformed taxa abundances (Legendre & De Cáceres, 2013) and calculated niche breadth only for taxa with at least five occurrences. Calculation was done with the 'niche' function in the R package 'adespatial' (Dray *et al.*, 2021). Taxa range sizes were calculated with convex hull areas (km^2) surrounding the sites where each taxon occurred. Site coordinates were in decimal degrees (Projection: EPSG: 4326) and the convex hull was measured with the 'area' function in the R package 'raster' (Hijmans *et al.*, 2020). Species that occurred in fewer than three sites were not included in range calculations, as at least three sites are needed to create a convex hull. Taxa were said to occur in a latitudinal zone if they were found in at least one site within that zone.

To assess if taxa tended to inhabit more sites at high vs. low latitudes, we measured occupancy as the proportion of sites within a latitudinal zone where a taxon occurred. We calculated for each site species richness (α-diversity) and for each latitudinal zone: mean species richness (mean α-diversity), compositional variability as pairwise Bray-Curtis dissimilarity (β-diversity), total species richness (γ-diversity).

Data analyses

Diversity metrics, occupancy, mean range size and mean niche breadth of the taxa found in each of the latitudinal zones (49-50 sites) were regressed against the median latitude of the sites within these zones.

To obtain spatial predictors, we performed a distance-based Moran eigenvector maps (dbMEM) using the 'listw.candidates' function in the R package 'adespatial' (Dray *et al.*, 2021) and retained the positively correlated MEMs for further analyses. We measured the relative contribution of niche and dispersal processes on α-diversity with multiple regression and variance partitioning. We regressed sample α-diversity within each latitudinal zone using forward selection against i) environmental variables and their squares and ii) positively correlated MEMs, using 'forward.sel' in 'adespatial' (Dray *et al.*, 2021). Then, we performed variance partitioning with 'varpart' in 'vegan' (Oksanen *et al.*, 2019). We measured the relative contribution of niche and dispersal processes on β-diversity with distance-based RDA (dbRDA) with variance partitioning. We employed stepwise forward selection of the Bray-Curtis dissimilarity matrix against i) environmental predictors and their squares and ii) positively correlated MEMs with the arguments 'ordistep' and 'capscale' in the package 'vegan' (Oksanen *et al.*, 2019). We applied Lingoes correction to our dbRDA to prevent calculation of negative eigenvalues owing to the non-Euclidean properties of Bray-Curtis dissimilarity (Legendre & Anderson, 1999). Adjusted R^2 values (from the pure environmental and pure spatial fractions) for α- and β-diversity in each latitudinal zone were regressed against the median latitude of the sites within this zone. The adjusted R^2 values for the environmental and spatial components were compared with a paired t-test, as both components were derived from the same sites with each latitudinal zone. To test the effects of water physicochemistry, climate, and their covariance, we performed the above analyses of α- and β-diversity with environmental data split into water physicochemistry and climate and positively correlated MEMs.

Results

Latitudinal trends

There were 466 algal species, 171 insect genera, and 266 fish species in our datasets. Only in fish, mean α-diversity, and regional richness (γ-diversity) declined marginally non-significantly or significantly at higher latitudes, respectively ($p \leq 0.05$), consistent with the latitudinal diversity gradient (Table 2.1). β-diversity was high across groups (0.7-0.8) and did not vary with latitude ($p > 0.05$, Table 2.1), indicating high dissimilarity between communities, regardless of latitude. Only in fish niche breadth and range size exhibited a marginally non-significant and significant positive linear relationships with latitude, respectively ($p \leq 0.05$, Table 2.1). These results indicated a tendency in fish toward wider niche breadths at higher latitudes, consistent with the latitude-niche breadth hypothesis, and broader geographic distributions, as predicted by Rapoport's rule. Occupancy was generally low, ranging from 10% to 11% for algae, 11% to 15% for insects and 13% to 18% for fish, and did not display significant latitudinal trends (Table 2.1). Considering α-diversity of all sites, irrespective of zone, we found significant quadratic relationships with latitude in all three datasets. Algae had the highest α-diversity at low and high latitudes, insects, at high latitudes, while fish, at low to mid latitudes (Fig. 2.2). However, the relationship in insects was very weak.

Tests of hypotheses

Opposing latitudinal trends in niche (positive) and dispersal (negative), as predicted by our hypothesis, were found only for α-diversity of algae and fish (Fig. 2.3a-c). In the case of insects, niche processes had a significant positive relationship with latitude, while dispersal had no relationship, thus only partially supporting our hypothesis. The covariance of niche and dispersal processes did not respond to latitude. In all three species groups, niche processes and the covariance between niche and dispersal processes explained overall more variance in species richness than dispersal processes (ANOVA and Tukey’s post-hoc test) (Fig. 2.3a-c).

With respect to metacommunity β-diversity, niche processes in insects had a significantly negative latitudinal trend, and dispersal processes in algae had a significantly positive trend (Fig. 2.4a-c), contrary to what we predicted. The covariance of niche and dispersal processes increased with latitude but only in algae, while in the other two species groups, it showed no latitudinal dependence. In all three species groups, niche processes explained more variance in β-diversity than dispersal processes but had a comparable effect to this of the covariance of niche and dispersal processes (Fig. 2.4a-c).

To assess the origins of the niche effects, we compared the explained variance of α- and β-diversity by climate, physicochemistry, and their covariance together with space. Pure physicochemistry tended to explain best both α-diversity and β-diversity, followed by the covariance of physicochemistry and climate and pure climate, which were generally comparable (Fig. 2.5).

Discussion

In this comprehensive investigation of stream algae, insects, and fish, we tested whether niche effects increased with latitude, while dispersal decreased with latitude. Our results supported our hypothesis only for α-diversity but not β-diversity. Regardless of latitude, niche processes emerged as stronger predictors of biodiversity.

Latitudinal trends

Species richness represented here as raw and average α-diversity, and γ-diversity, had a negative relationship with latitude only in fish, consistent with the classical latitudinal pattern. In algae and insects, the latitudinal responses of these metrics were nonlinear, weak, or non-significant. Our results thus align with previous work, showing the classical latitudinal diversity gradient in North American fish (Griffiths, 2015; Hanly *et al.*, 2017; Miller & Román-Palacios, 2021) but not in diatoms (Passy, 2010; Soininen *et al.*, 2016) or insects (Vinson & Hawkins, 2003). However, as we examine here both soft algae and diatoms, our results on the latitudinal diversity patterns are more general.

Notably, in all three species groups, we found little support for latitudinal gradients in β-diversity in agreement with other freshwater studies reporting weak trends in macrophytes (Alahuhta *et al.*, 2017a) and ephemeral pond macroinvertebrates (Kneitel, 2016). These findings contrast a global meta-analysis, showing generally lower β-diversity at high latitudes across a wide range of species, including bacteria, birds, and marine and freshwater fish (Soininen *et al.*, 2018). Declines in β-diversity at high latitudes have been related to increased range sizes, consistent with Rapoport’s rule (Qian *et al.*, 2009). Here we confirmed Rapoport’s rule only in stream fish, in agreement with past research (Stevens, 1989). Additionally, we demonstrated that this rule applies not only to freshwater fish above 40 degrees latitude, as previously documented (Rohde *et al.*, 1993), but also to stream fish between 30 and 40 degrees latitude. Prior to this study, Rapoport's rule had not been tested in algae (Soininen & Teittinen, 2019), and our lack of support for this rule suggests that algae may not have the same latitudinal gradients in dispersal limitation as larger organisms. In insects, we did not support Rapoport's rule, similar to results for aquatic diving beetles (Heino & Alahuhta 2019), but with a larger and more diverse dataset. However, regardless of whether Rapoport's rule was observed, no species group displayed a latitudinal β-diversity gradient. We attribute this finding in part to the low mean occupancy across latitudinal zones (average 10-15% of samples), given that low occupancy drives high β-diversity (Xu *et al.*, 2015). Additionally, the fragmented nature of stream systems tends to increase β-diversity even in the presence of Rapoport's rule (Soininen *et al.*, 2007).

We also tested the latitude-niche breadth hypothesis and found no support in algae and insects. Our results thus add to the findings of a recent global meta-analysis, which did not detect an overall correlation between latitude and environmental niche breadth in other species groups (Granot & Belmaker, 2020). However, the tendency of fish niche breadths to decrease with latitude (the trend was only marginally non-significant) opens the question of whether ecological specialization in streams may be more likely in less dispersive species.

Importance of niche vs. dispersal processes as drivers of biodiversity

Local species richness of all three species groups was more strongly controlled by niche processes at higher latitudes, in agreement with our hypothesis. These results are consistent with research on α-diversity of desmid communities, showing that the effect of habitat filtering increased at high latitudes (Bestová *et al.*, 2018). Congruent with our findings for fish, increased habitat filtering with latitude was also observed in French fish (Blanchet *et al.*, 2014). Strong environmental filtering effects at high latitudes were also reported for algae, macroinvertebrates, and fish from the boreal zone (Heino *et al.*, 2016; de Mendoza *et al.*, 2018; Brittain *et al.*, 2020). Our comparative analysis enhances this knowledge by demonstrating that the niche processes controlling local richness of three major stream groups are latitudinally constrained at subtropical to temperate climates. Unlike α-diversity, the niche effects on β-diversity had a significant negative relationship with latitude in insects but no relationship in algae and fish, contrary to our first hypothesis. The spatially structured environment, represented by the covariance term, had a significant latitudinal dependence only in algal β-diversity, indicating that for the most part, only the pure effects varied latitudinally.

We predicted that the role of niche processes would increase with latitude, assuming increased climatic stress to be the dominant factor driving community assembly. We did observe latitudinal increase in niche effects on α-diversity, but they were largely contributed by local physicochemistry. Previous research observed local physicochemical and climatic dependence of algal and fish community composition (Qu *et al.*, 2019; Gálvez *et al.*, 2020) and reported that climate alone cannot fully characterize algal biodiversity patterns across latitudes (Passy *et al.*, 2018). This suggests that community assembly along a latitudinal gradient has more complex environmental underpinnings. We predicted that if taxa niche breadths and range sizes increased with latitude, the dispersal effect on both biodiversity metrics would decrease. We found some evidence for the latitude-niche breadth hypothesis and confirmed Rapoport's rule only in fish, yet we supported our hypothesis for α-diversity not only in fish, but also in algae. These results imply that the dispersal-ecological specialization tradeoff hypothesis may operate to some degree in fish but not in algae or insects. The algal results further indicated that historic factors other than species range size contributed to greater dispersal effects at lower latitudes. Conversely, the effect of space on β-diversity did not follow our predicted latitudinal trend for all three species groups, which may be due to the low occupancy across latitudes.

We observed that niche processes and the spatially structured environment had greater contributions to α- and β-diversity than dispersal in all cases. While research on stream macroinvertebrates over smaller areas, allowing for greater connectivity between sites, found the opposite pattern (Jiang *et al.*, 2021), our results are consistent with studies on stream groups, indicating that habitat filtering is often the primary driver of diversity at large scales (Cottenie, 2005; Heino *et al.*, 2016). Climate and physicochemistry have pronounced impacts on the distribution of diatom, insect, and fish taxa (Magnuson *et al.*, 1979; Hasnain *et al.*, 2013; Healy *et al.*, 2018; Pound *et al.*, 2021) and global changes in these conditions may alter biodiversity (Biswas *et al.*, 2017; Fraker *et al.*, 2020; Pound *et al.*, 2021). While dispersal effects were weaker than niche effects in our study, they were still notable. Therefore, we recommend that future projections of aquatic biodiversity under global change specifically incorporate dispersal effects.

In conclusion, α- and β-diversity were controlled by both niche and dispersal processes, but niche processes were generally stronger. The weaker dispersal effects may be due in part to the way dispersal was measured (overland as opposed to instream) or to the scale of our latitudinal zones, where dispersal may have not been too limiting. Clear latitudinal trends in niche and dispersal effects were observed for α-diversity and rarely for β-diversity. However, the niche effects were mostly due to physicochemistry, and the dispersal effects were not always associated with variability in species niche breadths and range sizes. Therefore, the Dobzhansky-MacArthur hypothesis, the physiological tolerance hypothesis, and the dispersal-ecological specialization tradeoff hypothesis, predicting stronger environmental filtering at higher latitudes but more pronounced dispersal effects at lower latitudes as a result of climatic variability, do not have support in our streams. The reasons for the deviation of our results from these hypotheses may lie in physicochemistry potentially offsetting the effect of climate and the overall lack of latitudinal variability in specialization and range size.

References

Alahuhta, J., Kosten, S., Akasaka, M., Auderset, D., Azzella, M.M., Bolpagni, R., Bove, C.P., Chambers, P.A., Chappuis, E., Clayton, J., de Winton, M., Ecke, F., Gacia, E., Gecheva, G., Grillas, P., Hauxwell, J., Hellsten, S., Hjort, J., Hoyer, M. V., Ilg, C., Kolada, A., Kuoppala, M., Lauridsen, T., Li, E.H., Lukács, B.A., Mjelde, M., Mikulyuk, A., Mormul, R.P., Nishihiro, J., Oertli, B., Rhazi, L., Rhazi, M., Sass, L., Schranz, C., Søndergaard, M., Yamanouchi, T., Yu, Q., Wang, H., Willby, N., Zhang, X.K. & Heino, J. (2017a) Global variation in the beta diversity of lake macrophytes is driven by environmental heterogeneity rather than latitude. *Journal of Biogeography*, **44**, 1758–1769.

Alahuhta, J., Virtala, A., Hjort, J., Ecke, F., Johnson, L.B., Sass, L. & Heino, J. (2017b) Average niche breadths of species in lake macrophyte communities respond to ecological gradients variably in four regions on two continents. *Oecologia*, **184**, 219–235.

Arita, H.T., Rodríguez, P. & Vázquez-Domínguez, E. (2005) Continental and regional ranges of North American mammals: Rapoport's rule in real and null worlds. *Journal of Biogeography*, **32**, 961–971.

Baselga, A. (2010) Partitioning the turnover and nestedness components of beta diversity. *Global Ecology and Biogeography*, **19**, 134–143.

Benito, X., Fritz, S., Steinitz-Kannan, M., Tapia, P.M., Kelly, M.A. & Lowell, T.V. (2018) Geo-climatic factors drive diatom community distribution in tropical South American freshwaters. *Journal of Ecology*, **106**, 1660–1672.

Bestová, H., Munoz, F., Svoboda, P., Škaloud, P. & Violle, C. (2018) Ecological and biogeographical drivers of freshwater green algae biodiversity: from local communities to large-scale species pools of desmids. *Oecologia*, **186**, 1017–1030.

Biswas, S.R., Vogt, R.J. & Sharma, S. (2017) Projected compositional shifts and loss of ecosystem services in freshwater fish communities under climate change scenarios. *Hydrobiologia*, **799**, 135–149.

Blanchet, S., Helmus, M.R., Brosse, S. & Grenouillet, G. (2014) Regional vs local drivers of phylogenetic and species diversity in stream fish communities. *Freshwater Biology*, **59**, 450–462.

Brittain, J.E., Heino, J., Friberg, N., Aroviita, J., Kahlert, M., Karjalainen, S.M., Keck, F., Lento, J., Liljaniemi, P., Mykrä, H., Schneider, S.C. & Ylikörkkö, J. (2020) Ecological correlates of riverine diatom and macroinvertebrate alpha and beta diversity across Arctic Fennoscandia. *Freshwater Biology*, 1–15.

Brown, J.H. (2014) Why are there so many species in the tropics? *Journal of Biogeography*, **41**, 8–22.

Cirtwill, A.R., Stouffer, D.B. & Romanuk, T.N. (2015) Latitudinal gradients in biotic niche breadth vary across ecosystem types. *Proceedings of the Royal Society B: Biological Sciences*, **282**.

Cottenie, K. (2005) Integrating environmental and spatial processes in ecological community dynamics. *Ecology Letters*, **8**, 1175–1182.

Currie, D.J., Mittelbach, G.G., Cornell, H. V., Field, R., Guégan, J.F., Hawkins, B.A., Kaufman, D.M., Kerr, J.T., Oberdorff, T., O'Brien, E. & Turner, J.R.G. (2004) Predictions and tests of climate-based hypotheses of broad-scale variation in taxonomic richness. *Ecology Letters*, **7**, 1121–1134.

Dolédec, S., Chessel, D. & Gimaret-Carpentier, C. (2000) Niche separation in community analysis: A new method. *Ecology*, **81**, 2914–2927.

Dray, S., Bauman, D., Guillaume Blanchet, F., Borcard, D., Clappe, S., Guenard, G., Jombart, T., Larocque, G., Legendre, P., Madi, N. & Wagner, H.H. (2021) Package 'adespatial.' *CRAN*.

Dudgeon, D., Arthington, A.H., Gessner, M.O., Kawabata, Z.I., Knowler, D.J., Lévêque, C., Naiman, R.J., Prieur-Richard, A.H., Soto, D., Stiassny, M.L.J. & Sullivan, C.A. (2006) Freshwater biodiversity: Importance, threats, status and conservation challenges. *Biological Reviews of the Cambridge Philosophical Society*, **81**, 163–182.

Fick, S.E. & Hijmans, R.J. (2017) 2017 - WorldClim 2 new 1-km spatial resolution climate surfaces for global land areas-annotated.pdf. *International Journal of Climatology*, **37**, 4302–4315.

Fraker, M.E., Keitzer, S.C., Sinclair, J.S., Aloysius, N.R., Dippold, D.A., Yen, H., Arnold, J.G., Daggupati, P., Johnson, M.V. V., Martin, J.F., Robertson, D.M., Sowa, S.P., White, M.J. & Ludsin, S.A. (2020) Projecting the effects of agricultural conservation practices on stream fish communities in a changing climate. *Science of the Total Environment*, **747**, 1–14.

Gálvez, Á., Aguilar-Alberola, J.A., Armengol, X., Bonilla, F., Iepure, S., Monrós, J.S., Olmo, C., Rojo, C., Rueda, J., Rueda, R., Sasa, M. & Mesquita-Joanes, F. (2020) Environment and space rule, but time also matters for the organization of tropical pond metacommunities. *Frontiers in Ecology and Evolution*, **8**, 1–13.

Gill, B.A., Kondratieff, B.C., Casner, K.L., Encalada, A.C., Flecker, A.S., Gannon, D.G., Ghalambor, C.K., Guayasamin, J.M., Poff, N.L., Simmons, M.P., Thomas, S.A., Zamudio, K.R. & Funk, W.C. (2016) Cryptic species diversity reveals biogeographic support for the 'mountain passes are higher in the tropics' hypothesis. *Proceedings of the Royal Society B: Biological Sciences*, **283**, 7–12.

Granot, I. & Belmaker, J. (2020) Niche breadth and species richness: Correlation strength, scale and mechanisms. *Global Ecology and Biogeography*, **29**, 159–170.

Griffiths, D. (2015) Connectivity and vagility determine spatial richness gradients and diversification of freshwater fish in North America and Europe. *Biological Journal of the Linnean Society*, **116**, 773–786.

Gutiérrez-Pesquera, L.M., Tejedo, M., Olalla-Tárraga, M.A., Duarte, H., Nicieza, A. & Solé, M. (2016) Testing the climate variability hypothesis in thermal tolerance limits of tropical and temperate tadpoles. *Journal of Biogeography*, **43**, 1166–1178.

Hanly, P.J., Mittelbach, G.G. & Schemske, D.W. (2017) Speciation and the latitudinal diversity gradient: Insights from the global distribution of endemic fish. *The American Naturalist*, **189**, 604–615.

Harrison, S.P., Ross, S.J. & Lawton, J.H. (1992) Beta diversity on geographic gradients in Britain. *Journal of Animal Ecology*, **61**, 151–158.

Hasnain, S.S., Shuter, B.J. & Minns, C.K. (2013) Phylogeny influences the relationships linking key ecological thermal metrics for North American freshwater fish species. *Canadian Journal of Fisheries and Aquatic Sciences*, **70**, 964–972.

He, S., Soininen, J., Chen, K. & Wang, B. (2020) Environmental factors override dispersal-related factors in shaping diatom and macroinvertebrate communities within stream networks in China. *Frontiers in Ecology and the Environment*, **8**, 141.

Healy, T.M., Brennan, R.S., Whitehead, A. & Schulte, P.M. (2018) Tolerance traits related to climate change resilience are independent and polygenic. *Global Change Biology*, **24**, 5348–5360.

Heino, J. & Alahuhta, J. (2019) Knitting patterns of biodiversity, range size and body size in aquatic beetle faunas: significant relationships but slightly divergent drivers. *Ecological Entomology*, **44**, 413–424.

Heino, J., Soininen, J., Alahuhta, J., Lappalainen, J. & Virtanen, R. (2016) Metacommunity ecology meets biogeography: effects of geographical region, spatial dynamics and environmental filtering on community structure in aquatic organisms. *Oecologia*, **183**, 121–137.

Hijmans, R.J., Cameron, S.E., Parra, J.L., Jones, P.G. & Jarvis, A. (2005) Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, **25**, 1965–1978.

Hijmans, R.J., Sumner, M., Cheng, J., Baston, D., Bevan, A., Bivand, R., Buseto, L., Canty, M., Forrest, D., Golicher, D., Gray, J., Greenberg, J.A., Hiemstra, P., Hingee, K., Karney, C., Mattiuzzi, M., Mosher, S., Nowosad, J., Pebesma, E., Lamigueiro, O.P., Racine, E.B., Rowlingson, B., Shortridge, A., Venables, B. & Wueest, R. (2020) Package "Raster."

Hijmans, R.J., Williams, E. & Vennes, C. (2019) Package 'geosphere.'

Hillebrand, H. (2004) On the generality of the latitudinal diversity gradient. *The American Naturalist*, **163**, 192–211.

Janzen, D.H. (1967) Why mountain passes are higher in the tropics. *The American Naturalist*, **101**, 233–249.

Jiang, J., Sharma, A., Sivakumar, B. & Wang, P. (2014) A global assessment of climate-water quality relationships in large rivers: An elasticity perspective. *Science of the Total Environment*, **468–469**, 877–891.

Jiang, X., Pan, B., Jiang, W., Hou, Y., Yang, H., Zhu, P. & Heino, J. (2021) The role of environmental conditions, climatic factors and spatial processes in driving multiple facets of stream macroinvertebrate beta diversity in a climatically heterogeneous mountain region.

Ecological Indicators, **124**, 107407.

Jocque, M., Field, R., Brendonck, L. & De Meester, L. (2010) Climatic control of dispersal-ecological specialization trade-offs: A metacommunity process at the heart of the latitudinal diversity gradient? *Global Ecology and Biogeography*, **19**, 244–252.

Keck, F., Franc, A. & Kahlert, M. (2018) Disentangling the processes driving the biogeography of freshwater diatoms: A multiscale approach. *Journal of Biogeography*, **45**, 1582–1592.

Kneitel, J.M. (2016) Climate-driven habitat size determines the latitudinal diversity gradient in temporary ponds. *Ecology*, **97**, 961–968.

Kubota, Y., Kusumoto, B., Shiono, T. & Ulrich, W. (2018) Multiple filters affect tree species assembly in mid-latitude forest communities. *Oecologia*, **187**, 245–253.

Landeiro, V.L., Magnusson, W.E., Melo, A.S., Espírito-Santo, H.M.V. & Bini, L.M. (2011) Spatial eigenfunction analyses in stream networks: Do watercourse and overland distances produce different results? *Freshwater Biology*, **56**, 1184–1192.

Legendre, P. & Anderson, M.J. (1999) Distance-based redundancy analysis: testing multispecies responses in multifactorial ecological experiments. *Ecological Monographs*, **69**, 1–24.

Legendre, P. & De Cáceres, M. (2013) Beta diversity as the variance of community data: Dissimilarity coefficients and partitioning. *Ecology Letters*, **16**, 951–963.

Magnuson, J.J., Crowder, L.B. & Medvick, P.A. (1979) Temperature as an ecological resource. *American Zoologist*, **19**, 331–343.

de Mendoza, G., Kaivosoja, R., Grönroos, M., Hjort, J., Ilmonen, J., Kärnä, O.M., Paasivirta, L., Tokola, L. & Heino, J. (2018) Highly variable species distribution models in a subarctic stream metacommunity: Patterns, mechanisms and implications. *Freshwater Biology*, **63**, 33–47.

Miller, E.C. & Román-Palacios, C. (2021) Evolutionary time explains the global distribution of freshwater fish diversity. *Global Ecology and Biogeography*, **30**, 749–763.

Moore, R.B. & Dewald, T.G. (2016) The road to NHDPlus — advancements in digital stream networks and associated catchments. *Journal of the American Water Resources Association*, **52**, 890–900.

Morueta-Holme, N., Enquist, B.J., Mcgill, B.J., Boyle, B., Jørgensen, P.M., Ott, J.E., Peet, R.K., Šímová, I., Sloat, L.L., Thiers, B., Violle, C., Wiser, S.K., Dolins, S., Donoghue, J.C., Kraft, N.J.B., Regetz, J., Schildhauer, M., Spencer, N. & Svenning, J.C. (2013) Habitat area and climate stability determine geographical variation in plant species range sizes. *Ecology Letters*, **16**, 1446–1454.

Moulton II, S.R., Kennen, J.G., Goldstein, R.M. & Hambrook, J.A. (2002) Revised protocols for sampling algal, invertebrate, and fish communities as part of the national water-quality assessment program. *USGS.*

Oksanen, J., Blanchet, F.G., Friendly, M., Kindt, R., Legendre, P., Mcglinn, D., Minchin, P.R., O’hara, R.B., Simpson, G.L., Solymos, P., Stevens, H.M., Szoecs, E. & Wagner, H. (2019)

Package “vegan”, Community Ecology Package. *CRAN*.

Passy, S.I. (2010) A distinct latitudinal gradient of diatom diversity is linked to resource supply. *Ecology*, **91**, 36–41.

Passy, S.I., Larson, C.A., Jamoneau, A., Budnick, W.R., Leboucher, T., Tison-Rosebery, J. & Soininen, J. (2018) The biogeographical patterns of species richness and abundance distribution in stream diatoms are driven by climate and water chemistry. *The American Naturalist*, **192**, 607–617.

Pianka, E.R. (1966) Latitudinal gradients in species diversity: A review of concepts. *The American Naturalist*, **100**, 33–46.

Pound, K.L., Larson, C.A. & Passy, S.I. (2021) Current distributions and future climate-driven changes in diatoms, insects and fish in U.S. streams. *Global Ecology and Biogeography*, **30**, 63–78.

Qian, H. (2009) Beta diversity in relation to dispersal ability for vascular plants in North America. *Global Ecology and Biogeography*, **18**, 327–332.

Qian, H., Badgley, C. & Fox, D.L. (2009) The latitudinal gradient of beta diversity in relation to climate and topography for mammals in North America. *Global Ecology and Biogeography*, **18**, 111–122.

Qu, X., Peng, W., Liu, Y., Zhang, M., Ren, Z., Wu, N. & Liu, X. (2019) Networks and ordination analyses reveal the stream community structures of fish, macroinvertebrate and benthic algae, and their responses to nutrient enrichment. *Ecological Indicators*, **101**, 501–511.

Rohde, K., Heap, M. & Heap, D. (1993) Rapoport’s rule does not apply to marine teleosts and cannot explain latitudinal gradients in species richness. *The American Naturalist*, **142**, 1–16.

Soininen, J., Heino, J. & Wang, J. (2018) A meta-analysis of nestedness and turnover components of beta diversity across organisms and ecosystems. *Global Ecology and Biogeography*, **27**, 96–109.

Soininen, J., Jamoneau, A., Rosebery, J. & Passy, S.I. (2016) Global patterns and drivers of species and trait composition in diatoms. *Global Ecology and Biogeography*, **25**, 940–950.

Soininen, J., McDonald, R. & Hillebrand, H. (2007) The distance decay of similarity in ecological communities. *Ecography*, **30**, 3–12.

Soininen, J. & Teittinen, A. (2019) Fifteen important questions in the spatial ecology of diatoms. *Freshwater Biology*, **64**, 2071–2083.

Stevens, G.C. (1989) The latitudinal gradient in geographical range: How so many species coexist in the tropics. *The American Naturalist*, **133**, 240–256.

Su, G., Logez, M., Xu, J., Tao, S., Villéger, S. & Brosse, S. (2021) Human impacts on global freshwater fish biodiversity. *Science*, **371**, 835–838.

US Environmental Protection Agency (2013) National Rivers and Streams Assessment 2013-2014: Field Operations Manual – Wadeable.

Vázquez, D.P. & Stevens, R.D. (2004) The latitudinal gradient in niche breadth: concepts and evidence. *The American Naturalist*, **164**.

Vinson, M.R. & Hawkins, C.P. (2003) Broad-scale geographical patterns in local stream insect genera richness. *Ecography*, **26**, 751–767.

Weiher, E. & Keddy, P.A. eds. (2004) *Ecological Assembly Rules Perspectives, advances, retreats*, Cambridge University Press, Cambridge, UK.

Wiens, J.J. (2016) Climate-related local extinctions are already widespread among plant and animal species. *PLoS Biology*, **14**, 1–18.

Willig, M.R., Kaufman, D.M. & Stevens, R.D. (2003) Latitudinal gradients of biodiversity: Pattern, process, scale, and synthesis. *Annual Review of Ecology, Evolution, and Systematics*, **34**, 273–309.

Xu, W., Chen, G., Liu, C. & Ma, K. (2015) Latitudinal differences in species abundance distributions, rather than spatial aggregation, explain beta-diversity along latitudinal gradients. *Global Ecology and Biogeography*, **24**, 1170–1180.

Table 2.1. Metacommunity statistics with corresponding slopes and p-values from regressions against the median latitude of each zone. Variability in median latitude is caused by latitudinal differences among the three species groups. † Indicates marginal non-significance.

Group	Dataset	Parameter	Median latitude							Slope of regression	p-value
			31°N	34°N	37°N	40°N	42-43°N	44°N	47°N		
Algae	Full metacommunity	Mean α	56.18	56.26	43.40	50.96	46.84	55.52	63.61	0.29	0.610
		Mean β	0.81	0.80	0.84	0.77	0.74	0.81	0.78	>0.01	0.367
		γ	585	569	457	487	466	560	584	-0.72	0.881
		Mean occupancy	9.60	9.89	9.50	10.46	10.05	9.91	10.89	0.06	0.069
		Mean range size	7,077,101	7,125,559	7,519,767	7,081,045	7,474,615	7,018,691	7,243,207	0.00	0.802
		Niche breadth	3.0	3.0	2.7	2.6	2.7	2.7	3.0	-0.01	0.429
Insects	Full metacommunity	Mean α	23.74	24.50	29.42	23.28	22.38	26.04	32.22	0.30	0.310
		Mean β	0.73	0.75	0.71	0.70	0.74	0.76	0.73	0.00	0.616
		γ	201	207	198	163	171	230	236	1.62	0.476
		Mean occupancy	11.8	11.8	14.9	14.3	13.1	11.3	13.7	0.04	0.700
		Mean range size	7,123,893	7,365,505	7,294,908	7,528,854	7,577,587	7,348,461	7,164,454	0.00	0.658
		Niche breadth	2.3	2.3	2.1	2.2	2.3	2.3	2.3	0.00	0.746
Fish	Full metacommunity	Mean α	18.86	20.70	20.86	16.26	18.34	16.90	12.90	-0.37	0.052†
		Mean β	0.76	0.74	0.79	0.75	0.76	0.79	0.75	>0.01	0.789
		γ	132	147	161	101	109	95	83	-4.09	**0.033**
		Mean occupancy	14.29	14.08	12.96	16.10	16.83	17.79	15.54	0.21	0.085
		Mean range size	2,882,262	3,417,561	2,983,526	3,680,327	3,738,899	3,660,063	3,785,298	0.02	**0.020**
		Niche breadth	1.5	1.6	1.4	1.7	1.7	1.7	1.8	0.02	0.051†

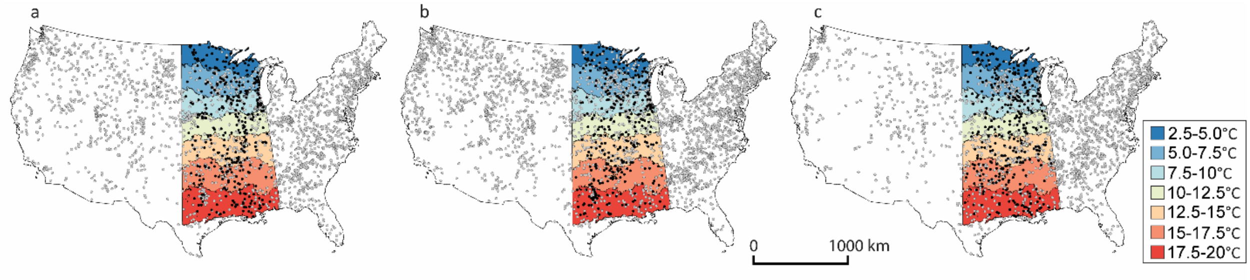

Figure 2.1. Map of sample locations for algae (a), insects (b), and fish (c) in the seven studied latitudinal zones, based on 2.5°C isotherms of mean annual temperature. Sites used to assess, range size and niche breadth are grey, while sites used to assess the relative importance of dispersal and niche processes are black (projection is EPSG: 2163).

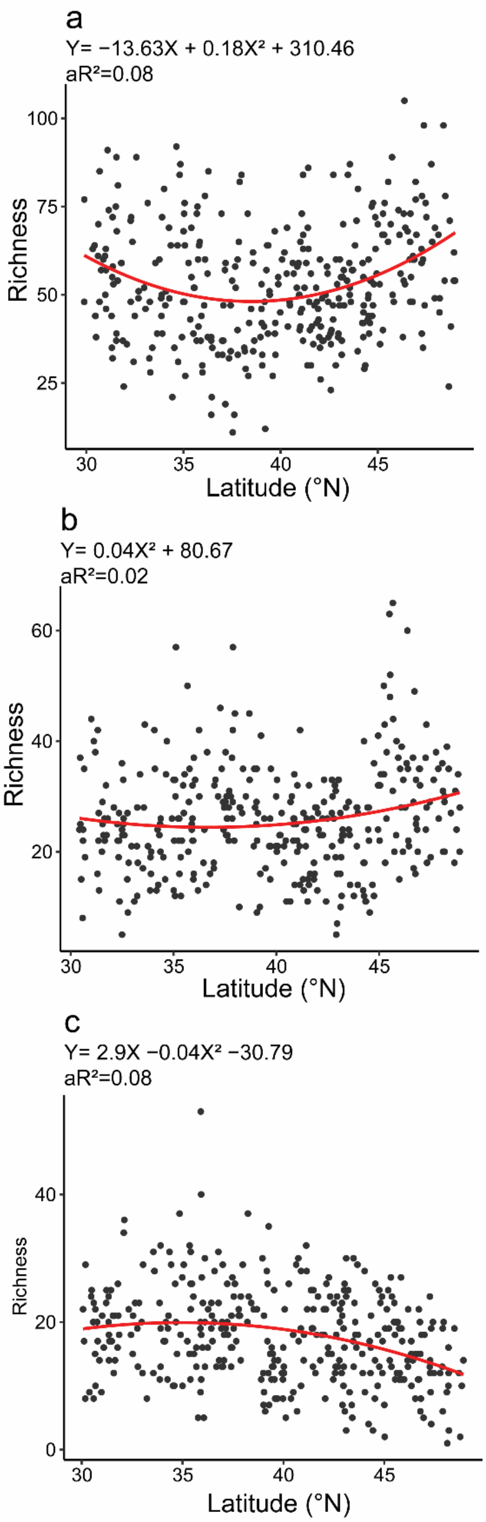

Figure 2.2. Regressions of α-diversity (Y variable) against latitude (X variable) of algae (a), insects (b) and fish (c). All regression parameters significant at $p < 0.05$. aR^2 = adjusted R^2.

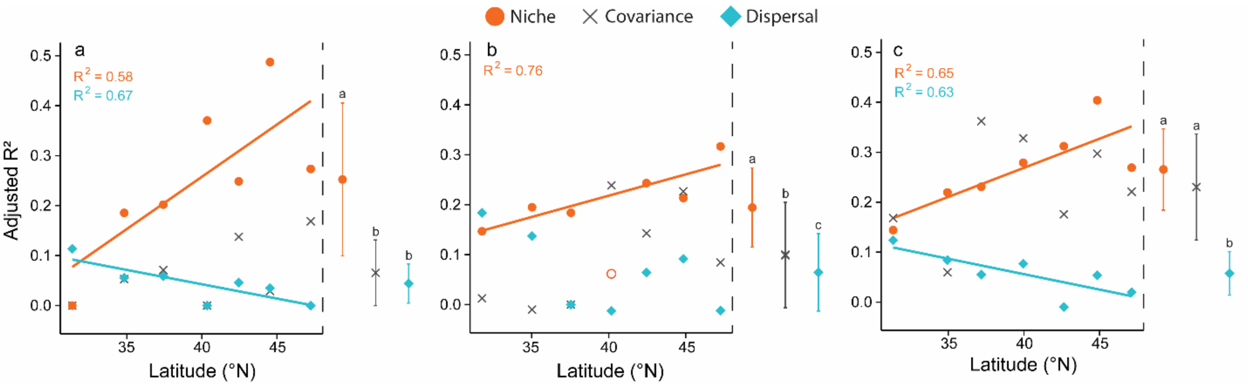

Figure 2.3. Regressions of adjusted R^2 from variance partitioning analyses of α-diversity of algae (a), insects (b), and fish (c) against the median latitude of each latitudinal zone. Significant regressions are indicated with a solid line. One point was identified as an outlier (shown as an open symbol) and not included in the regression analysis. Mean adjusted R^2 values of niche and dispersal processes and their covariance across latitudes are shown to the right of the vertical dashed line in each panel. Significant differences determined by a Tukey's post-hoc test are shown with different letters.

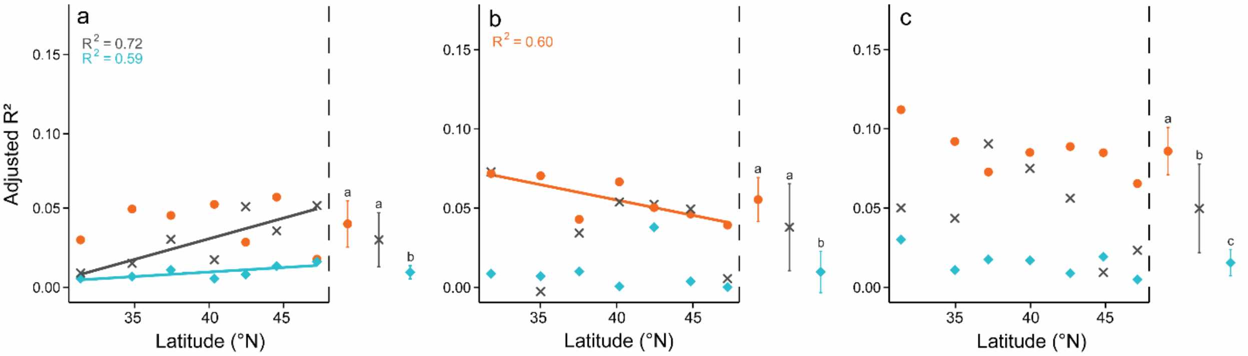

Figure 2.4. Regressions of adjusted R^2 from variance partitioning analyses of β-diversity of algae (a), insects (b), and fish (c) against the median latitude of each latitudinal zone. Significant regressions are indicated with a solid line. Mean adjusted R^2 values of niche and dispersal processes and their covariance across latitudes are shown to the right of the vertical dashed line in each panel. Significant differences determined by a Tukey's post-hoc test are shown with different letters.

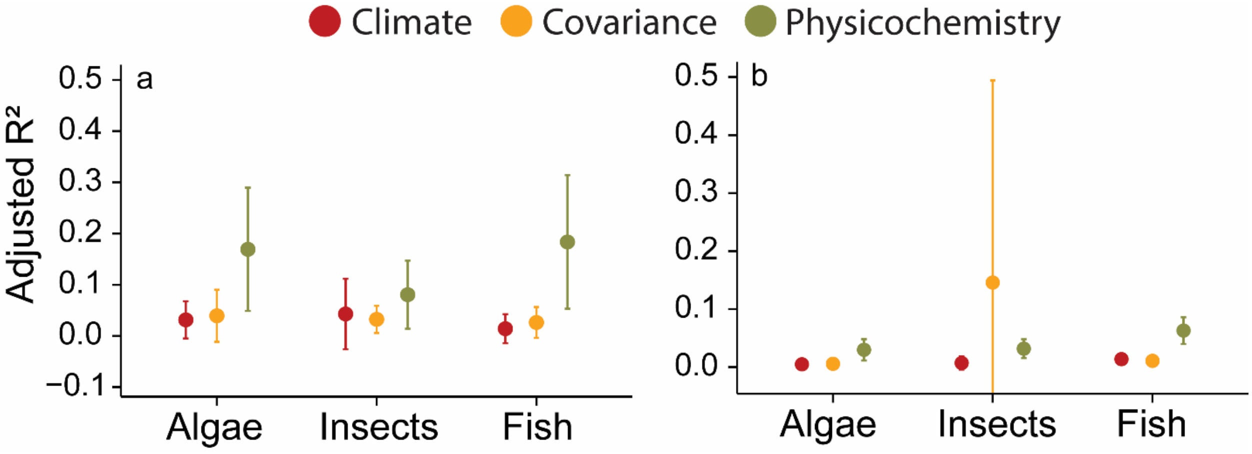

Figure 2.5: Effects of climate, physicochemistry, and their covariance following variance partitioning of climate, physicochemistry, and space (space is not shown). Adjusted R^2 for α-diversity of the metacommunity (a), and β-diversity of the metacommunity (b).

Chapter 3

Niche and dispersal processes drive metacommunity composition and co-occurrence network topology in stream diatoms, insects, and fish

Joseph L. Mruzek[1], William R. Budnick[2], Chad A. Larson[3], and S.I. Passy[1*]

[1]Department of Biology, University of Texas at Arlington, 501 S. Nedderman Dr., Arlington Texas, 76019, USA

[2]Department of Fisheries and Wildlife, Michigan State University, 480 Wilson Road, East Lansing, Michigan, 48824, USA

[3]Washington State Department of Ecology, Environmental Assessment Program, 300 Desmond Dr SE, Lacey, Washington, 98503, USA

Authors' emails and ORCID IDs:
Joseph L. Mruzek: joseph.mruzek@mavs.uta.edu; https://orcid.org/0000-0002-8067-3824
William R. Budnick: budnickw@msu.edu; https://orcid.org/0000-0001-9288-6782
Chad A. Larson: clar461@ecy.wa.gov; https://orcid.org/0000-0002-0329-8979
Sophia I. Passy: sophia.passy@uta.edu; https://orcid.org/0000-0002-8230-9380

*Corresponding Author

Introduction

A central question in metacommunity ecology is how environmental processes, and spatial processes shape community composition (MacArthur, 1972; Leibold *et al.*, 2004; Cottenie, 2005; Brown *et al.*, 2011). Both processes have been shown to control metacommunity composition, especially at large scales (Gravel *et al.*, 2006; Liu *et al.*, 2015). Large scale studies in freshwaters have also demonstrated environmental filtering, working in concert with dispersal limitation to determine the species composition of diatoms (Heino *et al.*, 2016), insects (Cañedo-Argüelles *et al.*, 2015; Brittain *et al.*, 2020), and fish (Blanchet *et al.*, 2014). However, the relative importance of environmental and spatial processes may depend on body size. In freshwater streams, small and better dispersing organisms, such as microbes, are more strongly structured by environmental niche processes (Beisner *et al.*, 2006; De Bie *et al.*, 2012), although spatial mechanisms, namely dispersal limitation, are still important, especially at large spatial scales (Vilmi *et al.*, 2016).

Conversely, very little is known about the relative importance of environmental and spatial processes on co-occurrence networks (but see Morueta-Holme *et al.*, 2016; Budnick *et al.*, 2021). Co-occurrence networks are collections of nodes and edges, representing respectively taxa and their significant pairwise correlations. The shape of a network – its topology – can be described in terms of number of nodes and edges, the connectivity among nodes, and its modularity. These metrics have ecological interpretations and are thought to be driven by both niche and dispersal processes. For instance, a more connected network will have more alternate pathways for connecting two taxa, thus reducing the impact of a node removal (i.e., a species extirpation) on overall network topology (Dunne *et al.*, 2002; Stouffer & Bascompte, 2011). A network characterized by a few well connected nodes can dramatically lower the average distance between all networked nodes, and allow disturbances to propagate more quickly (Shi *et al.*, 2016; Gilarranz *et al.*, 2017). Modularity characterizes the subdivision of the network into groups of more closely connected nodes due to dispersal (Carstensen *et al.*, 2012; Hu *et al.*, 2017) or niche similarity (Zhao *et al.*, 2016; Jones & Hallin, 2018). Network topology is sensitive to the environment, as shown in stream insect networks, which became more connected with water quality deterioration (Simons *et al.*, 2019) or as a result of seasonal variation in dissolved nutrients causing eutrophication-tolerant taxa to group together (Zhao *et al.*, 2016). Dispersal limitation can constrain network topology by reducing potential co-occurrences (Lentendu & Dunthorn, 2021), as was the case in stream biofilms (Widder *et al.*, 2014) and alpine plants (D'Amen *et al.*, 2018). Alternately, highly dispersive species can serve to link together disparate parts of a network (Borthagaray *et al.*, 2014), as was observed with highly abundant and dispersive microbes in the networks of a Chinese lake (Zhang *et al.*, 2020). However, niche and dispersal processes are rarely assessed simultaneously in networks, though in forest metacommunities and stream algae, these ecological processes had a strong impact on all aspects of network topology (Morueta-Holme *et al.*, 2016; Budnick *et al.*, 2021).

Notably, metacommunity composition and network topology may have distinct responses to the environment vs. dispersal limitation, as recently shown for algae inhabiting streams of different nutrient concentrations and stoichiometry (Budnick *et al.*, 2021). Specifically, niche effects and dispersal limitation on network topology were comparable, while the effect of dispersal limitation on metacommunity composition were stronger than niche effects. However, these results might have been influenced by the fact that nutrient content, which is a major driver of metacommunity composition, was already accounted for in the construction of the metacommunities. Therefore, it is still unclear whether the drivers of metacommunities vs.

networks are discrepant across stream organisms and dependent on body size. To ameliorate this gap, we performed a comprehensive comparison of environmental vs. spatial effects on metacommunity composition vs. co-occurrence networks of three major stream groups – diatoms, insects, and fish, differing greatly in body size and consequently, dispersal capacity. We predicted stronger spatial effects, in the form of dispersal limitation, than environmental effects on the metacommunities and networks of the more dispersal limited fish but relatively stronger environmental effects on the metacommunities and networks of the more dispersive diatoms and insects. We do not have any reason to predict differences in the strength of niche and dispersal effects between metacommunities vs. networks.

At large scales, co-occurrences are expected to be structured primarily by environmental filtering, which will group taxa based on niche overlap, and by dispersal limitation, grouping taxa spatially. The role of interspecific interactions is thought to be minimal, as these are generally local level processes (Zobel, 1997; Belmaker *et al.*, 2015), whereas environmental filtering and dispersal limitation will characterize large scale patterns of co-occurrence along environmental or biogeographic gradients. Strong habitat filtering at large scales may lead to increased phylogenetic similarity between co-occurring taxa (Webb *et al.*, 2002; HilleRisLambers *et al.*, 2012; Lentendu & Dunthorn, 2021) because phylogenetically related species may exhibit trait similarity due to phylogenetic niche conservatism (Wiens *et al.*, 2010; Fig. 3.1). Thus, we expected that if we controlled for environmental effects on network topology and eliminated co-occurrences that were explained by shared environmental preference, the co-occurrence networks composed of the remaining, unexplained edges, would show lower level of phylogenetic similarity. Much less is known about how dispersal limitation may affect the patterns of phylogenetic similarity. If the ability to disperse is conserved among closely related species as environmental preferences, then we would expect controlling for space vs. the environment to generate similar outcomes with respect to taxonomic similarity.

We used US stream data on diatoms, insects, and fish to address the following three objectives: 1) characterize the relative influence of environmental and spatial processes on metacommunity composition vs. network topology; 2) to test for evidence of phylogenetic niche conservation among those taxa that are connected within the networks and 3) evaluate whether the drivers in objectives 1) and 2) depend on body size. We hypothesize that metacommunity composition and network topology would share drivers, with environmental processes being relatively more important than dispersal processes for the highly dispersive diatom and insects groups. We further predicted that closely related taxa would be connected by links explained by environmental or spatial processes, following phylogenetic niche conservation.

Methods

Datasets:

We assembled three metacommunity datasets: stream diatoms, insects, and fish, each with 1400 sites across the conterminous United States (Fig. 3.2). Sites were compiled from both the National Water-Quality Assessment Program of the US Geologic Survey and the National Rivers and Streams Assessment of the US Environmental Protection Agency, which used similar collection methods. Samples were collected between 1993 and 2019, but the majority of the samples were collected between 2007 and 2010. Diatoms and insects were collected from a predetermined area of substrate in the warm months, May-September. Fish were sampled throughout the year using backpack electrofishing and seining. Diatoms and fish data consisted of species counts, and insect data, of genus counts.

In selecting our 1400 sites from datasets of 2175 diatom, 1760 insect, and 1617 fish samples, we controlled for spatial aggregation by ensuring that the minimum pairwise distance between any two sites was 1 km. We included the physicochemical parameters of nitrate (NO_3, µg/L), pH, total phosphorous (TP, µg/L), specific conductance (µS/cm), water temperature (C°), which were available from USGS or USEPA datasets. Elevation (m) was obtained from the WorldClim database (Fick & Hijmans, 2017) and used to calculate slope (% grade) with the package 'raster' (Hijmans *et al.*, 2020), and averaged across a 5 km buffer around all sites. Climate variables included mean temperature and precipitation minima, maxima, averages, ranges, and seasonality across a 5 km buffer extending from each site (WorldClim V1.4; 19 bioclimatic variables) (Hijmans *et al.*, 2005). Environmental variables were ln-transformed if normality was improved, except slope which was arcsine square root-transformed, and pH, which was not transformed. All variables were standardized (mean = 0, standard deviation = 1). Pairwise distances between all sites were calculated in km using the 'distm' function in the R package 'geosphere' (Hijmans *et al.*, 2019).

Co-occurrence Networks

We created a weighted network based on the partial Spearman correlation matrix of the relative standardized species abundances. To minimize spurious correlations caused by rare taxa, only those taxa occurring in 15 or more sites (1% of all sites) were including in network creation. We used random matrix thresholding in the R-package 'RMThreshold' (Menzel, 2016) with modification from Budnick *et al.* (2021) to objectively determine thresholds between meaningful and spurious correlations. The resultant adjacency matrix was used to create a weighted and an unweighted network with the R-package 'igraph' (Csárdi & Nepusz, 2006). In these networks, nodes represented taxa, while edges represented the co-occurrence between them. In the weighted network, the strength of the partial Spearman correlation was used as the edge weight (either positive or negative), while in the unweighted network all edges were given either a weight of 1 (present) or 0 (absent). We used a fast-greedy clustering algorithm (function: 'cluster_fast_greedy') to identify modules in the network. We then calculated the following network topology measures, node and edge count, mean node degree (mean number of edges per node), connectance, mean shortest path length, modularity, and number of modules. Mean shortest path length was calculated with a harmonic mean and was based on the unweighted network. Connectance was calculated as the ratio of observed over all possible edges. We also calculated the average number of edges between a certain node and those nodes within its module (intramodular degree, K_{in}), and outside its module (extramodular degree K_{out}). These parameters provide information on how connected or isolated modules are relative to the network as a whole (Guimerà & Amaral, 2005).

RDA and Variance Partitioning:

To analyze the effect of environment vs. space on metacommunity composition in each dataset, we performed redundancy analysis with variance partitioning, both using all taxa that occurred in at least 1% of sites, and again using only taxa that were included in the respective networks. We Hellinger transformed species abundances and utilized the first four axes of a PCA of the centered and standardized climatic and physicochemical parameters. For space, we used the first four axes of a PCA of the pairwise distance between all sites in km. This utilization of PCA axes was done to match the statistical procedures used to assess spatial effects on network topology (described in the next paragraph). We selected environmental and spatial variables with a forward selection procedure via ‘forward.sel’ in the R-package ‘adespatial’ (Dray *et al.*, 2021). Variance partitioning was conducted with ‘varpart’, and RDA with ‘rda’, both in ‘vegan’ (Oksanen *et al.*, 2019) on species abundances vs. significant environmental and spatial PCA axes. The RDA was tested for significance using a permutation ANOVA.

Gradient effects on co-occurrence networks:

To analyze the influence of environment, space, and their combination on co-occurrence patterns, we adopted the methodology outlined in Morueta-Holme *et al.* (2016) and the associated functions from the R-package ‘netassoc’ (Blonder & Morueta-Holme, 2017). Briefly, the ‘netassoc’ procedure compares the observed partial Spearman correlation values from our weighted network to the partial Spearman correlation scores derived from a probabilistic metacommunity. The probabilistic metacommunity was created by assigning individuals to sites based on their probability of occurrence derived from boosted regression trees (BRTs). These trees generated probabilities of occurrence for each species in each site based on environmental data, spatial data, or both. If a species was not observed in a site, we set its probability of occurrence to zero, to prevent assigning species to sites where they do not occur. This process was repeated 1000 times in order to build a standard effect size (SES) and a Benjamani–Hochberg adjusted p-value between the probabilistic and observed rates of co-occurrence. Significant edges were those that significantly differed from those generated by the gradient effect, and therefore were not explained, and, thus, were retained in the network, while non-significant edges were removed because they represented co-occurrences due to a shared response to a gradient.

We ran BRTs with a Bernoulli link-function, trained with 10-fold cross-validation (Elith *et al.*, 2008) independently on each species to calculate probability of occurrence. As BRTs are sensitive to the number of predictors, we used the first four PCA axes of each set of predictors – environment (climate and physicochemistry) and space. For space we based the PCA on the pair-wise distance between all sites in km. Our combined model incorporated all eight PCA axes. The BRTs were run with a bag fraction of 0.5, a tree complexity of 4, and maximum of 3000 trees. Only BRT results for those species with an AUC equal or above 0.7 were retained (between 70-85% of diatoms and insects taxa, and 95-99% of fish taxa). We converted the probability estimate generated by the BRTs to abundance by multiplying each species' probability by the observed abundance of the species. We then transformed this abundance dataset to relative abundance, thus creating a probability of occurrence matrix where probabilities were weighted by the observed abundance of each species. For species with an AUC ≤ 0.7, indicating the BRT failed to converge on a suitable model, we used their observed abundances in the predicted matrix.

We then used the predicted abundance matrices to reconstruct the networks, which now contained only edges that could not be explained by environment, space, or environment + space. A limitation to this approach, however, is that the null model may identify and subsequently eliminate the links between taxa which were not adequately predicted by BRT (taxa with an AUC ≤ 0.7). To avoid these issues, which could lead to an overestimation of the effect of environment or space, we ensured all edges originating from such taxa were maintained through the lottery process. As this process only generates one network per gradient effect, we qualitatively compared network parameters of the original observed network with the networks, controlled for environmental, spatial, and environmental + spatial effects.

Phylogenetic similarity:

We assembled higher-order taxa for each metacommunity, consisting of phylum, class, order, family, genus, species, variety, and form for diatoms; phylum, class, order, family, and genus for insects, and finally, phylum, class, order, family, genus, species, and subspecies for fish. We measured taxonomic distance between all species pairs linked in the networks with the argument 'taxa2dist' in the package 'Vegan' in R (Oksanen *et al.*, 2019), and represented average taxonomic distance of the species present in the network. Distance was scaled to 100, with 100 representing maximal distance (i.e., most distantly related taxa) and 0 representing maximal similarity (i.e., most closely related), to improve interpretability, we converted these data to phylogenetic similarity by subtracting the average distance from 100. To assess whether there was overdispersion, taxonomic similarities less than expected at random, or clustering, greater taxonomic similarity than expected at random (Warwick & Clarke, 1998; Clarke & Warwick, 2001), we randomly rewired the network 1000 times and measured taxonomic similarity. To preserve the network's structure while rewiring, we maintained the same node degree distribution in each pull, following an Erdös-Réyni null model. To prevent connections between species that cannot co-occur naturally, we only permitted edges between species that were found in the same two-digit hydrological unit (Seaber *et al.*, 1987). We then compared the average phylogenetic similarity of the observed network to the 95% confidence interval of the average phylogenetic similarity in the 1000 rewired networks. We then calculated average phylogenetic similarity in the networks controlled for environment, space, and both and compared them with the average taxonomic similarity of the observed networks with ANOVA followed by a Tukey's post-hoc test.

Results

Determining the drivers of metacommunity composition

To assess the role of environmental and spatial processes on metacommunity composition, we conducted RDA and variance partitioning on both the entire metacommunity dataset, and the subset of taxa that occurred in the networks. We found these datasets to provide very similar results (less than 1% difference), thus only the results of the networked taxa were discussed further. RDA showed that both environment and space significantly explained community composition in all three datasets. Generally, the combined model, environment + space, explained the most variance in community composition, between 11.4% and 16.8% of variance. In diatoms and insects, environment outperformed space, whereas space outperformed environment in fish (Fig. 3.3a). Variance partitioning provided similar results, though generally explained less variance, ranging from a total of 7.8% in insects to 16.8% in fish. Covariance between environment and space explained the greatest amount of variance, between 4% and 8.8%. Pure environment explained more than pure space in diatoms and insects, while in fish space outperformed pure environment (Fig. 3.3b).

Assessing the drivers of network topology

The raw diatom network was large, with 489 nodes and 886 edges, but not well connected, having a low average degree, 3.62, and low overall connectivity, 0.007. The diatom network was modular, with 15 modules, well connected to each other (K_{out} = 0.96), and an overall modularity of 0.65. Controlling for environment, space and both, greatly reduced network size, i.e., loss of both nodes and edges, reduced connectedness but increased the modularity. The three controls had similar effects on network topology, and all resulted in more fragmented networks with lower average node degree and higher modularity (Table 1, Fig. 3;4).

In insects, the raw network contained 256 nodes and 478 edges and had an average degree of 3.73 and a connectance of 0.015. It was modular, with a modularity of 0.62 and 14 modules which were well connected externally (K_{out} = 0.87) and internally (K_{in} = 2.87). The three controls produced similar networks, and all effects reduced the size, with the greatest reductions being driven by the combined effect model. Network parameters related to connectance: mean node degree, connectance, and path length depended on the control, as degree decreased in all cases (-48.1 – -56.8% change), generally indicated a more fragmented network, while connectance increased with a corresponding decrease in mean path length, under the space and combined gradient effects indicating the opposite. The networks controlled for gradient effect were also more modular (37.7 – 44.2% change), in all cases, with the greatest increase in the combined control network. In all cases, K_{out} and K_{in} declined, indicating that the modules were less connected, both to each other and internally (Table 1, Fig. 3;4).

In fish, the raw network was the smallest of the three datasets, with 160 nodes and 183 edges. It had low mean node degree (2.29), low connectance (0.014), but high mean path length (5.63). Additionally, the raw network was highly modular, with 19 modules and a modularity of 0.82, with an average K_{out} of only 0.25. The controls generated similar networks for fish, which were smaller in size, more connected, and more modular. Mean node degree decreased in the fish networks (-41.5 – -51.1% change) after control for gradient effects, while connectance increased (141 – 159% change) indicating the remaining nodes were better connected to each other. Modularity increased in all three controls, however there was a reduction in the number modules resulting from the overall decrease in network size. There was additionally a decrease in K_{in} as well as the complete loss of intermodular connections, K_{out}, indicating these smaller modules were isolated from each other (Table 1, Fig 3;4).

On average, there was a shift of between 45.2% and 71.4% in network topology following control of gradient effects, with the greatest changes in the fish networks, and the smallest changes in the insect networks. There were no large differences among the gradient effects (Fig. 5)

Measuring the drivers of phylogenetic similarity:

The mean phylogenetic similarity of connected species in the raw networks was greater than what would be expected by random chance in all three organismal groups. Diatoms had the lowest level of phylogenetic similarity, 12.4 (rewired = 5.7), followed by insects, 20.1 (rewired = 11.7), while fish had the greatest level of phylogenetic similarity, 39.0 (rewired = 29.8) (Fig. 6). In diatoms, the phylogenetic similarity of the raw vs. the controlled networks were not significantly different (ANOVA F = 1.895, p =0.128). Conversely, in insects, the raw network had significantly lower phylogenetic similarity than that of any controlled network, which did not differ from each other (ANOVA F = 34.97, Tukey $p < 0.01$). Fish exhibited the opposite trend: the phylogenetic similarity of the raw network was significantly greater than that of the controlled networks, which also did not differ from one another (ANOVA F = 12.24, Tukey $p < 0.01$; Fig. 7).

Discussion

Following our first and third objective, we compared the influence of environmental and spatial factors on metacommunity composition vs. network topology in organismal groups varying in body size and dispersal capacity. With respect to metacommunity composition, the combined effect of environment and space explained the greatest amount of variance in all three datasets, but much of this effect was due to covariance of environment and space, as indicated by variance partitioning. Similar dominance of the shared component has been reported in stream diatoms, insects and fish (Cottenie, 2005; Heino *et al.*, 2016). Individually, the environment outperformed space in both diatoms and insects, consistent with our expectation for a stronger environmental control in these small organisms. Conversely, in fish space was a stronger predictor than the environment, indicating that fish were subject to stronger dispersal limitation, also in agreement with expectation. Previous research in lakes and ponds found a greater influence of habitat filtering on small-bodied aquatic organisms, but of dispersal on large species (De Bie *et al.*, 2012; Padial *et al.*, 2014). Here we strengthen these claims by demonstrating similar patterns in stream organisms across a large subcontinental range.

With respect to network topology, the environmental and spatial effects were strong and comparable across the three studied groups, although individual topological metrics responded differently to the three controls. Similarly, metacommunity composition was driven primarily by the interaction of space and environment, with only minor differences in the pure components, as discussed. Our results were thus inconsistent with Budnick et al. (2021) who reported environmental processes, namely climatic filtering, dominating over spatial processes on both network topology and metacommunity composition. This discrepancy was likely due to our inclusion of physicochemical data explicitly in our assessment of environment, which is critical for freshwater algae (Passy *et al.*, 2018). We further extended these conclusions to insects and fish, providing further knowledge of the roles of niche and dispersal mechanisms along a broad gradient of dispersal capacity.

Networks controlled for environment, space, or both were substantially smaller and more modular, but their connectance and path length differed among the three groups. The reduction in size was most notable in the fish networks, which suggests that co-occurrences in fish were more strongly driven by both environmental and spatial processes. Smaller network size in fish was further associated with increased connectance and shorter path lengths. Thus, while the majority of the fish nodes were lost, those that remained were more connected. Connectance did not change appreciably in diatoms and insects following environmental and/or spatial control, and path length showed a divergent response—it increased in diatoms but decreased in insects. This could be due partially to the low initial connectance in our networks (0.01-0.03), likely a result of our large spatial scale, as other network studies conducted over smaller areas report higher connectance values of 0.05 to 0.3 (Thompson *et al.*, 2012). The increased connectance in the controlled fish networks may be evidence that the remaining taxa had strong ecological relationships with each other, such as facilitation through nest construction, as previously documented in stream fish networks (Peoples *et al.*, 2015). We additionally measured mean shortest path length, which is related to connectance as high connectance increases the number of alternate paths in a network, thus lowering path length. We confirmed this trend across our controls, except in the combined control of diatoms, which had a small increase in both connectance and path length.

All networks were modular, indicating high fragmentation (Newman & Girvan, 2004), however, modularity increased after the application of controls. Modules can indicate species that are grouped together due to shared niche or shared dispersal history (Jones & Hallin, 2018), and our results indicated that both processes were at play, as modules were broken up when we removed edges explainable by niche or spatial overlap. Our results further revealed a trend of decreasing intramodular degree, K_{in}, confirming that the modules in the controlled networks contained taxa with fewer connections to each other. We also observed a decrease in intermodular degree, K_{out}, allowing us to attribute intermodular connections, which are critical for network cohesion (Yang *et al.*, 2021), to spatial or environmental processes. These intermodular links could be due to cosmopolitan species occurring broadly. These taxa would have multiple co-occurrences with range restricted taxa represented in multiple different modules (McGarvey & Veech, 2018). Overall, we concluded that species groupings were mediated by both niche and dispersal processes, regardless of dataset.

In all cases networks were substantially reduced by the environment, space, and environment + space controls, i.e., between 68% and 90% of all edges were removed after control, suggesting that most co-occurrences were driven by abiotic processes. Co-occurrence networks have been used to assess species interactions, including competition (Berry & Widder, 2014) and facilitation (Blonder *et al.*, 2018), but at large scales evidence shows that edges are due to filtering or dispersal (Blanchet *et al.*, 2020). Here we provide the first empirical support of this notion in fish and insect networks, and confirm a previous finding in diatoms that network topology is strongly driven by environmental and spatial processes (Budnick *et al.*, 2021). While it is possible that some of our unexplained edges could be due to interactions as our methods do not account for biotic factors (Morueta-Holme *et al.*, 2016), it is more likely they were due to processes poorly accounted for in our controls, such as unmeasured niche dimensions (e.g., substrate quality).

To address our third objective, we assessed the relatedness of co-occurring taxa in our networks. We compared the phylogenetic similarity in raw vs. randomly assembled networks and found a pattern of phylogenetic clustering in all three datasets. These results revealed that co-occurring taxa were more phylogenetic similar than expected by chance, but the reasons for this pattern may differ across the three groups, as evident from our comparisons of taxonomic similarity in raw vs. controlled networks (discussed below). Previous work has shown that microbes were more likely to connect with closely related taxa (Hu *et al.*, 2017; Zhang *et al.*, 2020; Lentendu & Dunthorn, 2021), but here we detect this trend across eukaryotic organisms, ranging from diatoms to fish.

We next tested whether the controlled networks had higher taxonomic similarity relative to the observed network, which would occur if environmental niches and dispersal history were phylogenetically conserved. The three groups showed divergent patterns. In diatoms, phylogenetic distances in raw vs. controlled networks were not significantly different from each other, suggesting that neither environmental niches nor dispersal history were phylogenetically conserved. Previous research has shown that diatoms exhibit a weak niche conservatism due to their high level of local adaptation, particularly to climate (Soininen *et al.*, 2019). Here we broaden our knowledge by demonstrating the lack of phylogenetic signal in niche or dispersal mediated co-occurrences. This can be explained in part with the high dispersal ability of diatoms, which allows them to broadly occur where the environment permits them, irrespective of their phylogenetic relations (Soininen *et al.*, 2019). The raw insect network was had lower taxonomic similarity relative to the controlled networks, suggesting that closely related insects had divergent environmental preferences and spatial patterns. This was consistent with niche partitioning, which drives an inverse relationship between phylogenetic relatedness and environmental preference by selecting for species with similar traits (Cavender-Bares *et al.*, 2004; Starko *et al.*, 2020). Fish conformed to our prediction of phylogenetic niche conservatism, as taxonomic similarity decreased after controls, indicating that phylogenetic clustering was driven by shared niches and dispersal history. This conclusion was supported by past research documenting niche conservatism in freshwater fish, particularly among range restricted taxa (McNyset, 2009).

In this first comparative continental analysis of metacommunity composition vs. co-occurrence networks across stream organisms, we demonstrated that both metacommunity composition and co-occurrence networks share the same macroecological drivers of habitat filtering and dispersal limitation. We additionally confirmed that metacommunity drivers were weakly dependent on the dispersal ability, as diatoms and insects showed a greater contribution of pure environment relative to pure space, while fish demonstrated greater pure space, though covariance between environment and space was greatest in all cases. We did not, however, extend this conclusion to network topology drivers, as parity between environment and space was found across all three datasets. This allowed us to conclude that a majority of freshwater co-occurrence network topology was a result of environmental or spatial processes. Finally, we showed that, at large scales, these freshwater groups were more likely to co-occur with closely related taxa than expected by chance. We then provided novel insight into this pattern by demonstrating that taxa which co-occur due to shared environmental preference were closely related, consistent with phylogenetic niche conservation, only in fish, whereas the opposite, phylogenetic niche dispersion, was true in insects, and diatoms, due to their broad dispersal capacity demonstrated neither niche conservation nor dispersion.

References

Beisner, B.E., Peres-Neto, P.R., Lindstrom, E.S., Barnett, A. & Longhi, M.L. (2006) The role of environmental and spatial processes in structuring lake communities from bacteria to fish. *Ecology*, **87**, 2985–2991.

Belmaker, J., Zarnetske, P., Tuanmu, M.N., Zonneveld, S., Record, S., Strecker, A. & Beaudrot, L. (2015) Empirical evidence for the scale dependence of biotic interactions. *Global Ecology and Biogeography*, **24**, 750–761.

Berry, D. & Widder, S. (2014) Deciphering microbial interactions and detecting keystone species with co-occurrence networks. *Frontiers in Microbiology*, **5**, 1–14.

De Bie, T., De Meester, L., Brendonck, L., Martens, K., Goddeeris, B., Ercken, D., Hampel, H., Denys, L., Vanhecke, L., Van der Gucht, K., Van Wichelen, J., Vyverman, W. & Declerck, S.A.J. (2012) Body size and dispersal mode as key traits determining metacommunity structure of aquatic organisms. *Ecology Letters,* **15**, 740–747.

Blanchet, F.G., Cazelles, K. & Gravel, D. (2020) Co-occurrence is not evidence of ecological interactions. *Ecology Letters*, **23**, 1050–1063.

Blanchet, S., Helmus, M.R., Brosse, S. & Grenouillet, G. (2014) Regional vs local drivers of phylogenetic and species diversity in stream fish communities. *Freshwater Biology*, **59**, 450–462.

Blonder, B., Kapas, R.E., Dalton, R.M., Graae, B.J., Heiling, J.M. & Opedal, Ø.H. (2018) Microenvironment and functional-trait context dependence predict alpine plant community dynamics. *Journal of Ecology*, **106**, 1323–1337.

Blonder, B. & Morueta-Holme, N. (2017) Package 'netassoc': Inference of species associations from co-occurence data. *CRAN*.

Borthagaray, A.I., Arim, M. & Marquet, P.A. (2014) Inferring species roles in metacommunity structure from. *Proceedings of the Royal Society B-Biological Sciences*, **281**, 20141425.

Brittain, J.E., Heino, J., Friberg, N., Aroviita, J., Kahlert, M., Karjalainen, S.M., Keck, F., Lento, J., Liljaniemi, P., Mykrä, H., Schneider, S.C. & Ylikörkkö, J. (2020) Ecological correlates of riverine diatom and macroinvertebrate alpha and beta diversity across Arctic Fennoscandia. *Freshwater Biology*, 1–15.

Brown, B.L., Swan, C.M., Auerbach, D.A., Campbell Grant, E.H., Hitt, N.P., Maloney, K.O. & Patrick, C. (2011) Metacommunity theory as a multispecies, multiscale framework for studying the influence of river network structure on riverine communities and ecosystems. *Journal of the North American Benthological Society*, **30**, 310–327.

Budnick, W.R., Mruzek, J.L., Larson, C.A. & Passy, S.I. (2021) The impacts of nutrient supply and imbalance on subcontinental co-occurrence networks and metacommunity composition of stream algae. *Ecography*, **44**, 1109–1120.

Cañedo-Argüelles, M., Boersma, K.S., Bogan, M.T., Olden, J.D., Phillipsen, I., Schriever, T.A. & Lytle, D.A. (2015) Dispersal strength determines meta-community structure in a dendritic riverine network. *Journal of Biogeography*, **42**, 778–790.

Carstensen, D.W., Dalsgaard, B., Svenning, J.C., Rahbek, C., Fjeldså, J., Sutherland, W.J. & Olesen, J.M. (2012) Biogeographical modules and island roles: A comparison of Wallacea and the West Indies. *Journal of Biogeography*, **39**, 739–749.

Cavender-Bares, J., Ackerly, D.D., Baum, D.A. & Bazzaz, F.A. (2004) Phylogenetic overdispersion in Floridian oak communities. *American Naturalist*, **163**, 823–843.

Clarke, K.R. & Warwick, R.M. (2001) A further biodiversity index applicable to species lists: Variation in taxonomic distinctness. *Marine Ecology Progress Series*, **216**, 265–278.

Cottenie, K. (2005) Integrating environmental and spatial processes in ecological community dynamics. *Ecology Letters*, **8**, 1175–1182.

Csárdi, G. & Nepusz, T. (2006) The igraph software package for complex network research. *InterJournal, Complex Systems*, **1695**, 1–9.

D'Amen, M., Mod, H.K., Gotelli, N.J. & Guisan, A. (2018) Disentangling biotic interactions, environmental filters, and dispersal limitation as drivers of species co-occurrence. *Ecography*, **41**, 1233–1244.

Dray, S., Bauman, D., Guillaume Blanchet, F., Borcard, D., Clappe, S., Guenard, G., Jombart, T., Larocque, G., Legendre, P., Madi, N. & Wagner, H.H. (2021) Package 'adespatial.' *CRAN*.

Dunne, J.A., Williams, R.J. & Martinez, N.D. (2002) Food-web structure and network theory: The role of connectance and size. *Proceedings of the National Academy of Sciences*, **99**, 12917–12922.

Elith, J., Leathwick, J.R. & Hastie, T. (2008) A working guide to boosted regression trees. *Journal of Animal Ecology*, **77**, 802–813.

Gilarranz, L.J., Rayfield, B., Liñán-Cembrano, G., Bascompte, J. & Gonzalez, A. (2017) Effects of network modularity on the spread of perturbation impact in experimental metapopulations. *Science*, **357**, 199–201.

Gravel, D., Canham, C.D., Beaudet, M. & Messier, C. (2006) Reconciling niche and neutrality: The continuum hypothesis. *Ecology Letters*, **9**, 399–409.

Guimerà, R. & Amaral, L.A.N. (2005) Functional cartography of complex metabolic networks. *Nature*, **433**, 895–900.

Heino, J., Soininen, J., Alahuhta, J., Lappalainen, J. & Virtanen, R. (2016) Metacommunity ecology meets biogeography: effects of geographical region, spatial dynamics and environmental filtering on community structure in aquatic organisms. *Oecologia*, **183**, 121–137.

Hijmans, R.J., Cameron, S.E., Parra, J.L., Jones, P.G. & Jarvis, A. (2005) Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, **25**, 1965–1978.

Hijmans, R.J., Williams, E. & Vennes, C. (2019) Package 'geosphere.'

HilleRisLambers, J., Adler, P.B., Harpole, W.S., Levine, J.M. & Mayfield, M.M. (2012) Rethinking community assembly through the lens of coexistence theory. *Annual Review of Ecology, Evolution, and Systematics*, **43**, 227–248.

Hu, A., Ju, F., Hou, L., Li, J., Yang, X., Wang, H., Mulla, S.I., Sun, Q., Bürgmann, H. & Yu, C.P. (2017) Strong impact of anthropogenic contamination on the co-occurrence patterns of a riverine microbial community. *Environmental Microbiology*, **19**, 4993–5009.

Jones, C.M. & Hallin, S. (2018) Geospatial variation in co-occurrence networks of nitrifying microbial guilds. *Molecular Ecology*, **28**, 293–306.

Leibold, M.A., Holyoak, M., Mouquet, N., Amarasekare, P., Chase, J.M., Hoopes, M.F., Holt, R.D., Shurin, J.B., Law, R., Tilman, D., Loreau, M. & Gonzalez, A. (2004) The metacommunity concept: A framework for multi-scale community ecology. *Ecology Letters*, **7**, 601–613.

Lentendu, G. & Dunthorn, M. (2021) Phylogenetic relatedness drives protist assembly in marine and terrestrial environments. *Global Ecology and Biogeography*, **30**, 1532–1544.

Liu, Y., Tang, Z. & Fang, J. (2015) Contribution of environmental filtering and dispersal limitation to species turnover of temperate deciduous broad-leaved forests in China. *Applied Vegetation Science*, **18**, 34–42.

MacArthur, R.H. (1972) *Geographical ecology, patterns in the distribution of species,* 1st edn. Princeton University Press, Princeton, NJ.

McGarvey, D.J. & Veech, J.A. (2018) Modular structure in fish co-occurrence networks: A comparison across spatial scales and grouping methodologies. *PLoS ONE*, **13**, 1–20.

McNyset, K.M. (2009) Ecological niche conservatism in North American freshwater fishes. *Biological Journal of the Linnean Society*, **96**, 282–295.

Menzel, U. (2016) Package ‘RMThreshold’, Signal-noise separation in random matrices by using eigenvalue spectrum analysis. *CRAN*.

Morueta-Holme, N., Blonder, B., Sandel, B., McGill, B.J., Peet, R.K., Ott, J.E., Violle, C., Enquist, B.J., Jørgensen, P.M. & Svenning, J.C. (2016) A network approach for inferring species associations from co-occurrence data. *Ecography*, **39**, 1139–1150.

Newman, M.E.J. & Girvan, M. (2004) Finding and evaluating community structure in networks. *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics*, **69**, 1–16.

Oksanen, J., Blanchet, F.G., Friendly, M., Kindt, R., Legendre, P., Mcglinn, D., Minchin, P.R., O’hara, R.B., Simpson, G.L., Solymos, P., Stevens, H.M., Szoecs, E. & Wagner, H. (2019) Package “vegan”, Community Ecology Package. *CRAN*.

Padial, A.A., Ceschin, F., Declerck, S.A.J., De Meester, L., Bonecker, C.C., Lansac-Tôha, F.A., Rodrigues, L., Rodrigues, L.C., Train, S., Velho, L.F.M. & Bini, L.M. (2014) Dispersal ability determines the role of environmental, spatial and temporal drivers of metacommunity structure. *PLoS ONE*, **9**, 1–8.

Passy, S.I., Larson, C.A., Jamoneau, A., Budnick, W.R., Leboucher, T., Tison-Rosebery, J. &

Soininen, J. (2018) The biogeographical patterns of species richness and abundance distribution in stream diatoms are driven by climate and water chemistry. *The American Naturalist*, **192**, 607–617.

Peoples, B.K., Blanc, L.A. & Frimpong, E.A. (2015) Lotic cyprinid communities can be structured as nest webs and predicted by the stress-gradient hypothesis. *Journal of Animal Ecology*, **84**, 1666–1677.

Seaber, P.R., Kapinos, F.P. & Knapp, G.L. (1987) *Hydrologic unit maps*, Denver.

Shi, S., Nuccio, E.E., Shi, Z.J., He, Z., Zhou, J. & Firestone, M.K. (2016) The interconnected rhizosphere: High network complexity dominates rhizosphere assemblages. *Ecology Letters*, **19**, 926–936.

Simons, A.L., Mazor, R. & Theroux, S. (2019) Using co-occurrence network topology in assessing ecological stress in benthic macroinvertebrate communities. *Ecology and Evolution*, 1–13.

Soininen, J., Jamoneau, A., Rosebery, J., Leboucher, T., Wang, J., Kokociński, M. & Passy, S.I. (2019) Stream diatoms exhibit weak niche conservation along global environmental and climatic gradients. *Ecography*, **42**, 346–353.

Starko, S., Demes, K.W., Neufeld, C.J. & Martone, P.T. (2020) Convergent evolution of niche structure in Northeast Pacific kelp forests. *Functional Ecology*, **34**, 2131–2146.

Stouffer, D.B. & Bascompte, J. (2011) Compartmentalization increases food-web persistence. *Proceedings of the National Academy of Sciences*, **108**, 3648–3652.

Thompson, R.M., Dunne, J.A. & Woodward, G. (2012) Freshwater food webs: Towards a more fundamental understanding of biodiversity and community dynamics. *Freshwater Biology*, **57**, 1329–1341.

Vilmi, A., Karjalainen, S.M., Hellsten, S. & Heino, J. (2016) Bioassessment in a metacommunity context: Are diatom communities structured solely by species sorting? *Ecological Indicators*, **62**, 86–94.

Warwick, R.M. & Clarke, K.R. (1998) Taxonomic distinctness and environmental assessment. *Journal of Applied Ecology*, **35**, 532–543.

Webb, C.O., Ackerly, D.D., McPeek, M.A. & Donoghue, M.J. (2002) Phylogenies and community ecology. *Annual Review of Ecology and Systematics*, **33**, 475–505.

Widder, S., Besemer, K., Singer, G.A., Ceola, S., Bertuzzo, E., Quince, C., Sloan, W.T., Rinaldo, A. & Battin, T.J. (2014) Fluvial network organization imprints on microbial co-occurrence networks. *Proceedings of the National Academy of Sciences*, **111**, 12799–12804.

Wiens, J.J., Ackerly, D.D., Allen, A.P., Anacker, B.L., Buckley, L.B., Cornell, H. V., Damschen, E.I., Jonathan Davies, T., Grytnes, J.A., Harrison, S.P., Hawkins, B.A., Holt, R.D., McCain, C.M. & Stephens, P.R. (2010) Niche conservatism as an emerging principle in ecology and conservation biology. *Ecology Letters*, **13**, 1310–1324.

Yang, Y., Shi, Y., Kerfahi, D., Ogwu, M.C., Wang, J., Dong, K., Takahashi, K., Moroenyane, I.

& Adams, J.M. (2021) Elevation-related climate trends dominate fungal co-occurrence network structure and the abundance of keystone taxa on Mt. Norikura, Japan. *Science of the Total Environment*, **799**, 149368.

Zhang, H., Hou, F., Xie, W., Wang, K., Zhou, X., Zhang, D. & Zhu, X. (2020) Interaction and assembly processes of abundant and rare microbial communities during a diatom bloom process. *Environmental Microbiology*, **22**, 1707–1719.

Zhao, D., Shen, F., Zeng, J., Huang, R., Yu, Z. & Wu, Q.L. (2016) Network analysis reveals seasonal variation of co-occurrence correlations between cyanobacteria and other bacterioplankton. *Science of the Total Environment*, **573**, 817–825.

Zobel, M. (1997) The relative role of species pools in determining plant species richness: An alternative explanation of species coexistence? *Trends in Ecology and Evolution*, **12**, 266–269.

Table 3.1. Parameters of observed network and networks controlled for the environment, space, and environment + space (combined).

Group	Control	No. nodes	No. edges	Degree	Connectance	Mean shortest path	Modularity	No. modules	Kout	Kin
Diatoms	None	489	886	3.62	0.007	4.71	0.65	15	0.96	2.67
	Environment	292	281	1.92	0.007	7.39	0.88	37	0.14	1.79
	Space	288	281	1.95	0.007	6.36	0.87	36	0.15	1.81
	Combined	230	198	1.72	0.008	5.29	0.92	40	0.05	1.67
Insects	None	256	478	3.73	0.015	4.28	0.62	14	0.87	2.87
	Environment	131	119	1.82	0.014	4.56	0.85	27	0.15	1.66
	Space	130	126	1.94	0.015	3.54	0.85	22	0.14	1.80
	Combined	98	79	1.61	0.017	2.22	0.89	25	0.06	1.55
Fish	None	160	183	2.29	0.014	5.63	0.82	19	0.25	2.04
	Environment	32	18	1.13	0.038	1.10	0.91	14	0	1.13
	Space	34	19	1.12	0.035	1.12	0.91	15	0	1.12
	Combined	35	21	1.20	0.036	1.32	0.87	14	0	1.20

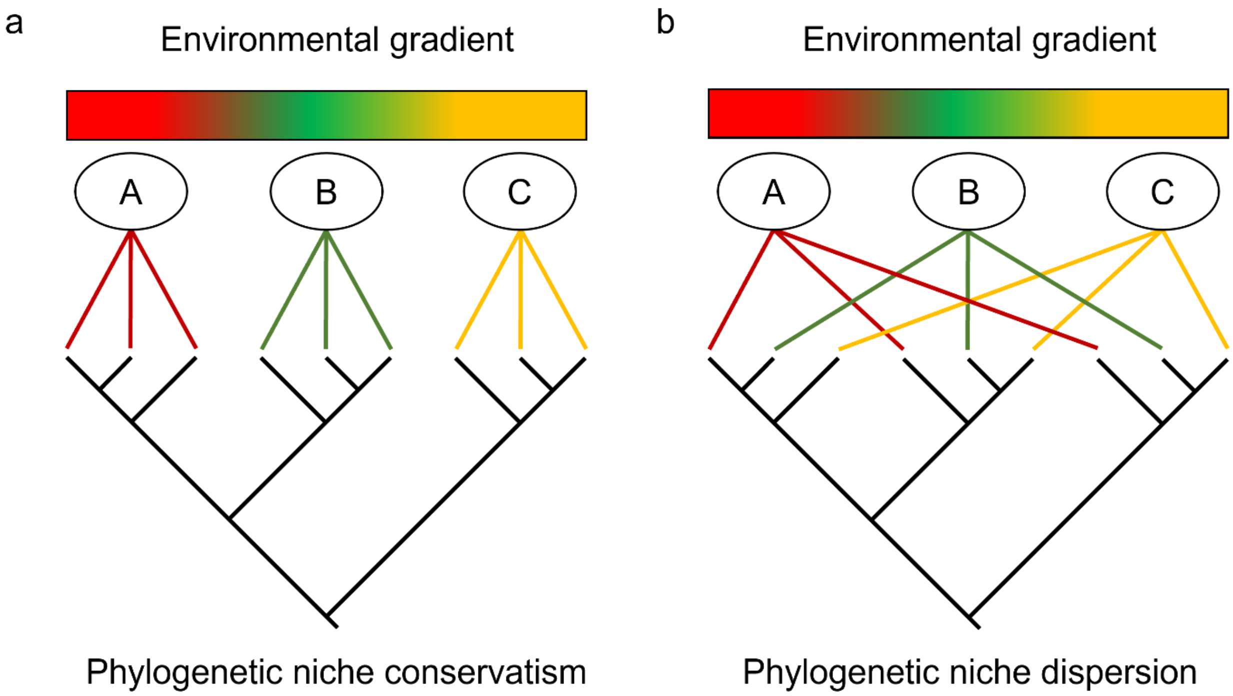

Fig. 3.1. Conceptual diagram of phylogenetic niche conservatism or dispersion. In scenario a), species (colored lines) co-occur in sites A, B or C based on color due to habitat filtering along and environmental gradient. Following phylogenetic niche conservatism, taxa that are closely related have similar traits, and therefore co-occurring taxa have taxonomic similarity, represented by the short distances between co-occurring taxa in the phylogenetic tree. In scenario b) species again co-occur in sites A, B or C based on color due to habitat filtering along an environmental gradient. However, due to phylogenetic niche dispersion, taxonomically similar species have divergent traits, represented by the longer distances between co-occurring taxa in the phylogenetic tree. Figure modified from Starko *et al.*, (2020).

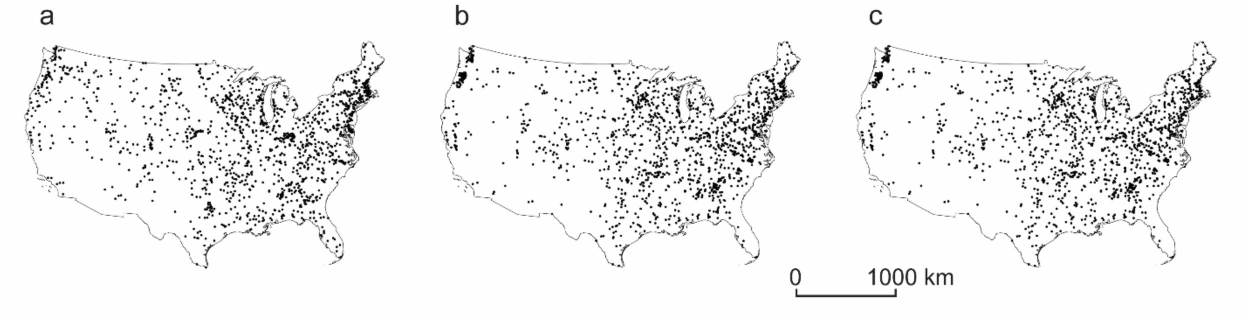

Fig. 3.2. Map of sampled diatom sites (a), insects sites (b) and fish sites (c). (Projection: EPSG: 2163).

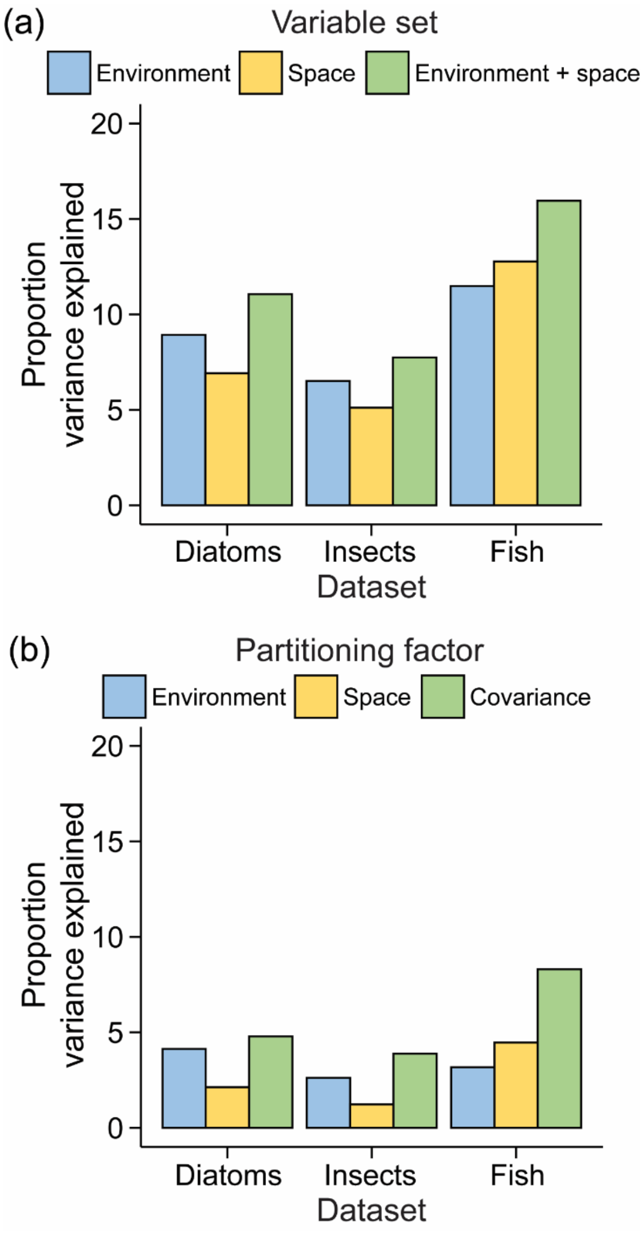

Fig. 3.3. (a) Adjusted R^2 from redundancy analysis of metacommunity composition against environmental, spatial, and environmental + spatial predictors. (b) Adjusted R^2 of variance partitioning, decomposing explained variance of metacommunity composition into pure environment, pure space, and their covariance.

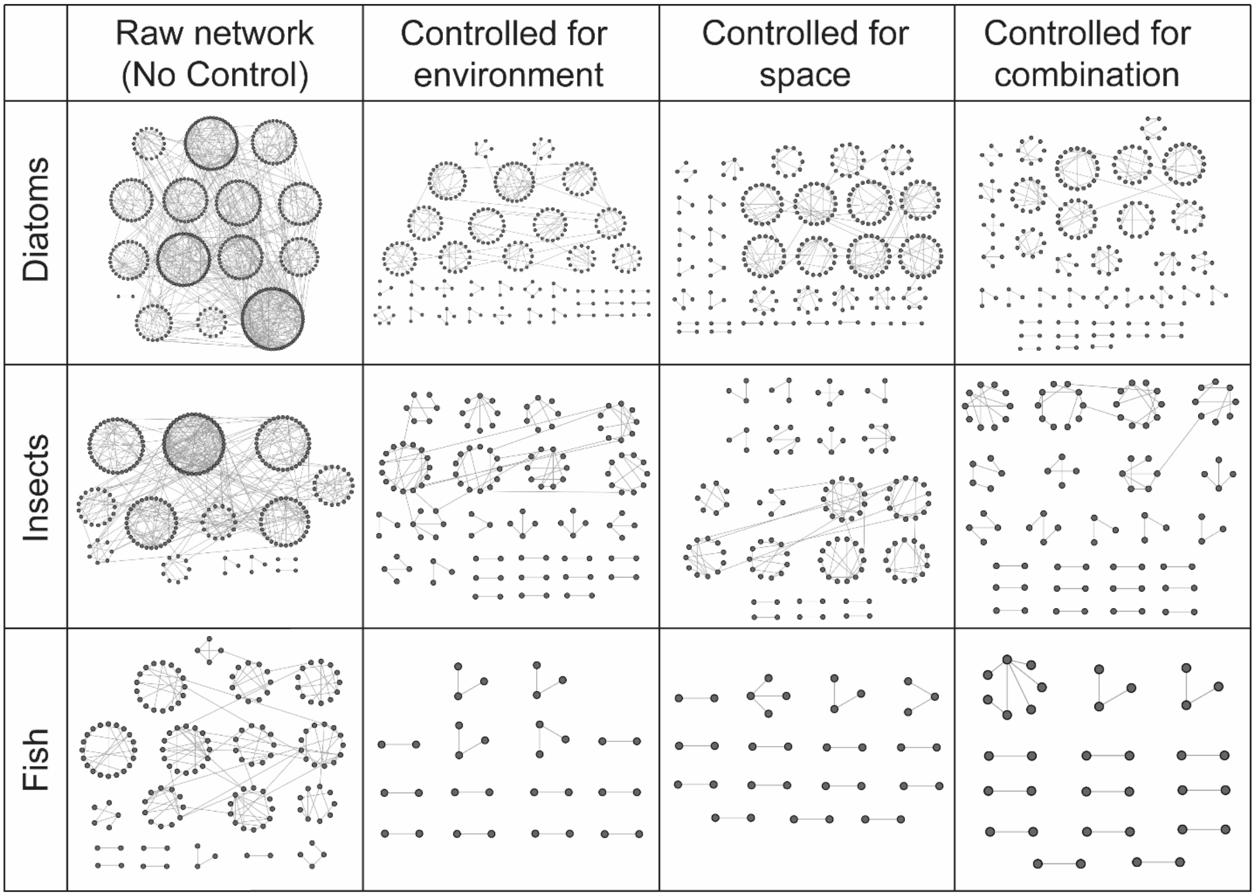

Fig 3.4. Visualizations of observed networks and the networks controlled for environment, space, and environment + space.

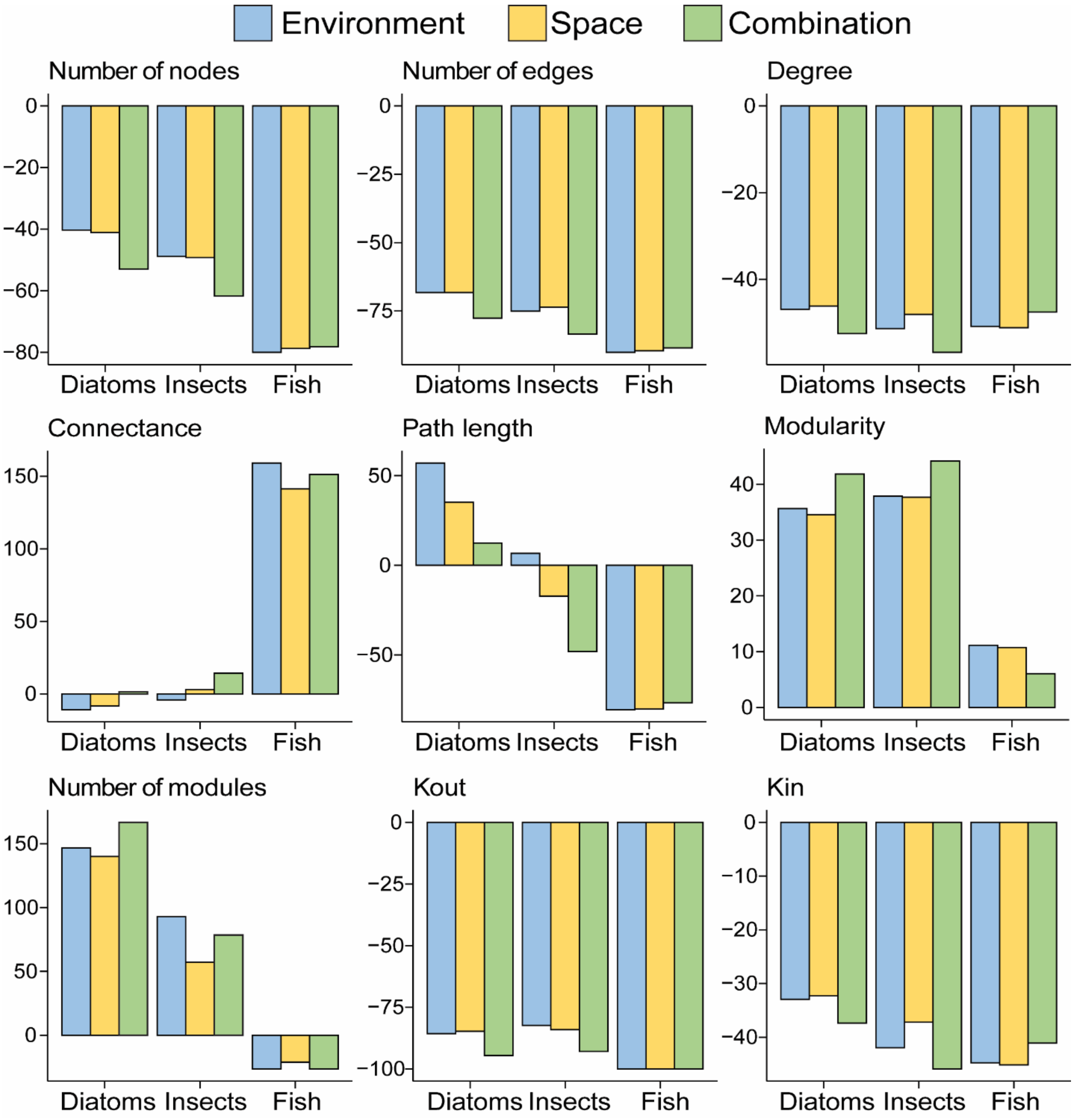

Fig 3.5. Percent change in each topologic parameter between the observed networks and the networks controlled for environment, space, and environment + space.

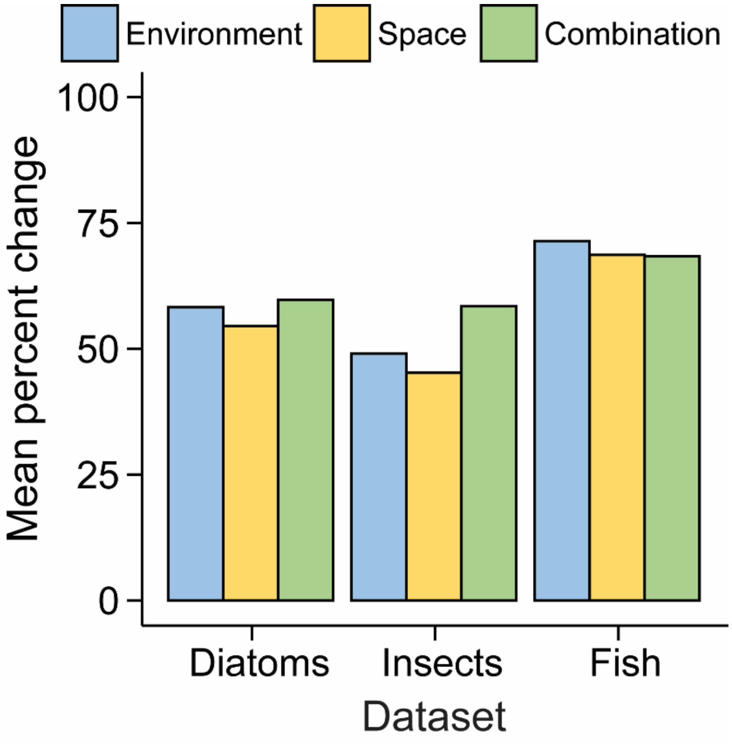

Fig. 3.6. Average percent change (absolute values) in network parameters between the observed networks and the networks controlled for environment, space, and environment + space.

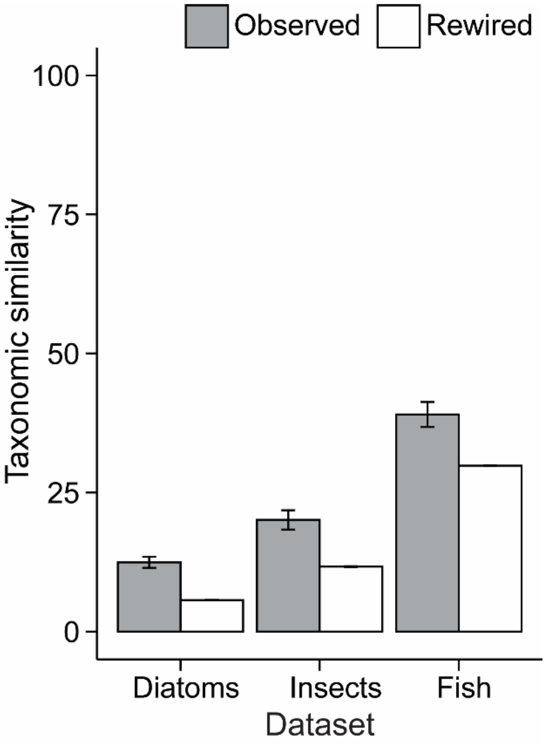

Fig. 3.7. Taxonomic similarities in the observed and rewired networks.

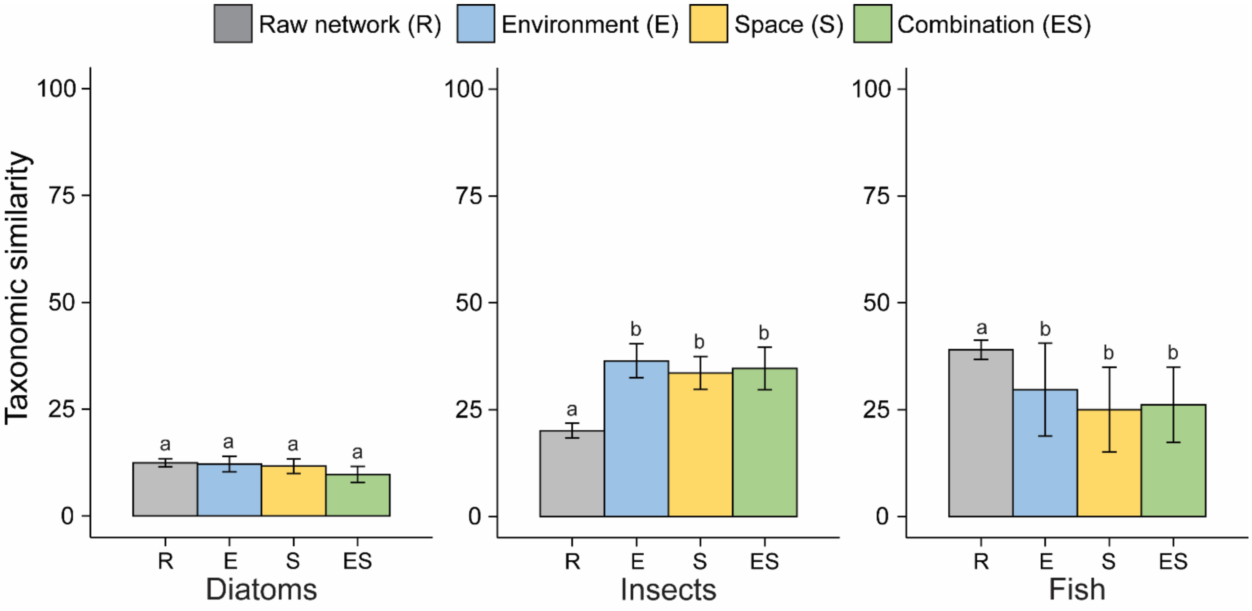

Fig. 3.8. Taxonomic similarity of the observed networks (R) vs. networks controlled for environment(E) space(S) and environment + space (ES). Error bars represent a 95% confidence interval of the mean taxonomic distance. Means were compared with ANOVA followed by a Tukey post-hoc test.

Chapter 4

Environmental filtering and dispersal effects on the relationship between local and regional species richness in streams depend on scale and organismal group

Joseph L. Mruzek[1], Chad A. Larson[2], William R. Budnick[3], Thibault Leboucher[4] and S.I. Passy[1*]

[1]Department of Biology, University of Texas at Arlington, 501 S. Nedderman Dr., Arlington Texas, 76019, USA

[2]Washington State Department of Ecology, Environmental Assessment Program, 300 Desmond Dr SE, Lacey, Washington, 98503, USA

[3]Department of Fisheries and Wildlife, Michigan State University, 480 Wilson Road, East Lansing, Michigan, 48824, USA

[4]Laboratory for Continental Environments, National Scientific Research Center, University of Lorraine, Metz, France

Authors' emails and ORCID IDs:
Joseph L. Mruzek: joseph.mruzek@mavs.uta.edu; https://orcid.org/0000-0002-8067-3824
Chad A. Larson: clar461@ecy.wa.gov; https://orcid.org/0000-0002-0329-8979
William R. Budnick: budnickw@msu.edu; https://orcid.org/0000-0001-9288-6782
Thibault Leboucher: thibault.leboucher@outlook.fr; https://orcid.org/0000-0003-0353-8896
Sophia I. Passy: sophia.passy@uta.edu; https://orcid.org/0000-0002-8230-9380

*Corresponding Author

Introduction

What mechanisms drive species richness is a central research topic in ecology. Local species richness, defined as the number of species in a single locality, is constrained by local processes, e.g., habitat filtering and species interactions, and regional processes, namely dispersal from the species pool (Leibold *et al.*, 2004). A common tool used by macroecologists to assess whether local or regional processes are dominant is the local species richness (LSR) - regional species richness (RSR) relationship. A curvilinear, saturating relationship is thought to be indicative of dominance of local processes, while a linear relationship, indicative of dominance of regional processes (Terborgh & Faaborg, 1980; Ricklefs, 1987). A saturated trend has been interpreted as indicative that all available niches are occupied and intense competition (Ricklefs, 1987; Cornell & Lawton, 1992), habitat filtering (Brown *et al.*, 2001), or disturbances (Harrison & Cornell, 2008; DeVantier *et al.*, 2020) prevents immigrating taxa from establishing. Conversely, a linear increase of local with regional richness indicated niche availability, and a community open to immigration (Cornell, 1993; Mateo *et al.*, 2017) but limited by the regional species pool (Caley & Schluter, 1997; Shurin *et al.*, 2000). Research has demonstrated that both saturated and unsaturated patterns are common across communities types (Szava-Kovats *et al.*, 2013; Cornell & Harrison, 2014). Linear trends were found in systems with strong regional environmental gradients, such as stream diatoms (Passy, 2009; Soininen *et al.*, 2009), macroinvertebrates (Grönroos & Heino, 2012), or weak local interactions, as observed in fish (Oberdorff *et al.*, 1998). Saturated patterns, on the other hand, were observed in communities structured by competition, e.g. in global fern communities (Weigand *et al.*, 2020), and are more likely in habitats with strong local environmental filters, e.g. in grasses (Michalet *et al.*, 2015).

Spatial scale affects the LSR-RSR relationship because the mechanisms driving this relationship operate on different scales--habitat filtering and species interactions at small scales and dispersal limitation at large scales (Hillebrand & Blenckner, 2002) This has been shown in diatoms, where the LSR-RSR relationship transitioned from curvilinear at small regional scales (a single stream) to linear at a large local and regional scales (multiple streams across a drainage basin) (Soininen *et al.*, 2009). Scale dependence of the LSR-RSR relationship was similarly demonstrated in rocky tidal invertebrates, though local scale rather than regional scale was varied, with small local scale driving saturation and large local scale resulting in linearity (Rivadeneira *et al.*, 2002). However, research with coral communities showed a linear trend regardless of the local scale considered, possibly owing to the high levels of local heterogeneity present in coral reefs (Cornell *et al.*, 2008).

In addition to spatial scale, the LSR-RSR relationship may be related to body size and dispersal ability of the organisms, which are inversely correlated in freshwater streams (De Bie *et al.*, 2012; Padial *et al.*, 2014). Larger species with poor dispersal capacity may possess non-saturated LSR-RSR relationships due to dispersal limitation, while small species with good dispersal abilities may be more strongly controlled by the environment and have saturated LSR-RSR relationships (Hillebrand & Blenckner 2002). Evidence for this pattern is conflicting, given that both microscopic diatoms (Passy 2009) and large bodied fish (Oberdorff *et al.*, 1998) exhibited linear LSR-RSR relationships. However, this discrepancy could be a result of differences in scale, given that the study of Passy (2009) was sub-continental, where dispersal limitation is strong and may have resulted in linear LSR-RSR. To disentangle the effects of scale and body size on the LSR-RSR relationship, we examined this relationship while maintaining constant local scale (single stream reach), while varying the regional scale, generating twelve distinct regional sizes ranging from 160,000 km^2 to 6,760,000 km^2. We analyze three major stream species groups, diatoms, insects, and fish, which differ greatly in body size and dispersal ability, allowing us to further elucidate the drivers of local richness.

Here we combined a recently developed landscape window approach (Leboucher *et al.*, 2019) with multiple regression and variance partitioning to test not only the scale dependence of the LSR-RSR relationships, but also the scale dependence of the drivers of local richness of stream organisms. Landscape windows (windows henceforth) were squares, representing differing scales. We first hypothesized that as regional scale increased, the LSR-RSR relationship would shift from curvilinear to linear as a result of reduced relative importance of environmental filtering but increased relative importance of dispersal limitation. Environmental filtering was assessed by measuring the pure influence of local physicochemistry and climate, while dispersal limitation was measured by the pure effect of regional richness on median local richness in each landscape window. We secondly hypothesized that the shift from curvilinear to linear LSR-RSR relationship would occur at smaller scales for less dispersive species, such as fish.

Methods

Datasets

Stream diatoms, insects, and fish were collected from 2278, 2270 and 2296 distinct localities, respectively, by the National Water-Quality Assessment (NAWQA) Program of the US Geological Survey and the National Rivers and Streams Assessment (NRSA) of the US Environmental Protection Agency (Fig. 4.1). Samples were taken between 1993 and 2019, though the majority of sampling occurred in 2008 and 2009 for diatoms and insects, and in 2008, 2009, 2013, and 2014 for fish. Diatoms and insects were sampled from a predefined area of stream substrate, while fish were collected with electrofishing and seines. Diatoms and fish were identified to species, and insects to genus. As local richness depends on the number of counted individuals and regional richness depends on the number of counted samples, we subsampled down each sample to 400 individuals for diatoms and 100 individuals for insects and fish and pulled at random 60 sites from each window (Fig. 2a, b). Climate data were retrieved from the WorldClim V1.4 database (Hijmans *et al.*, 2005) using a five-kilometer buffer for each site. They included temperature and precipitation averages, minima, maxima, ranges, seasonality, and extremes (19 variables). For each site, watershed and physicochemistry data were available, including slope (percent grade) and elevation (m), each based on a five-kilometer buffer, drainage area (km^2), as well as pH, water temperature (°C), specific conductance (µS/cm), total phosphorus (µg/L) and nitrates (µg/L).

Spatial design

We defined the local scale as a single stream reach, but varied regional scale by utilizing twelve differently sized landscape windows, following Leboucher *et al.*, (2019). We defined the windows as squares ranging from 16,000 km^2 (400 km edges) to 6,760,000 km^2 (2600 km edges) by increasing edge size by 200 km. To improve readability, we refer to windows by their edge size rather than area. To ensure that each window was adequately sampled, we further subdivided it into 9 sub-windows (3 × 3 grid) and only considered a window valid if 7/9 of the sub-windows contained samples. Invalid windows were resampled until the 7/9 rule was satisfied. Each window contained unique sites because they overlapped by at most two thirds. Some overlap was necessary to cover the whole US, particularly for the largest windows, as two adjacent 2600 km windows do not fit within the US, and would fail to meet the 7/9 requirement. To control for pseudoreplication due to windows overlapping and sharing samples, we implemented a randomization approach, whereby windows were subsampled at random without replacement in proportion to the number of sites it contained. We pulled 1000 times 60 sites at each scale, generating 1000 windows for each scale or a total of 12,000 windows per species group.

Statistical analysis

We calculated median local richness and regional richness (γ-diversity) from the 60 selected sites in each window. We compared the relationship between local richness (LSR) and regional richness (RSR) with quadratic regression.

(1) $Y = b_0 + b_1X + b_2X^2 + \varepsilon$

Where Y was median local richness, X was centered and standardized regional richness, b_0 was the intercept, b_1 was slope, and b_2 determined the curvature of the relationship. Median local richness was chosen over mean local richness in order to reduce the effect of extremes in local richness. To assess if saturation of the LSR-RSR relationship was scale dependent, we regressed parameter b_2, which describes the shape of the quadratic fit, against scale. When parameter b_2 is significantly negative, the fit is concave and indicative of saturation. When parameter b_2 is significantly positive or non-significant, the fit is, respectively, convex, or linear and indicative of non-saturation.

We also calculated the median of all physicochemical and climatic parameters in each window, as well as the heterogeneity of climate and physicochemistry as median Euclidean distance. To assess the drivers of median local richness across scales, we performed multiple regression with variance partitioning at each scale. The median local richness was the dependent variable, while the centered and standardized regional richness and its square, the centered and standardized environmental medians and their squares, and the centered and standardized environmental heterogeneity and its square were the predictors. We selected regional richness and environmental variables through forward selection using AIC. Then, we performed variance partitioning and determined the pure effect of regional richness, the pure effect of the environment, and their covariance effect as adjusted R^2. To assess the scale dependency of the drivers of median local richness, we regressed the adjusted R^2 of all terms against the log-transformed scale and its square. To assess the drivers of saturation, we regressed parameter b_2 against the adjusted R^2 of all variance partitioning terms.

Results

LSR-RSR relationship

The form of the LSR-RSR relationship, as assessed by parameter b_2, was dependent on scale in all groups, but a transition from generally curvilinear to linear pattern was observed only in diatoms and insects. In fish, the relationship remained curvilinear across scales. Specifically, diatoms were unsaturated at the smallest regional scale (400 km), saturated from the 600 km to 2000 km scale, and unsaturated concave from 2200 km to 2600 km scale. Similarly, insects also were unsaturated at the 400 km and 600 km scale, and saturated from 800 km to 1400 km. However, the b_2 parameter was non-significant from 1600 km to 2400 km, and positive in the 2600 km window, indicating linearity. Fish were significantly saturated from the smallest to the largest window, with the exception of the 2000 km window. However, the b_2 parameter in fish, while significant, was small and trended towards zero and lacked a clear asymptote from the

1800 km and larger windows, suggesting that local diversity was not in fact limited (Fig. 4.3 – 4.5). Regressing parameter b_2 against log-scale revealed a parabolic relationship in all three datasets with maximal saturation occurring in the middle extents (Fig. 4.6).

Drivers of local richness

We examined the contribution of regional richness, environment, and their covariance to median local richness with multiple regression and variance partitioning. In diatoms, we observed significant quadratic relationships with scale for pure regional richness, pure environment, and the covariance term (Fig. 4.7a). The regional richness effect increased with scale, the environmental effect peaked at smaller scales, and the covariance effects decreased with scale. A comparison of the pure effects showed that the environment was more important at all but the largest scales. In insects, the pure environmental effect and the covariance effect showed significant correlations with scale, similar to diatoms, while the regional richness was not significantly related to scale (Fig. 4.7b). The environmental effect peaked at the 1000 km window and remained dominant in all large windows. The covariance effect decreased with scale and peaked at 400 km window. In fish, the environmental effect had a U-shaped relationship with scale (Fig. 4.7c). The role of regional richness marginally increased with scale and explained the smallest portion of variance at all scales. The covariance effect was hump-shaped and dominant in all but the smallest scale. Pure regional richness had an upward parabolic relationship with insect b_2, and a downward relationship with fish b_2, and pure environment had a negative relationship with b_2 in both diatoms and insects, while covariance had significant upward parabolic relationships in diatoms and insects, but none in fish (Fig. 4.8).

Discussion

Our first hypothesis predicted a shift from saturated to non-saturated LSR-RSR relationship with spatial scale due to reduced environmental filtering and increased dispersal limitation. We partially confirmed this hypothesis in diatoms, and insects, but rejected it in fish. In diatoms, the LSR-RSR relationship generally transitioned with scale as expected, with the exception of the smallest scale, where the relationship was linear. These results are consistent with Soininen *et al.* (2009), who observed that LSR-RSR was saturated when regional richness was assessed at small scales, a single stream reach, but linear when large scales, a drainage basin with multiple streams, were considered. However, we further related the changes in the shape of the LSR-RSR relationship to driving mechanisms. In our study, these changes were indeed associated with a decrease in the role of pure environment with scale and an increasing role of regional richness in determining median local richness. When the LSR-RSR relationship was saturated, environment filtering was generally the strongest. Previous studies have reported strong filtering from water chemistry and climate on local diatom richness (Passy, 2009; Soininen *et al.*, 2016; Passy *et al.*, 2018), but here we show that the strength of this process is scale-dependent. The linearity of the LSR-RSR at large scales, as well as the increase in the importance of regional richness, indicated that dispersal limitation was influential for local richness, adding to a growing body of research demonstrating that communities of good dispersers, such as microorganisms, are not structured solely by environmental filtering but also by dispersal limitation (Vilmi *et al.*, 2016; Bestová *et al.*, 2018).

In insects, the LSR-RSR relationship was unsaturated at smallest scales, 400 km, and 600 km, saturated at mid-scales, but flat or non-saturated at scales of 1600 km or larger. At the scales where the LSR-RSR relationship was saturated, the environmental effect was the strongest, suggesting that local control of median species richness may have resulted in curvilinearity of the LSR-RSR relationship. However, the pure regional richness effect on median species richness was generally low, and the environment always had a stronger effect, implying that dispersal limitation may have had little influence. Previous research in stream macroinvertebrates at the guild level similarly found a minor contribution of regional richness to local richness variability relative to the environment, however, the LSR-RSR relationship was linear (Grönroos & Heino, 2012). In contrast, an investigation of head water streams (Heino *et al.*, 2003) and an Australian study over a relatively small regional size (Marchant *et al.*, 2006) found linear RSR-LSR and evidence that regional richness more than the environment constrained local richness. Nevertheless, none of these three studies considered the effect of scale, which we show is inherently linked with the shape of the LSR-RSR relationship and its drivers.

In diatoms and insects, the LSR-RSR relationship was linear at the smallest scales before transitioning to a saturated pattern in the middle scales, which we attribute to mass effects. Mass effects offset the influence of local processes, such as environmental filtering, by allowing species to persist in unfavorable habitats through immigration (Shmida & Wilson, 1985) and may thus prevent local richness from saturating. Previous studies have demonstrated mass effects in diatoms (Jamoneau *et al.*, 2018; Leboucher *et al.*, 2020) and aquatic insects (Heino & Tolonen, 2017). However, to our knowledge, this process has not been implicated as a driver of the LSR-RSR relationship.

In diatoms and insects, the presence and level of saturation, measured by parameter b_2, was determined by environmental effects. Positive parameter b_2 values and values approaching zero, indicative, respectively of lack of saturation or weak saturation, were generally associated with weaker environmental filtering. This is consistent with the original ideas of Terborgh &

Faaborg (1980), and Ricklefs (1987), who both posited that environmental conditions can limit the number of species that can persist at a site. This follows as past stream observations have revealed that local environmental conditions, such as climate and water chemistry (Passy *et al.*, 2018; Brittain *et al.*, 2020), limit local richness (Poff, 1997; Heino, 2011).

In fish, the LSR-RSR relationship also trended towards linearity with increasing spatial scales, which was associated with a significant increase in the effect of pure regional richness, confirming our prediction. However, unlike diatoms and insects, where parameter b_2 decreased under stronger filtering, there was no clear connection between the shape of the fish LSR-RSR relationship and the strength of the environmental processes. Notably, most of the variance in median local richness in fish was explained by the covariance of environment and regional richness, indicating that environment influenced local richness primarily indirectly through regional richness. These results contrasted past research on the LSR-RSR relationship in fish which found linearity (Caley & Schluter, 1997; Oberdorff *et al.*, 1998), and attributed it to a strong influence of regional richness on local richness. We attribute this discrepancy to our inclusion of more environmental variables which could explain fish range limits and therefore regional richness patterns, namely climate (Griffiths, 2015), and physical dispersal barriers such as mountains (represented in our data with elevation and slope), or disconnected drainage basins (Carvajal-Quintero *et al.*, 2019). These barriers may also have led to the limited, though significant, saturation at large scales, as indicated by negative b_2 parameters. Many fish have limited ranges which were smaller than our larger windows (Chapter 2), which could have caused us to overestimate regional richness.

Our second hypothesis predicted that a transition from saturated to unsaturated LSR-RSR relationships would occur at a smaller scale for dispersal limited species, such as fish. We rejected this hypothesis, as diatoms and insects showed lack of saturation at the smallest scales, potentially due to mass effects, and fish remained saturated at all scales even though they became only weakly curvilinear at these scales. Saturation, albeit weak, at large scales for fish may have been a product of their lower dispersal ability, which drove smaller average range size, which precludes most taxa from being able to disperse to all sampled sites, and decoupling the regional species pool from local richness (Cornell & Lawton, 1992; Shurin *et al.*, 2000). Local processes of habitat filtering and limiting similarity were shown to constrain local richness in French fish (Blanchet *et al.*, 2014), but here we demonstrated that the effect of these processes was both direct and indirect through regional species richness.

In conclusion, we identified scale dependent relationships in the drivers of median local richness in stream diatoms insects, and fish. In both diatoms and insects, environmental filtering decreased with scale and was mostly responsible for the transition from saturated to non-saturated LSR-RSR relationship. In diatoms, and fish, this transition was further related to increased dispersal limitation. In fish, environmental filtering and regional richness operated jointly in determining median local richness. While we partially confirmed our second hypothesis that dispersal capacity would determine the scale at which dispersal limitation would occur, insects appeared to be less dispersal limited than diatoms, counter to expectations. Finally, our results indicate that the LSR-RSR relationship may be driven by mass effects at small scales in good dispersers, such as diatoms and insects.

References

Bestová, H., Munoz, F., Svoboda, P., Škaloud, P. & Violle, C. (2018) Ecological and biogeographical drivers of freshwater green algae biodiversity: from local communities to large-scale species pools of desmids. *Oecologia*, **186**, 1017–1030.

De Bie, T., De Meester, L., Brendonck, L., Martens, K., Goddeeris, B., Ercken, D., Hampel, H., Denys, L., Vanhecke, L., Van der Gucht, K., Van Wichelen, J., Vyverman, W. & Declerck, S.A.J. (2012) Body size and dispersal mode as key traits determining metacommunity structure of aquatic organisms. *Ecology Letters*, **15**, 740–747.

Blanchet, S., Helmus, M.R., Brosse, S. & Grenouillet, G. (2014) Regional vs local drivers of phylogenetic and species diversity in stream fish communities. *Freshwater Biology*, **59**, 450–462.

Brittain, J.E., Heino, J., Friberg, N., Aroviita, J., Kahlert, M., Karjalainen, S.M., Keck, F., Lento, J., Liljaniemi, P., Mykrä, H., Schneider, S.C. & Ylikörkkö, J. (2020) Ecological correlates of riverine diatom and macroinvertebrate alpha and beta diversity across Arctic Fennoscandia. *Freshwater Biology*, 1–15.

Brown, J.H., Morgan Ernest, S.K., Parody, J.M. & Haskell, J.P. (2001) Regulation of diversity: Maintenance of species richness in changing environments. *Oecologia*, **126**, 321–332.

Caley, M.J. & Schluter, D. (1997) The relationship between local and regional diversity. *Ecology*, **78**, 70–80.

Carvajal-Quintero, J., Villalobos, F., Oberdorff, T., Grenouillet, G., Brosse, S., Hugueny, B., Jézéquel, C. & Tedesco, P.A. (2019) Drainage network position and historical connectivity explain global patterns in freshwater fishes' range size. *Proceedings of the National Academy of Sciences of the United States of America*, **116**, 13434–13439.

Cornell, H. V. (1993) *Unsaturated patterns in species assemblages: the role of regional processes in setting local species richness*. *Species diversity in ecological communities* (ed. by R.E. Ricklefs) and D. Schluter), pp. 243–252. University of Chicago Press, Chicago.

Cornell, H. V. & Harrison, S.P. (2014) What are species pools and when are they important? *Annual Review of Ecology, Evolution, and Systematics*, **45**, 45–67.

Cornell, H. V., Karlson, R.H. & Hughes, T.P. (2008) Local-regional species richness relationships are linear at very small to large scales in west-central Pacific corals. *Coral Reefs*, **27**, 145–151.

Cornell, H. V. & Lawton, J.H. (1992) Species interactions, local and regional processes, and limits to the richness of ecological communities: A theoretical perspective. *The Journal of Animal Ecology*, **61**, 1–12.

DeVantier, L., Turak, E. & Szava-Kovats, R. (2020) Species richness and abundance of reef-building corals in the Indo-West Pacific: The local–regional relation revisited. *Frontiers in Marine Science*, **7**, 1–19.

Griffiths, D. (2015) Connectivity and vagility determine spatial richness gradients and diversification of freshwater fish in North America and Europe. *Biological Journal of the Linnean Society*, **116**, 773–786.

Grönroos, M. & Heino, J. (2012) Species richness at the guild level: Effects of species pool and local environmental conditions on stream macroinvertebrate communities. *Journal of Animal Ecology*, **81**, 679–691.

Harrison, S. & Cornell, H. (2008) Toward a better understanding of the regional causes of local community richness. *Ecology Letters*, **11**, 969–979.

Heino, J. (2011) A macroecological perspective of diversity patterns in the freshwater realm. *Freshwater Biology*, **56**, 1703–1722.

Heino, J., Muotka, T. & Paavola, R. (2003) Determinants of macroinvertebrate diversity in headwater streams: Regional and local influences. *Journal of Animal Ecology*, **72**, 425–434.

Heino, J. & Tolonen, K.T. (2017) Ecological drivers of multiple facets of beta diversity in a lentic macroinvertebrate metacommunity. *Limnology and Oceanography*, **62**, 2431–2444.

Hijmans, R.J., Cameron, S.E., Parra, J.L., Jones, P.G. & Jarvis, A. (2005) Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, **25**, 1965–1978.

Hillebrand, H. & Blenckner, T. (2002) Regional and local impact on species diversity - from pattern to processes. *Oecologia*, **132**, 479–491.

Jamoneau, A., Passy, S.I., Soininen, J., Leboucher, T. & Tison-Rosebery, J. (2018) Beta diversity of diatom species and ecological guilds: Response to environmental and spatial mechanisms along the stream watercourse. *Freshwater Biology*, **63**, 62–73.

Leboucher, T., Budnick, W.R., Passy, S.I., Boutry, S., Jamoneau, A., Soininen, J., Vyverman, W. & Tison-Rosebery, J. (2019) Diatom β-diversity in streams increases with spatial scale and decreases with nutrient enrichment across regional to sub-continental scales. *Journal of Biogeography*, **46**, 734–744.

Leboucher, T., Tison-Rosebery, J., Budnick, W.R., Jamoneau, A., Vyverman, W., Soininen, J., Boutry, S. & Passy, S.I. (2020) A metacommunity approach for detecting species influenced by mass effect. *Journal of Applied Ecology*.

Leibold, M.A., Holyoak, M., Mouquet, N., Amarasekare, P., Chase, J.M., Hoopes, M.F., Holt, R.D., Shurin, J.B., Law, R., Tilman, D., Loreau, M. & Gonzalez, A. (2004) The metacommunity concept: A framework for multi-scale community ecology. *Ecology Letters*, **7**, 601–613.

Marchant, R., Ryan, D. & Metzeling, L. (2006) Regional and local species diversity patterns for lotic invertebrates across multiple drainage basins in Victoria. *Marine and Freshwater Research*, **57**, 675–684.

Mateo, R.G., Mokany, K. & Guisan, A. (2017) Biodiversity models: What if unsaturation is the rule? *Trends in Ecology and Evolution*, **32**, 556–566.

Michalet, R., Maalouf, J.P., Choler, P., Clément, B., Rosebery, D., Royer, J.M., Schöb, C. &

Lortie, C.J. (2015) Competition, facilitation and environmental severity shape the relationship between local and regional species richness in plant communities. *Ecography*, **38**, 335–345.

Oberdorff, T., Hugueny, B., Compin, A. & Belkessam, D. (1998) Non-interactive fish communities in the coastal streams of North-Western France. *Journal of Animal Ecology*, **67**, 472–484.

Padial, A.A., Ceschin, F., Declerck, S.A.J., De Meester, L., Bonecker, C.C., Lansac-Tôha, F.A., Rodrigues, L., Rodrigues, L.C., Train, S., Velho, L.F.M. & Bini, L.M. (2014) Dispersal ability determines the role of environmental, spatial and temporal drivers of metacommunity structure. *PLoS ONE*, **9**, 1–8.

Passy, S.I. (2009) The relationship between local and regional diatom richness is mediated by the local and regional environment. *Global Ecology and Biogeography*, **18**, 383–391.

Passy, S.I., Larson, C.A., Jamoneau, A., Budnick, W.R., Leboucher, T., Tison-Rosebery, J. & Soininen, J. (2018) The biogeographical patterns of species richness and abundance distribution in stream diatoms are driven by climate and water chemistry. *The American Naturalist*, **192**, 607–617.

Poff, N.L. (1997) Landscape filters and species traits: towards mechanistic understanding and prediction in stream ecology. *Journal of the North American Benthological Society*, **16**, 391–409.

Ricklefs, R.E. (1987) Community diversity: relative roles of local and regional processes. *Science*, **235**, 167–171.

Rivadeneira, M.M., Fernández, M. & Navarrete, S.A. (2002) Latitudinal trends of species diversity in rocky intertidal herbivore assemblages: Spatial scale and the relationship between local and regional species richness. *Marine Ecology Progress Series*, **245**, 123–131.

Shmida, A. & Wilson, M. V. (1985) Biological determinants of species diversity. *Journal of Biogeography*, **12**, 1–20.

Shurin, J.B., Havel, J.E., Leibold, M.A. & Pinel-Alloul, B. (2000) Local and regional zooplankton species richness: A scale-independent test for saturation. *Ecology*, **81**, 3062–3073.

Soininen, J., Heino, J., Kokocinski, M. & Muotka, T. (2009) Local-regional diversity relationship varies with spatial scale in lotic diatoms. *Journal of Biogeography*, **36**, 720–727.

Soininen, J., Jamoneau, A., Rosebery, J. & Passy, S.I. (2016) Global patterns and drivers of species and trait composition in diatoms. *Global Ecology and Biogeography*, **25**, 940–950.

Szava-Kovats, R.C., Ronk, A. & Partel, M. (2013) Pattern without bias: Local-regional richness relationship revisited. *Ecology*, **94**, 1986–1992.

Terborgh, J.W. & Faaborg, J. (1980) Saturation of bird communities. *The American Naturalist*, **116**, 178–195.

Vilmi, A., Karjalainen, S.M., Hellsten, S. & Heino, J. (2016) Bioassessment in a metacommunity context: Are diatom communities structured solely by species sorting? *Ecological Indicators*, **62**, 86–94.

Weigand, A., Abrahamczyk, S., Aubin, I., Bita-Nicolae, C., Bruelheide, H., I. Carvajal-Hernández, C., Cicuzza, D., Nascimento da Costa, L.E., Csiky, J., Dengler, J., Gasper, A.L. de, Guerin, G.R., Haider, S., Hernández-Rojas, A., Jandt, U., Reyes-Chávez, J., Karger, D.N., Khine, P.K., Kluge, J., Krömer, T., Lehnert, M., Lenoir, J., Moulatlet, G.M., Aros-Mualin, D., Noben, S., Olivares, I., G. Quintanilla, L., Reich, P.B., Salazar, L., Silva-Mijangos, L., Tuomisto, H., Weigelt, P., Zuquim, G., Kreft, H. & Kessler, M. (2020) Global fern and lycophyte richness explained: How regional and local factors shape plot richness. *Journal of Biogeography*, **47**, 59–71.

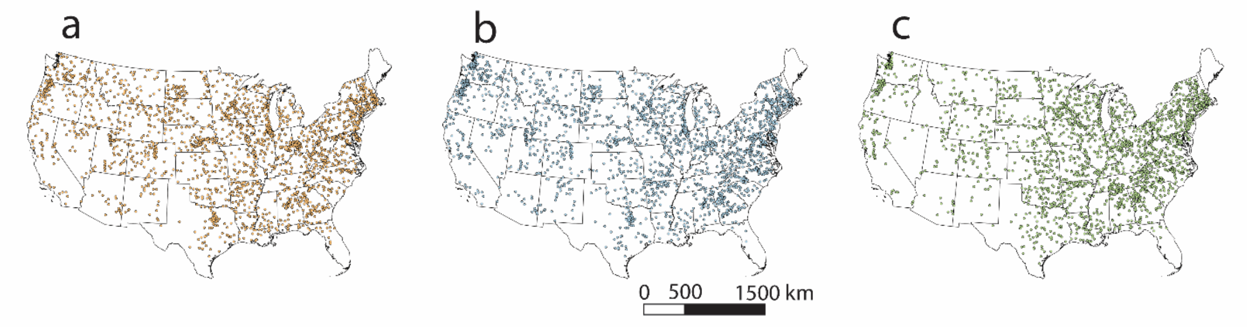

Fig. 4.1. Location of diatom (a), insect (b) and fish (c) stream samples. (Projection: EPSG: 2163)

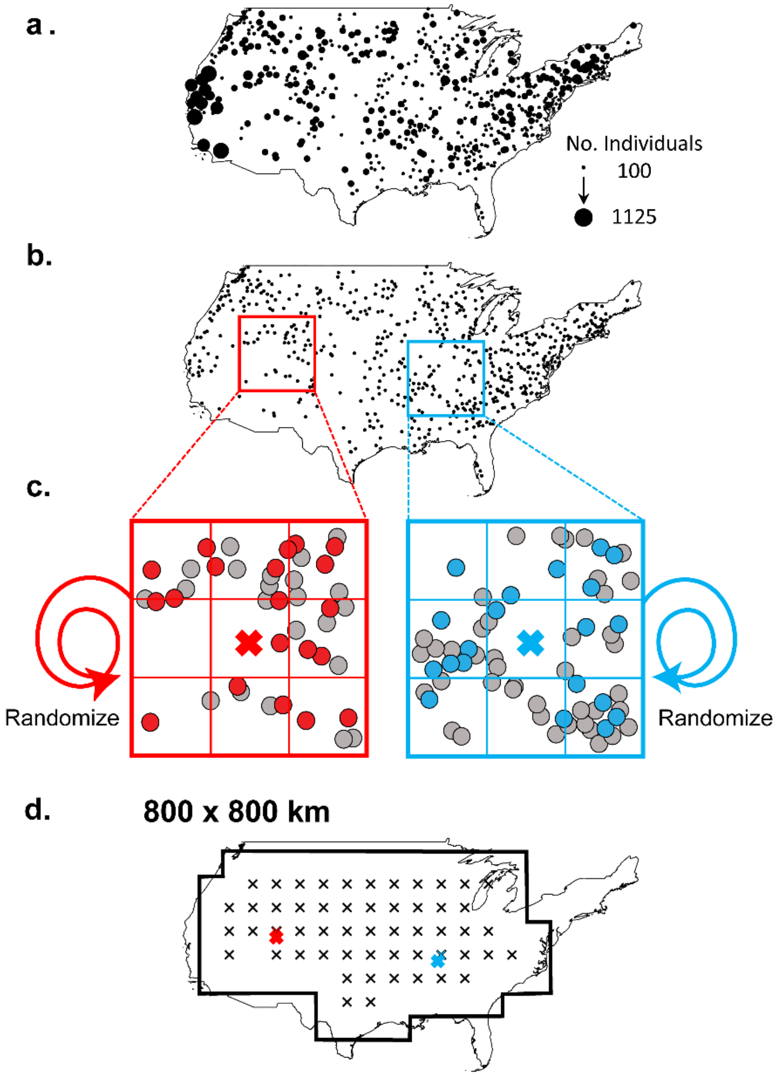

Fig. 4.2. Cartoon describing our methods. Stream samples are first sampled down from their original abundance, represented by point size in (a), to a constant abundance (b). Windows were then drawn, in this case with 800 km sides, and subdivided into 9 sub-windows. Sixty sites were randomly sampled from each window without replacement, and windows were only considered valid if a minimum of 7/9 sub-windows contained samples (c). The number of times a window was randomly resampled is proportional to the number of sites within, in this case the blue window has more sites than the red window, and was sampled proportionately more. (d) Origins of the windows sized 800 x 800 km with blue and red origins corresponding to the windows in c.

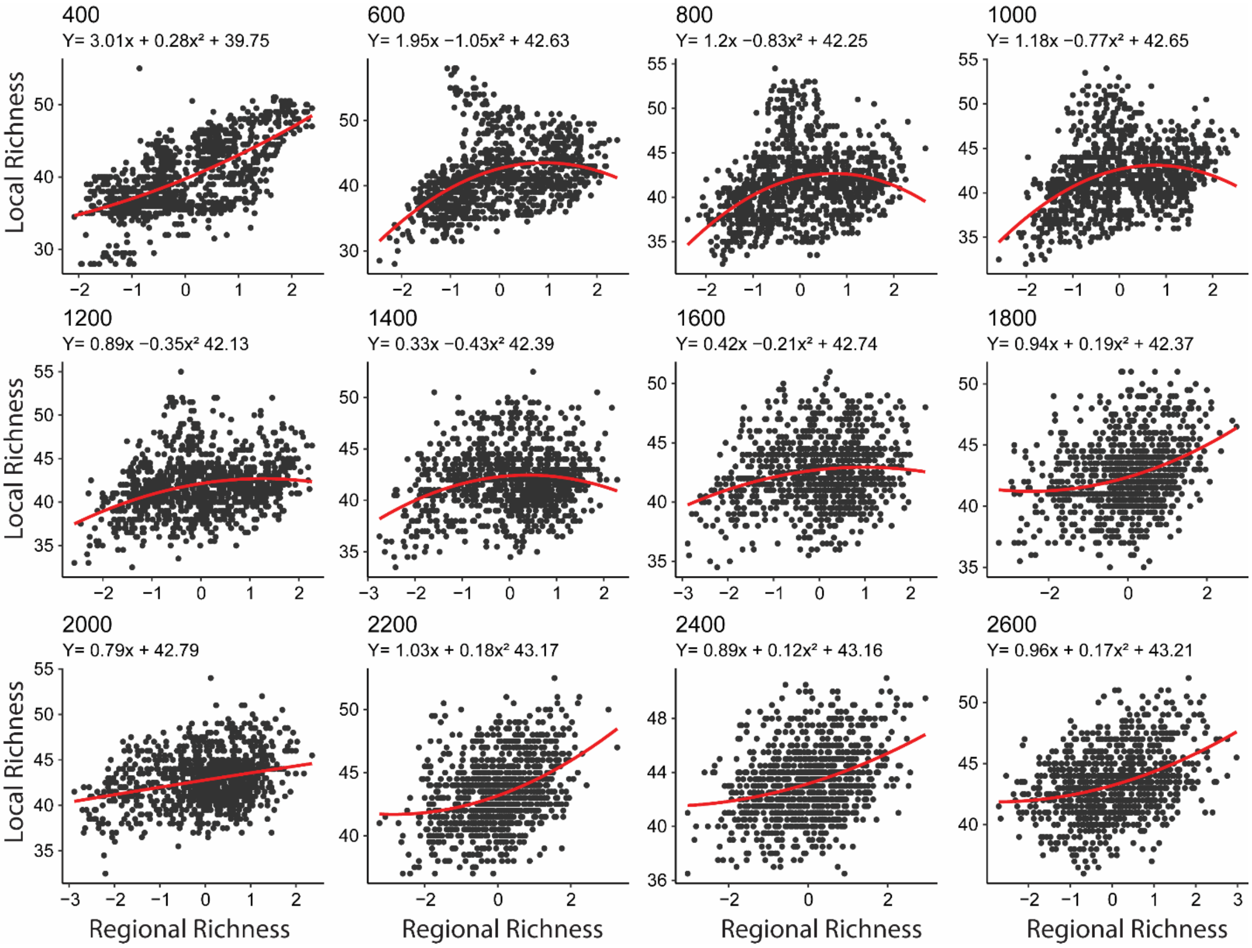

Fig. 4.3. Regressions of local and regional richness of diatoms across our 12 scales. Significant regression parameters areincluded in the equations, and significant trends shown with the red lines.

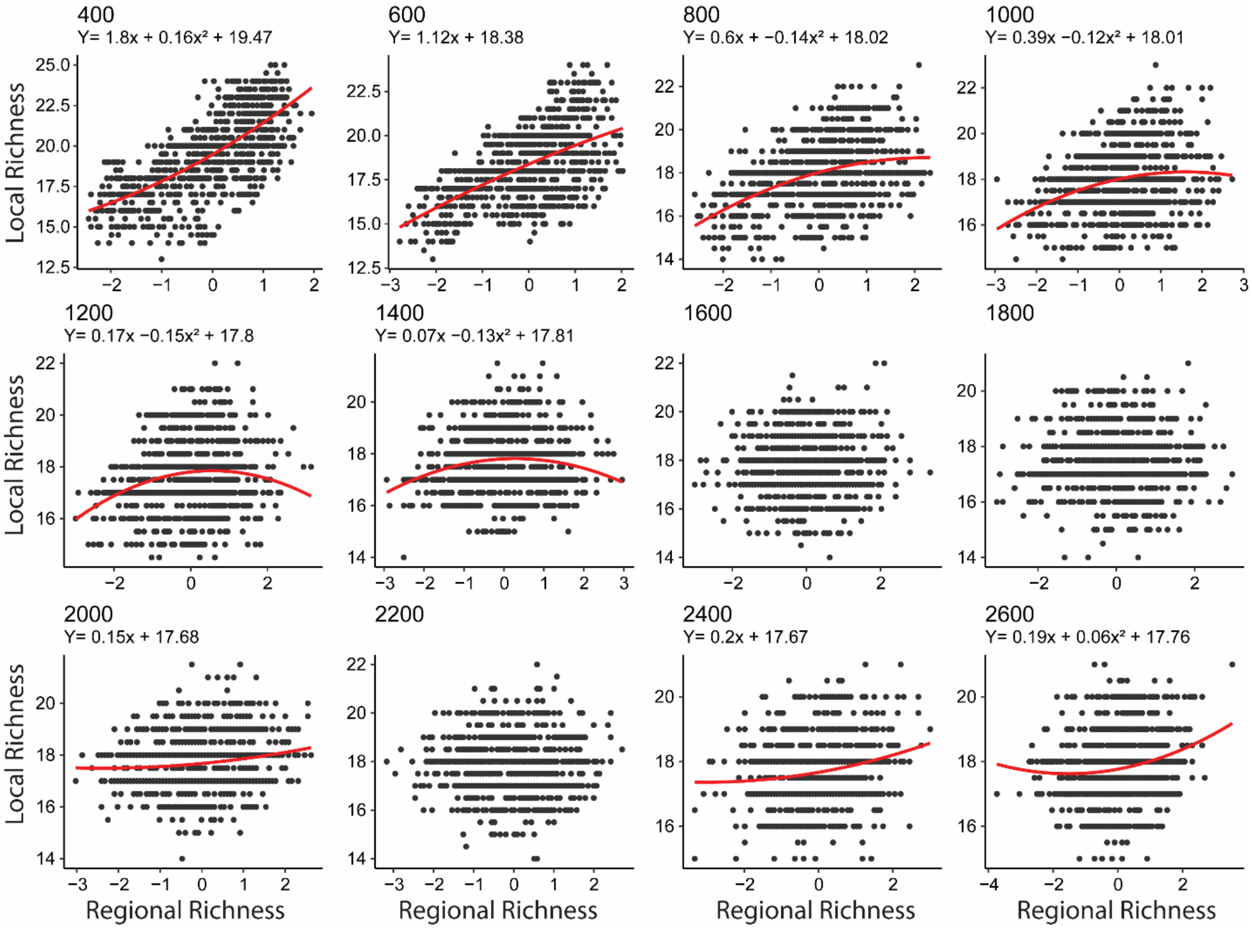

Fig. 4.4. Regressions of local and regional richness of insects across our 12 scales.

Significant regression parameters areincluded in the equations, and significant trends shown with the red lines.

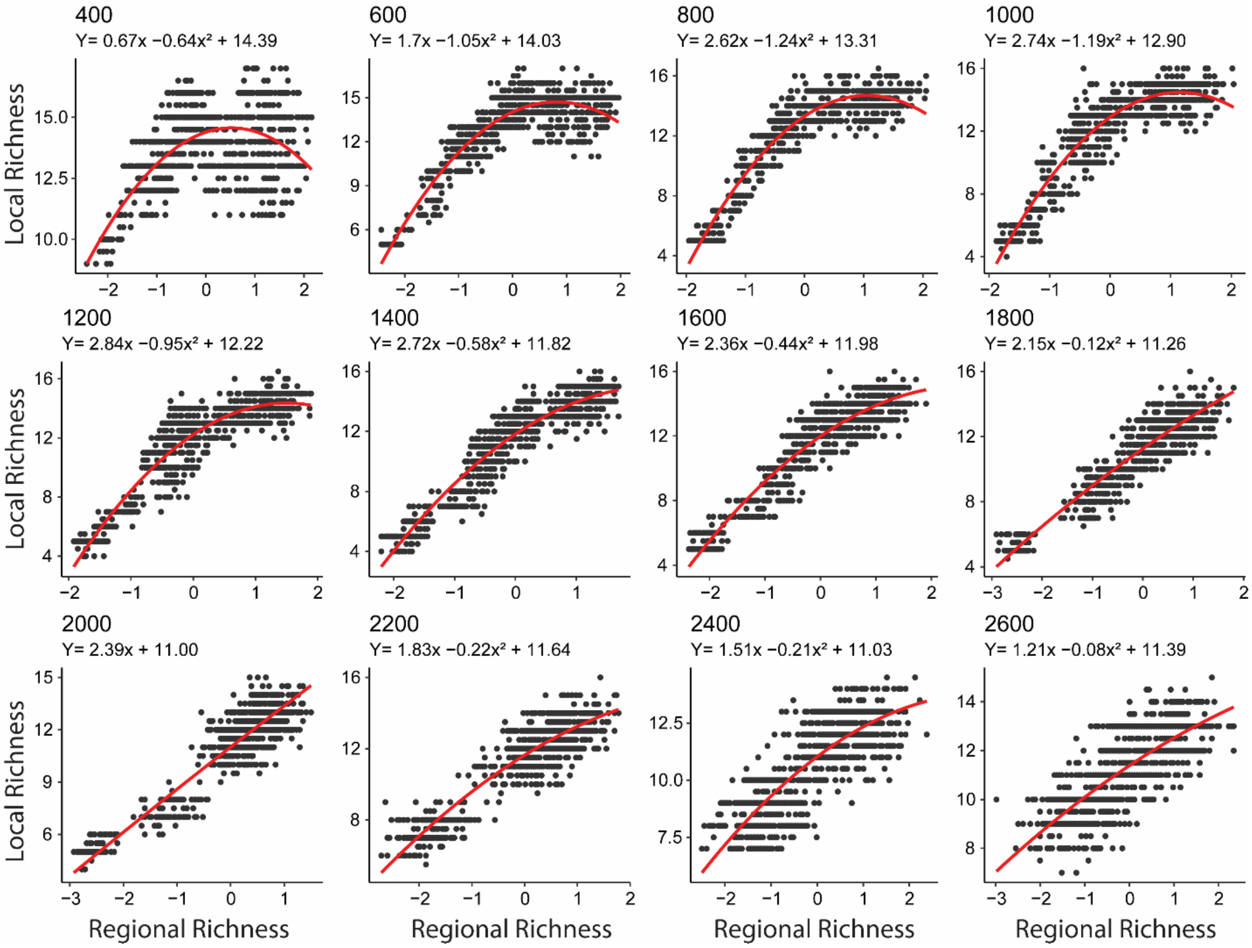

Fig. 4.5. Regressions of local and regional richness of fish across our 12 scales. Significant regression parameters are shown.

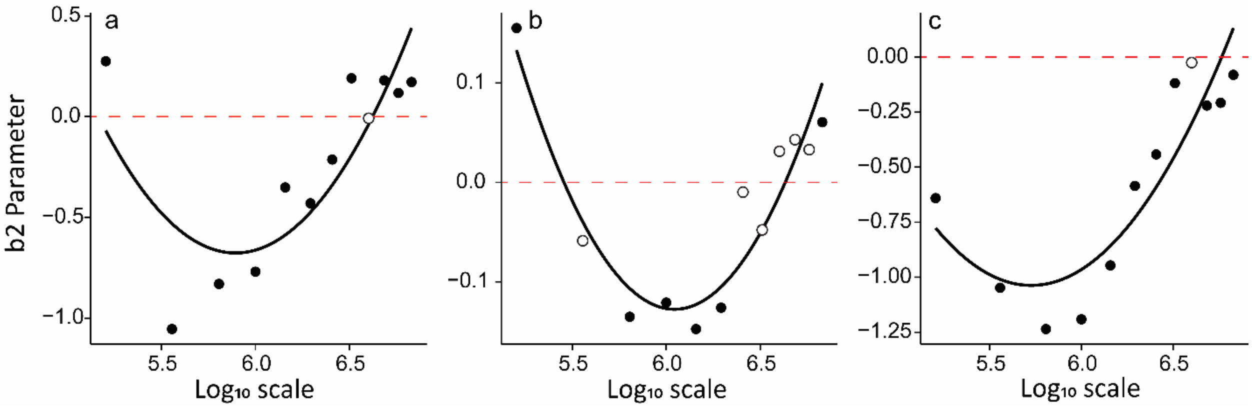

Fig. 4.6. Parameter b_2 from quadratic regression of local richness against regional richness regressed against the $\log_{10}$-transformed scale for diatoms (a), insects (b) and fish (c). Non-significant b_2 parameters are shown with open symbols but were still considered in this regression as they indicate an unsaturated trend.

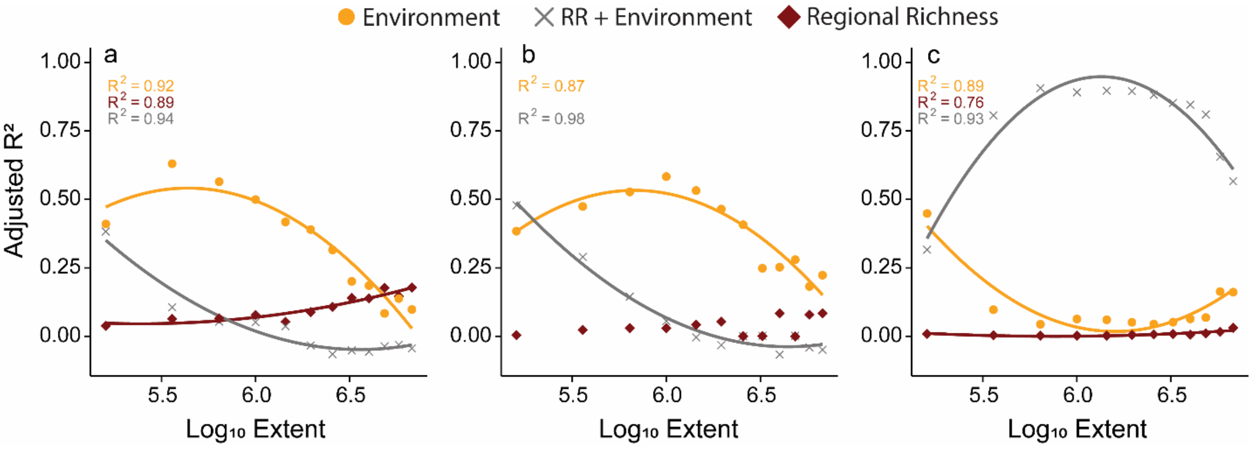

Fig 4.7. Explained variance in local richness (adjusted R^2) by environment, regional richness, and their covariance along the $\log_{10}$-transformed scale for diatoms (a), insects (b) and fish (c). Significant relationships with scale are shown with trendlines and their respective adjusted R^2 provided.

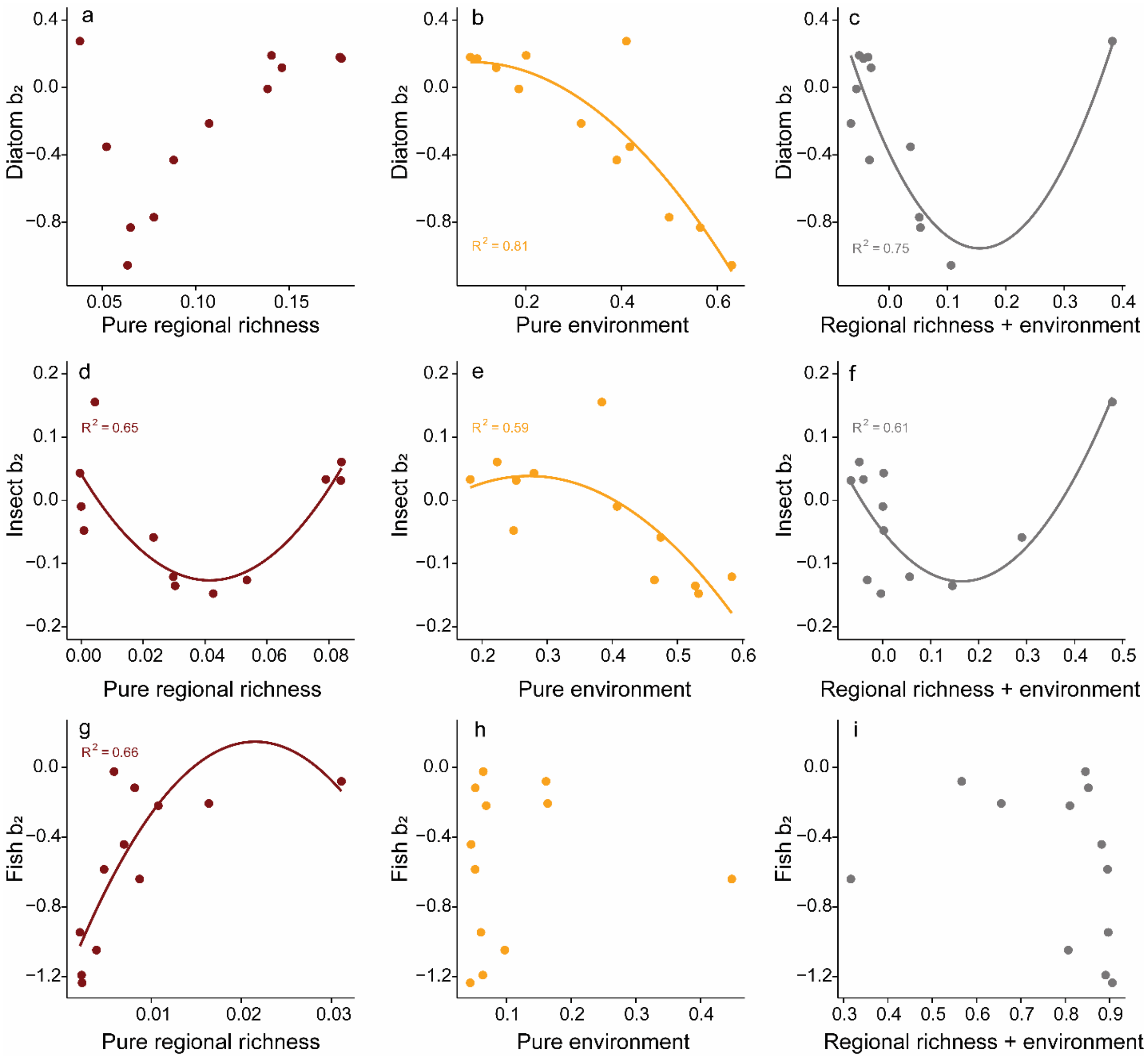

Fig. 4.8. Regressions of the b_2 parameter against the variance explained regional richness, environment, and their covariance for diatoms (a, d, g), insects (b, e, h) and fish (c, f, i). Significant relationships are shown with trendlines and their respective is adjusted R^2 provided

Chapter 5

General Conclusion

The overarching goal my dissertation was to explore how niche and dispersal processes control biodiversity, co-occurrence, and taxonomic relatedness of stream metacommunities in the US. I analyzed the effect of gradients in environmental stress, spatial scale, and species traits finding that the relative importance of niche versus dispersal processes was highly context dependent on latitude, scale, and species group.

In chapter 2, we analyzed latitudinal trends in the relative importance of niche and dispersal processes structuring α- and β-diversity of stream algae, insects, and fish. Latitude represents a gradient of climatic stress, with harsher and more seasonal climates at high latitudes, but more benign and stable climates at low latitudes, which lead to specialization and reduced range size. Therefore, we predicted that the role of niche processes would increase with latitude, while the role of dispersal limitation would decrease. We analyzed α-diversity with multiple regression and variance partitioning, and β-diversity with distance-based RDA and variance partitioning to assess the proportion of variance explained by the environment, space, and their covariance. We found that the importance of niche and dispersal processes varied predictably along the latitudinal gradient only for α-diversity but not β-diversity. However, the niche effects were driven mostly by physicochemistry, not climate, and dispersal effects were not always linked with range size. However, niche effects were generally stronger than dispersal effects on both α- and β-diversity. This suggests that climate-based biodiversity theories have limited relevance for the streams in our study, as non-climatic environmental gradients exerted strong filtering on both local diversity and species turnover.

In chapter 3, we investigated the drivers of co-occurrence network topology vs. metacommunity composition, predicting that niche processes would have a stronger effect on small and more dispersive taxa, while dispersal processes would more strongly constrain large and less dispersive organisms. To address this question, we examined how environmental and spatial variables, and their combination, affected network size, connectance and modularity. Controlling for each type of variables generated networks that were smaller (i.e., fewer nodes and edges), more sparsely connected (i.e., lower connectance and lower average degree), and more modular. We found that while metacommunity composition was primarily structured by the covariance of environment and space, the pure effects were dependent on dispersal ability, as diatoms and insects had a greater contribution of pure environment, while fish had greater pure space. Network topology, however, showed parity between environment, space, and environment + space controls. Finally, we tested for evidence of phylogenetic niche conservatism in networks, predicting that habitat filtering and dispersal limitation would cause co-occurring taxa to be more closely related than would be expected by chance, due to phylogenetic niche conservation. Only fish showed evidence of niche conservatism, while insects showed evidence of phylogenetic niche dispersion, and diatoms displayed no difference in taxonomic similarity resulting from network controls, owing to their high dispersal capability, allowing diatoms to co-occur broadly, irrespective of trait similarity.

In chapter 4, we investigated the drivers of local richness across spatial scale and species group by analyzing the RSR-LSR relationship. We predicted that dominance of local environmental filtering at small scales would result in a saturating RSR-LSR relationship, while dominance of dispersal limitation from the regional species pool would result in a linear relationship at large scales. We found a positive relationship between the effect of pure regional richness and scale in diatoms and fish, and a negative relationship between pure environment and scale in diatoms and insects, aligning with our predictions. However, we found pure regional richness to explain the least amount of variation in local richness in nearly all cases, indicating that the environment, either directly or by influencing regional richness patterns, was dominant in these stream datasets. We confirmed our prediction concerning the shape of the RSR-LSR relationship as all three groups were linear, or trended towards linearity in the case of fish, at large scales. Our final hypothesis, which predicted that the relatively dispersal limited fish would transition from saturated to unsaturated at smaller scales than diatoms or insects was rejected as the smallest windows of diatoms and insects, were in fact linear. This was explained in part by strong mass effects operating these highly dispersive groups.

We presented three chapters of original research describing the relative importance of niche and dispersal processes on metacommunity composition along a latitudinal gradient, on co-occurrence network topology, and across spatial scales through the RSR-LSR relationship. These works deepen our knowledge of how these classic macroecological processes interact to constrain freshwater stream biodiversity, which is critical to making informed conservation decisions.