Integrating the Top-down and Bottom-up Controls of Community Assembly

Gabriel Khattar

A Thesis in the Department of

Biology

Presented in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy (Biology)

Concordia University

Montreal, Quebec, Canada

October, 2023

CONCORDIA UNIVERSITY

SCHOOL OF GRADUATE STUDIES

This is to certify that the thesis prepared

By: Gabriel Khattar

Entitled: Integrating the Top-down and Bottom-up Controls of Community Assembly

and submitted in partial fulfillment of the requirements for the degree of

Doctor in Philosophy (Biology)

complies with the regulations of the University and meets the accepted standards with respect to originality and quality.

Signed by the final examining committee:

_____ Chair

Dr. Yves Gélinas

_____ External Examiner

Dr. Laura Melissa Guzman

_____ Arms-length Examiner

Dr. Rassim Khelifa

_____ Examiner

Dr. Eric Pedersen

_____ Examiner

Dr. Jean-Philippe Lessard

_____ Thesis Supervisor

Dr. Pedro R. Peres-Neto

Approved by

_____

Dr. Robert Weladji, Graduate Program Director

December 08, 2023

_____

Dr. Pascale Sicotte, Dean of Faculty Arts and Science

Abstract

Integrating the Top-down and Bottom-up Controls of Community Assembly

Gabriel Khattar, Ph.D.

Concordia University, 2023

Community assembly theory investigates the mechanisms through which species from a broader pool of potential colonizers form local communities at finer spatiotemporal scales. The theory is heuristic because it reduces the large number of possible mechanisms shaping communities into a tractable number of fundamental high-level processes. However, despite its heuristic value, community assembly theory is inherently context-dependent, i.e., its predictions regarding community dynamics are only valid within specific ecological conditions. Thus, a synthetic understanding of community assembly relies on identifying a few influential ecological axes that regulate the context-dependent nature of community dynamics. In this thesis, I set out to investigate community assembly along two latent ecological axes that determine the context of community dynamics. The first represents the top-down control of species pools on the membership of local communities. The second represents the bottom-up control of landscape features on species movement and interactions. By employing process-based simulation models that replicated community assembly across varied landscape structures and species pool compositions, I generated theoretical predictions about the isolated and interactive effects of both forms of control on: (i) spatiotemporal patterns in community composition; (ii) the ecological selection of prevailing life-history strategies observed in (meta)communities; (iii) the relative importance of assembly processes across space and time and throughout large-scale ecological gradients; and (iv) the trajectories of communities (towards differentiation or homogenization) in response to natural or anthropogenic disturbances. I provide empirical validation for these theoretical predictions by investigating the assembly of insect communities in distinct (bio)geographic contexts or by contrasting model predictions with empirical patterns observed in the literature. In parallel, I introduced new analytical frameworks that allowed the testing of the predictions outlined in this thesis and can be used to address other pertinent questions in community ecology. Collectively, the chapters in this thesis derive a mechanistic understanding of causal links between landscape-mediated bottom-up control, species pool-mediated top-down control, and the context-dependent nature of community assembly. Beyond its theoretical significance, this knowledge is crucial for predicting how the impact of human activities on landscapes and species pools can alter the structure, dynamics, and regulation of ecological communities.

Acknowledgements

I wish to express my sincere gratitude to my advisor, Dr. Peres-Neto (Pedro), for his invaluable guidance and continuous encouragement throughout the course of this thesis. He created an environment where I could think independently, dive deep into questions, and have the freedom to challenge ideas. Yet, whenever I found myself hitting too many walls, he was there to show me the way. Our meetings were both challenging and enjoyable, filled with a lively exchange of ideas that inspired me so much and were instrumental in the development of this thesis. Thank you for believing in me, for your invaluable insights, and for recognizing my needs and adapting to be the exact advisor I needed you to be. Lastly, I really appreciate all the time you have dedicated to helping me grow academically and your recent support when life threw me some curveballs. I began this Ph.D. with an immense respect for the scientist you are. I finish this Ph.D. with an even greater respect for the person I've come to know you as.

I am deeply grateful to have committee members whose research laid the foundation for this thesis. I'm especially grateful to Dr. Lessard (JP) and Dr. Pedersen (Eric) for their insightful feedback on my research and pedagogical training, and for the time they dedicated to help refine my work.

Special thanks to my lab mates, especially Little P., Paul(o), Ali, Alex, Gabby and Hammed. Thank you for your collaborative spirit, insights, and for being part of this journey with me.

I am eternally indebted to my dad, mom, and brothers for their endless love, patience, and encouragement. I also want to thank Angela for being my rock, reminding me who I am, and taking this leap with me.

Lastly, I want to express my gratitude to anyone who directly or indirectly contributed positively to this research and those who believed in me even when I doubted myself.

Dedication

In memory of Joseph Khattar, Vilma Bezerra Khattar, and Rozalina de Almeida Nogueira, who saw me starting this journey but could not see its completion.

Contribution of Authors

As primary author, I led the conception, planning, data collection/extraction/simulation, data analyses, and writing of all chapters in this thesis. Similarly, Dr. Peres-Neto contributed immensely to the conception, planning, interpretation of results, editing, and reviewing of all chapters.

For Chapter 2, Dr. Macedo and Dr. Monteiro assisted in the collection of data, identification of specimens, the curation of biological material, provided field support, contributed critically to the initial drafts and gave final approval for the final version and publication of the manuscript.

For Chapter 4, Dr. Savary contributed to the conception, planning, interpretation of results. They also contributed critically to initial and final drafts an gave final approval for the submission of this manuscript.

Chapter 2 was published online in August of 2021, in Ecography. Authors were credited following the order: Gabriel Khattar, Margarete Macedo, Ricardo Monteiro, and Pedro R. Peres-Neto. https://doi.org/10.1111/ecog.05726

Chapter 3 is under review in The American Naturalist as: Gabriel Khattar & Pedro R. Peres-Neto. The geography of metacommunities: landscape characteristics drive geographic variation in the assembly process through selecting species pool attributes. A preprint of this chapter has been deposited in BioRxiv. https://doi.org/10.1101/2023.06.20.545715

Chapter 4 is under review in Philosophical Transactions of the Royal Society B as: Gabriel Khattar, Paul Savary, and Pedro R. Peres-Neto. Ecological selection of dispersal strategies in metacommunities: impact of landscape features and competitive dynamics. A preprint of this chapter has been deposited in BioRxiv. https://doi.org/10.1101/2023.09.08.556897

Chapter 5 is in preparation to be submitted to Oikos as: Gabriel Khattar & Pedro R. Peres-Neto. Uncovering the trajectories of metacommunities: insights gained from the keystone community concept

Table of Contents

List of Figures

ability) and landscape features influence the attributes of local patches harbouring keystone communities. Keystone communities are those whose extirpation causes cascading secondary effects on extinction and colonization patterns, ultimately driving significant temporal changes in the spatial β-diversity (middle panel).

Figure 2.1: A summary of our conceptual model and quantitative framework. The heatmap in the left represents the spatiotemporal structure of environment in the study system (i.e., scores of elevation-by-month samples in the first principal component of the correlation matrix of climate variables). It is highly stratified elevationaly but not seasonally, which is typical for mountain systems (i.e., high vertical but low horizontal color variation in the heatmap). The panel in the middle represents four fictional communities varying in time (T1; T2) and space (S1; S2). The resulting Sørensen's dissimilarity matrix (right panel) is then used to partition the total variance of the species-by-site-by-time matrix (total beta-diversity, BD_{total}, Legendre and De Cáceres 2013) into its purely spatial (i.e., BD_{space}, red arrows), purely temporal (i.e., BD_{time}, green arrows), and spatiotemporal (i.e., BD_{SpxT}, purple arrows) components based on the equation at the bottom. Here we consider two types of species within each community (adapted from White et al. 2010). Core species (large font size) are well adapted to local environmental conditions and, consequently, maintain long-term viable local populations even in the absence of immigration. Occasional species (small font size) are not well adapted to local environmental conditions and, consequently, their local occurrence depends on the random immigration of individuals from neighboring communities. In this example, BD_{space} is given by the turnover of core species caused by the high stratification of environmental conditions in space (i.e., deterministic species-environment sorting mechanisms). The low stratification of environmental conditions in time allows the maintenance of the local population of core species. As such, BD_{time} represents the turnover of occasional species driven by stochastic events of colonization and local extinctions that cannot be distinguished from the random allocation of individuals from the regional species pool into local communities (i.e., stochastic sampling effects). BD_{SpxT} is then the outcome of the deterministic turnover of core species in space and stochastic turnover of occasional species in time.

Figure 2.2: Panel A summarizes (i.e., bar heights = mean, whiskers = Standard error of the mean) the average contribution of spatial (BD_{space}), temporal (BD_{time}) and spatiotemporal (BD_{SpxT}) beta-diversity to BD_{total} observed across the 10 taxonomic groups (seen in Panel B). Across taxa, the average contribution of BD_{time} to BD_{total} is significantly lower than the contributions of BD_{space} and BD_{SpxT}. Additionally, BD_{time} is significantly more nested than BD_{space} and BD_{SpxT}. Panel B, first row (top): Phengodidae (glow-worms), Lampyridae (fireflies), Carabidae (ground beetles). Second row: Eumolpinae (Leaf beetles), Antrhibidae (fungus weevils), and Cerambycidae (Longhorn beetles). Third row: Metopiinae, Pimplinae, Meteorus. Fourth row: Mesosotoinae.

Figure 2.3: Considering all taxonomic groups combined (n = 10) , dissimilarities in temporal beta-diversity (BD_{time}) are, on average, not different from null model expectation (mean standardized effect size -mean SES_{ij}- overlaps with dashed line in 0). Conversely, spatial (BD_{space}) and spatiotemporal (BD_{SpxT}) beta-diversities were higher than the null expectation (i.e., mean $SES_{ij} > 0$ in both cases) and significantly higher than in BD_{time}. A jitter effect was used to place taxonomic groups according to their mean SES_{ij} values across dimensions.

local abundances. Communities were sorted into deciles of keystoneness (i.e., Q1 = the 3 communities with the lowest levels of keystoneness, Q10 = the 3 communities with the highest keystoneness). Losses and gains on local species abundances over time sum up to total Δ Spatial β. Points represent average effect sizes across 50 simulation replicates per combination of levels of niche breadths and dispersal rates in the simulated regional pool. Whiskers represent 95% confidence intervals. Dashed line represent Δ Spatial $\beta = 0$.

Figure 5.4: Mapping the keystoneness of moths communities in space (across elevations) and time. High-elevation communities consistently exhibited a greater degree of keystoneness compared to low-elevation communities.

Figure 5.5 Results of random forest modeling considering local diversity (Panel A, $R^2 = 0.58$) and keystoneness (Panel B, $R^2 = 0.61$) as a function of climate, spatial MEMs, temporal AEMs, and light pollution (ALAN, in yellow) at different scales, spatial and temporal variables of the empirically studied moth metacommunity. Only predictors retained via model selection are reported here. The first column shows variables ranked in order of importance to model fit (estimated as % of the increase in mean squared errors when permuted). The second column shows their standardized effect sizes. Dashed line represents effect size = 0, and whiskers indicate a confidence interval at 95%. The third column reports partial dependence plots. They serve as a graphical depiction of how the average predicted values of the response (local diversity and keystoneness) change with variation in the level of ALAN at smaller (1km buffer) and larger (5km buffer) scales. The shaded area encompasses +- 1 Standard error of the mean.

Figure SI. 2.I Virtual landscape in which metacommunity models were set to run (only last 12 time steps are shown). The spatiotemporal structure of environmental conditions is similar to the one observed in the tropical slope where insect communities were sampled (see Fig. 2.1 in the main text)

Figure SI. 2.II: Average contribution of spatial (BD$_{space}$), temporal (BD$_{time}$), and spatiotemporal (BD$_{SpxT}$) changes in community composition to the total variance of species-by-site-by-time matrix (BD$_{total}$). In the first panel analyses were carried out considering both types of species across communities (Full matrix) and results were qualitatively similar to the results observed across insect communities sampled in this study (see Fig. 2.2 main text). The other two panels show the results when only occasional (middle) or core species (right) were kept within communities. As we can see, the composition of core species varies relatively little over time, while the composition of occasional species varies in all dimensions.

Figure SI. 2.III: Mean Standardized effect sizes (SES) observed across dimensions calculated considering full communities and after the removal of either core or occasional species (see caption of Figure SI 2. II). Patterns observed in the full matrix were qualitatively similar to the patterns observed in analyses considering real communities (see Fig. 2.3). When considering only occasional species (middle), there is a higher similarity between the null expectation and the final simulated matrix (SES closer to zero). This result indicates that stochastic events of colonization extinction are the main drivers of beta-diversity when only occasional species are taken into consideration in our analyses. Conversely, beta-diversity was very different from the null

expectation (i.e., above 0 in BD_{space} and BD_{SpxT}, but below zero 0 in BD_{time}), which indicates that other mechanisms drive the spatiotemporal distribution of this type of species across communities

We make use of simple simulations to demonstrate that averaging effectively controls for differences among BD_{space}, BD_{time}, and BD_{SpxT} that result solely from unbalances in the dissimilarity matrix **D**. We started by simulating 150 regional pools using three different species abundance distributions (Log-normal, Poisson-lognormal, and negative binomial distributions, 50 regional pools each). Each regional pool was composed of 1000 individuals distributed across 50 species. Each individual within any given pool was then randomly allocated across 90 samples with specific spatial and temporal coordinates. Then we calculated the pairwise dissimilarity among simulated communities and obtained the matrix **D_{sim}**. Note that the number of entries in **D_{sim}** representing D_{time}, D_{space}, and D_{SpXT} was set to be, respectively, 225, 630, and 3150; the same unbalance observed in our real **D** matrices. Since the composition of simulated communities is generated by chance alone, all dimensions should equally contribute to BD_{total}. However, without proper correction, we observe that $BD_{SpxT}> BD_{space}> BD_{time}$; a pattern that is caused by the observed unbalance in **D_{sim}** (Figure SI. 2. V). Only after dividing BD_{space}, BD_{time}, and BD_{SpxT} by the number of entries in **D_{sim}** representing D_{space}, D_{time}, and D_{SpxT}, we observe that all dimensions, on average, contribute equally to BD_{total} (Figure SI. 2. VI). Based on the same simulations, we also show that the results of the SES procedure (see main text) were not influenced by unbalances in **D.** More specifically, in these simulations where species were randomly distributed across communities, the mean SES_{ij} for each dimension is not significantly different from 0 (Figure SI. 2.V).

above which the decreases in Multivariate pseudo SE with additional samples are not significantly

List of Tables

Chapter 1 General Introduction

1.1 Community ecology may be lawless, but it is certainly not unprincipled

> *"[...] without a historical appreciation for the development of ecological ideas, ecologists can neither easily relate theory to reality nor detect the recycling of historical debates and issues."*
> (Graham and Dayton 2002)

The realization that mature disciplines are marked by the possession of general laws has led many ecologists to pursue distinctive ecological laws that would "carve nature at its joints" and solidify ecology's status among the foremost natural sciences (O'Hara 2005; Roughgarden 2009; Justus 2021). This motivation was particularly strong in the second half of the 20th century and is outlined in the opening remarks of Robert MacArthur's influential book *"Geographical Ecology"* (1972):

"To do science is to search for repeated patterns, not simply to accumulate facts [...] But not all naturalists want to do science; many take refuge in nature's complexity as a justification to oppose any search for patterns [...]".

MacArthur's rallying cry, which reads as a manifesto against phenomenological and descriptive studies in community ecology, echoed among influential ecologists at the time, including E. O. Wilson, Richard Levins, and Richard Lewontin[1]. However, the pursuit of distinctive and universal ecological laws manifested as repeated patterns in nature faced challenges when it became evident that different underlying mechanisms could produce analogous patterns in community size and composition (i.e., the "many-to-one" problem, *sensu* Levins and Lewontin 1980). Moreover, a large body of empirical evidence demonstrated that ecological relationships initially considered as universal (e.g., diversity-area relationships, latitudinal clines on species diversity, etc.) are marked by striking exceptions across taxa and biogeographic regions (e.g., Wardle et al. 1997; Cerezer et al. 2022). Consequently, the literature between the end of the 20th and the beginning of the 21st centuries is marked by a strong skepticism about the scientific value of community ecology due to its elusive study subjects (Ricklefs 2008) and its theories that are filled with "messy details" and devoid of predictive power (Mcintosh 1987; Lawton 1999).

Simultaneously, two distinct counterarguments gained momentum in the literature. The first acknowledges that while community ecology may lack general laws, it asserts its relevance by emphasizing that the effectiveness of conservation approaches significantly diminishes when the particularities of focal systems are deliberately disregarded (Simberloff 2004). The second

---

[1] In his autobiography, Edward O. Wilson [1994, Chapter 13] names this group of ecologists the "Marlboro Circle", referring to MacArthur's lakeside home at Marlboro, Vermont, where they met to discuss their research agendas and ongoing projects.

acknowledges that community ecology may be lawless but is certainly not devoid of principles (O'Hara 2005). That is, while community ecology may not possess general laws, it can be made predictive when approached from a "first principles" perspective, which breaks down the myriad of mechanisms driving community dynamics into a bearable number of fundamental high-level processes validated both theoretically and empirically (Roughgarden 2009; Vellend 2010; Marquet et al. 2014). This understanding sparked a paradigm shift in community ecology: A synthetic and general theory for community ecology should not be based on searching for universal correlational patterns among variables but rather on identifying the fundamental processes underlying these patterns.

Numerous studies have aimed to identify the universal processes governing communities and, consequently, contribute to the development of a synthetic theory for community ecology (e.g., Hutchinson 1959; Diamond 1975; Chesson 1985; Belyea and Lancaster 1999; Chase and Myers 2011). Arguably, the most comprehensive attempt proposes that all mechanisms underlying community size and composition can be categorized into four broad fundamental high-level processes: selection, ecological drift, dispersal, and speciation (Vellend 2010, 2018). Selection encompasses all mechanisms in which species characteristics (traits) interact with abiotic and biotic conditions to determine the composition of local communities. It includes species-environment sorting, competitive exclusion, and storage effects. Ecological drift refers to the influence of demographic events (e.g., birth, death, immigration, emigration) that occur at random with respect to species characteristics (including their identities) on governing the structure of communities (Vellend et al. 2014). Examples of mechanisms under ecological drift are regional sampling effects (Kraft et al. 2011) and demographic stochasticity (Shoemaker et al. 2020*a*). Dispersal is the unidirectional movement of individuals from one location to another (Jacobson and Peres-Neto 2010) or from one moment in time to another moment in the future (e.g., dormancy, Buoro and Carlson 2014). Dispersal is a process that results from plastic behavioral decisions involving an individual's departure, movement, and settlement (Clobert et al. 2012). Lastly, speciation operates at large biogeographic spatiotemporal scales and influences the size and composition of communities by its influence on the size and composition of the regional pool of species available for colonization (see more on different definitions of species pools below).

Sorting mechanisms into these four non-mutually exclusive fundamental processes enabled community ecology to move from "either/or" debates that have been recycled in many instances through its history (Levins and Lewontin 1980) to a consensus that these processes collectively shape the assembly of ecological communities, each with its unique degree of influence. The envelope of theories, models, and concepts in community ecology that can be organized around these four fundamental processes has been named "The theory of ecological communities" (Vellend 2018). However, here, I will refer to it as " community assembly theory" to emphasize its primary focus on inferring the relative importance of processes through which species from a broader pool of potential colonizers form horizontal communities (i.e., communities consisting of a single trophic level, but see Guzman et al. 2019) at finer scales (Weiher et al. 2011). I argue that

this terminology better delineates the body of theories and models in community ecology that fall within and out of the scope of this thesis. Notable examples of models and theories explicitly and implicitly encompassed in this thesis include modern coexistence theory (Chesson 1985; Barabás et al. 2018), metacommunity theory (Leibold and Chase 2018), and island-biogeography theory (Macarthur and Wilson 1967), among others. Conversely, pertinent theories in community ecology that lie outside the boundaries of our definition of "community assembly theory" will not be covered in this thesis. Examples include metabolic theory (Brown et al. 2004), biodiversity-function relationships (reviewed in van der Plas 2019), and ecological stoichiometry theory (e.g., Moe et al. 2005), to name a few.

1.2 Community assembly theory: a contingent yet heuristic framework

"Since all models are wrong, the scientist cannot obtain a 'correct' one by excessive elaboration. On the contrary, following William of Occam, he should seek an economical description of natural phenomena." (Box 1976)

Community assembly theory is inherently contingent, meaning that its predictions regarding community assembly are valid only within specific ecological conditions (Lawton 1999; Catford et al. 2021; Kolasa et al. 2021). Predictions stemming from the contingent framework of community assembly theory can be expressed through the following parsimonious verbal model:

"If conditions A and B hold, then the relative importance of process X in shaping community structure is Y."

In this context, *A* and *B* represent the states of study-specific ecological conditions and predictors (e.g., environmental types), while *X* is the assembly process under investigation, and *Y* corresponds to the estimated relative importance of this assembly process derived through multivariate models and variation partitioning, for example.

Despite its contingent nature, community assembly theory retains its heuristic value by organizing the multitude of mechanisms that can be invoked to represent *X* (e.g., competition, species-environment sorting, demographic stochasticity, sampling effects) into a manageable set of universal and non-mutually exclusive high-level processes (i.e., selection, dispersal, drift, and speciation). This reduction in the dimensionality of X can foster a common ground upon which community ecologists could reframe system-specific discoveries within a unified framework for studying the structure of ecological communities (Vellend 2018).

By applying a similar rationale, we can enhance the heuristic value of community assembly by proposing a reduced number of general ecological conditions (i.e., the "As" and "Bs" in the verbal model above) that dictate the relative importance of different assembly processes. Candidates to represent such conditions must possess a degree of generality that allows for their

recognition across studies investigating different biotic systems and should be readily adaptable within the framework of community assembly theory. Based on these criteria, three noteworthy potential indicators of such ecological conditions emerge in the literature: (i) the scale (spatial and temporal extent and resolution) at which community assembly is studied and its processes are measured (Levin 1992; Chave 2013); (ii) the characteristics of the landscapes where the assembly process takes place (Peres-Neto et al. 2012; Bar-Massada et al. 2014; Fournier et al. 2017, 2020); and (iii) the characteristics of the species pool from which ecological communities are assembled (Taylor et al. 1990; Lessard et al. 2012*a*; Zobel 2016).

The spatial and temporal extent and resolution at which one studies ecological communities ultimately determines one's capacity to detect the signal of assembly processes on community composition data. This is because it defines the variation in the dependent and independent variables considered in ecological analyses (Levin 1992). For instance, at broader spatial extensions, sampling encompasses greater environmental heterogeneity and, consequently, increases the signal of selection through species-environment on community assembly (Viana and Chase 2019). Conversely, at smaller scales, less habitat heterogeneity is encompassed by sampling and, consequently, the contribution of stochastic events (ecological drift) in underlying community variation increases (Chase 2014). The critical role of scale in modulating our inferences about community assembly has been well recognized and extensively studied for decades (e.g., Wiens 1989; O'Neill et al. 1996; Viana and Chase 2019). Therefore, for the rest of this thesis, I will not extend our discussion on the role of scale as a moderator (i.e., a factor determining the importance of assembly mechanisms) of community assembly. Instead, I will focus on expanding our understanding of how landscape attributes and characteristics of species pools influence community assembly.

It is important to emphasize that I acknowledge the inherent covariation between scale and species pool attributes (e.g., the species pool size is bound to increase with area) and the existing relationships between scale and landscape characteristics (e.g., average connectivity among habitat patches decays with spatial extent). Consequently, the theoretical and empirical examples used throughout this thesis share the fundamental assumption that the sampling scale appropriately addresses the ecological questions under investigation. That is, the scale of sampling designs here is comparable and sufficient to demonstrate the individual, interactive, and combined effects of species pools and landscape features on community assembly.

1.3 The top-down control of species pools over community assembly

"The decrease in species density with increasing habitat fertility level in temperate regions [...] is a consequence not of a general increase in the intensity of competition, but of a general decrease in the size of the pool of species that are suited or adapted to increasingly specialized (i.e., more fertile) habitat conditions."
(Taylor et al., 1990)

4

Species pools are broadly defined as the species that can potentially colonize a group of habitat patches (Cornell and Harrison 2014). While it is difficult to trace back the theoretical origins of this concept, we see it implied in the work of early plant ecologists investigating the relationship between local and regional richness of desert plants (e.g., Spalding 1909, p. 2-15). Nonetheless, the understanding that characteristics of species pools (e.g., richness, composition, and dominant life history traits of its member species) can explain observed spatiotemporal variation in ecological patterns has been formalized under the "species pool hypothesis" (Taylor et al. 1990; Partel et al. 1996; Zobel 2016). This hypothesis posits that the number of species found locally under different ecological conditions is determined by the number of species adapted to such conditions in the regional pool. For instance, suppose a particular biogeographic region where most species in the regional pool have evolved to perform better under high productivity levels. As such, if a group of communities is distributed along a productivity gradient, ecological selection through species-environment sorting would give rise to a positive relationship between community richness and productivity. Conversely, a negative relationship is expected in biogeographic regions where species have evolved to perform better under low productivity levels. The species pool hypothesis has been invoked to explain why the relationship between richness and ecological gradients (e.g., productivity) change directions in different parts of the globe (Zobel et al. 2011). Following the same rationale, recent developments extended the hypothesis to explain how biogeographic variation in the evolutionary and historical processes that have determined the characteristics of regional pools can underly broad-scale variation in the relative importance of different assembly processes (Lessard et al. 2012*a*).

The species pool hypothesis and its recent analogs imply the existence of a "top-down" control on community assembly[2]. In this case, species pool characteristics (size, dominant life histories, composition, etc.), which are components at a larger scale than local communities, trickle down to influence the importance of assembly processes operating within local communities. In contrast, these within-community processes have little influence on the characteristics of species pools. The existence of a top-down control is at the core of the so-called "Mainland-island models" for community assembly (e.g., MacArthur and Wilson 1967; or Keddy's filter model, (1992)), which represent species pools as being decoupled from local communities. Depicting species pools as being external to local communities is more than a stylish decision; it implies the assumption that species pools are little influenced by within-community mechanisms operating at fine spatiotemporal scales, serving as an external reservoir of species that are shaped by evolutionary and historical mechanisms operating at broad spatiotemporal scales. As such, this "top-down control" perspective links spatial variation in the importance of assembly processes to broad-scale

---

[2] "Top-down control" is commonly defined as the control of consumers on the abundance and distribution of resources. In the context of this thesis, it is defined as the effects of species pools characteristics on the membership of local communities

variation in the evolutionary and historical mechanisms that have uniquely shaped the size and composition of species pool (Kraft et al. 2011; Lessard et al. 2012*a*, 2012*b*; Carstensen et al. 2013).

A large body of empirical and theoretical evidence supports the existence of species pool-mediated top-down control over community assembly (Fukami 2004; Kraft et al. 2011; Lessard et al. 2012*a*; Karger et al. 2015). Theoretical studies observe the existence of such control when they manipulate the characteristics of simulated species pools and observe shifts in diversity patterns (e.g., Thompson et al. 2020) and the importance of different assembly processes (e.g., Gravel et al. 2006; Ovaskainen et al. 2019). For instance, by manipulating the degree of ecological specialization and dispersal propensity of species in the regional pool, one can create distinct community assembly archetypes (e.g., Sokol et al. 2020).

In empirical studies, support for the top-down control perspective becomes evident when one relies on distinct mechanism-based operational definitions of species pools (i.e., the set of species considered in ecological analyses) to test different hypotheses about the mechanisms determining the composition of species assemblages (e.g., Peres-Neto et al. 2001; Lessard et al. 2016; Braga et al. 2023). Different operational definitions of species pools represent distinct hypotheses regarding the processes filtering the pool into a subset of species that can co-occur (not necessarily coexist) across local communities. These distinct hypotheses are then tested through null models to assess how observed community composition deviates from what would be expected in the absence of the specific mechanism under investigation. The outcomes of such null models can be understood as the operationalization of null hypotheses in inferential statistical analyses. Therefore, by contrasting the results of null models against observed data, one can explore how different mechanism-based definitions of species pools elucidate the role of different assembly processes in filtering species from the regional pool into local communities (Peres-Neto et al. 2001; Lessard et al. 2012*b*, 2016).

1.4 The bottom-up control of landscapes over community assembly

> *"The pace of ecological change will be determined by complex processes of ecological succession influenced by landscape position, topography, climate, [...] and further geophysical forces. The current volcanic activity at Mount St Helens attests to its dynamic character [...] landform, and soil legacies of the 1980 eruption will influence ecological processes for centuries to come".*
> (Dale et al. 2005)

The recognition that the relative importance of different assembly processes is contingent upon the characteristics of the landscape (e.g., spatiotemporal variation in physical connectivity, environmental autocorrelation, and environmental heterogeneity) where community dynamics unfold has been formalized in numerous distinct hypotheses (see summary table I in Tscharntke et al. 2012). These hypotheses converge in the prediction that landscape characteristics modulate the costs and risks of species movement, the probability and outcomes of biotic interactions, and the

importance of demographic stochasticity through their influence on effective population sizes (see Tscharntke et al. 2012 and references within). These hypotheses imply that landscape characteristics impose a "bottom-up control" on the relative importance of assembly mechanisms.

Supporting evidence for landscape-mediated bottom-up control in community assembly can be found in empirical studies and theoretical models. For instance, research in floodplain ecosystems has demonstrated that seasonal fluctuation in the area of temporary ponds can significantly alter the rate of dispersal among patches, the frequency of interactions among consumers, and the role of competition for limiting resources (Fernandes et al. 2013, FitzGerald et al. 2017). Moreover, the temporal variation in the relative importance of assembly processes caused by seasonal variations in landscape topology and environmental conditions is commonly reported in the literature (Tonkin et al. 2017; Holyoak et al. 2020; Li et al. 2023). Theoretical studies have also illustrated the existence of a landscape-mediated bottom-up control on community assembly. In these studies, the heterogeneity, spatial distribution of environmental conditions, and connectivity of landscapes are commonly manipulated as a way to investigate how they affect coexistence patterns, the dominance of dispersal strategies, and the influence of demographic stochasticity on community dynamics (e.g., Büchi et al. 2009; Büchi and Vuilleumier 2012; Fournier et al. 2016; Marco Palamara et al. 2023).

1.5 Integrating both types of control to better understand community assembly

> *""[…] Future research should examine how landscape composition and configuration affect fragment community dynamics and species pools, which components of the species pools are then locally represented in different habitat types, and how the composition and configuration of habitat types in turn can feed back to determine the regional species pool."* (Tscharntke et al. 2012)

One might be inclined to view top-down and bottom-up controls of community assembly as mutually exclusive forms of regulation and question which type of control holds greater relevance in influencing community dynamics. Indeed, it is not uncommon to encounter such dichotomous viewpoints in the history of community ecology (Levins and Lewontin 1980; Graham and Dayton 2002). However, this thesis goes beyond this either/or debate and seeks to demonstrate that a comprehensive understanding of community assembly necessitates recognizing that both forms of control have a mutual influence on the importance of assembly processes.

To achieve that, I will rely on the conceptual framework of metacommunity ecology, a subbranch of assembly theory that studies the dynamics of local communities interconnected through the dispersal of potentially interacting species. In metacommunity ecology, landscapes transcend the mere juxtaposition of local communities in space; they shape metacommunity dynamics by limiting species movement and governing coexistence across local and regional

scales (Mouquet et al. 2013; Li et al. 2023). Moreover, metacommunity theory revolves around archetypes representing recurring patterns or types of interactions and dynamics within metacommunities, namely species sorting, mass-effects, patch dynamics, and neutral (Leibold et al. 2004; Cottenie 2005; Brown et al. 2017; but see Leibold and Chase 2018). These metacommunity archetypes assume that metacommunity dynamics arise from ecological differentiation, competitive hierarchies, and variation in species' dispersal abilities in the regional pool (Thompson et al. 2020). As a result, the conceptual foundation of metacommunity ecology seamlessly articulates the influence of species pool-mediated top-down and landscape-mediated bottom-up controls of community assembly within a cohesive framework.

Metacommunity ecology was initially conceptualized as a spatially-oriented framework for exploring the underlying causes of spatial variation in species distributions and community composition (Leibold et al. 2004; Cottenie 2005; Winegardner et al. 2012). However, recent developments fostered by empirical and theoretical research demonstrated that studying metacommunities from both spatial and temporal perspectives enhances our ability to estimate the relative importance of different assembly processes (Holyoak et al. 2020; Guzman et al. 2022). Embracing an integrative spatiotemporal approach to the study of metacommunities facilitates the development of conceptual models that incorporate top-down and bottom-up community assembly controls in at least two different ways (see Figure 1.1 for an illustrative conceptual representation). Firstly, integrative spatiotemporal approaches consider the spatial and temporal dimensions of species' life histories, their responses to environmental gradients in time and space, and the temporal fluctuations in landscape topology and connectivity. Consequently, integrative spatiotemporal metacommunity models improve inferences about the importance of processes driving community variation (Wisnoski et al. 2019; Record et al. 2021; Guzman et al. 2022; Li et al. 2023). Secondly, unlike traditional mainland-island models where species pools are decoupled from local communities, species pools in spatiotemporal metacommunity models are dynamic at fine spatiotemporal scales due to biotic interactions, habitat selection, and spatial dispersal and temporal dispersal through dormancy (Wisnoski et al. 2019). Consequently, spatiotemporal metacommunity models seamlessly acknowledge the existence of a feedback loop in which dynamics within and between local communities scale up to induce changes in the regional pool forming metacommunities, while the dominant characteristics of species in the metacommunity pool ultimately influence community dynamics (Mittelbach and Schemske 2015; Lamy et al. 2021)[3].

_____

[3] It should be noted that species pools in mainland-island models (described above) and metacommunity models are not mutually exclusive (Fukami 2015); instead, they are nested and connected through spatial (long and short distance) and temporal dispersal (e.g., dormancy, diapause) events (see figure 1.1 for a conceptual illustration). See Chapter 3 for a in-depth discussion about operational definitions of species pools and their implied assumptions.

In this thesis, I draw upon empirical and theoretical spatiotemporal metacommunities to demonstrate the significance of acknowledging the mutual impact of top-down and bottom-up controls on community assembly. More precisely, I aim to expand our understanding of the spatial and temporal variations in metacommunity patterns and the underlying fundamental processes that drive these variations. It is important to note that while the chapters in this thesis can be viewed as independent studies with their own concepts, questions, and objectives, they collectively propose
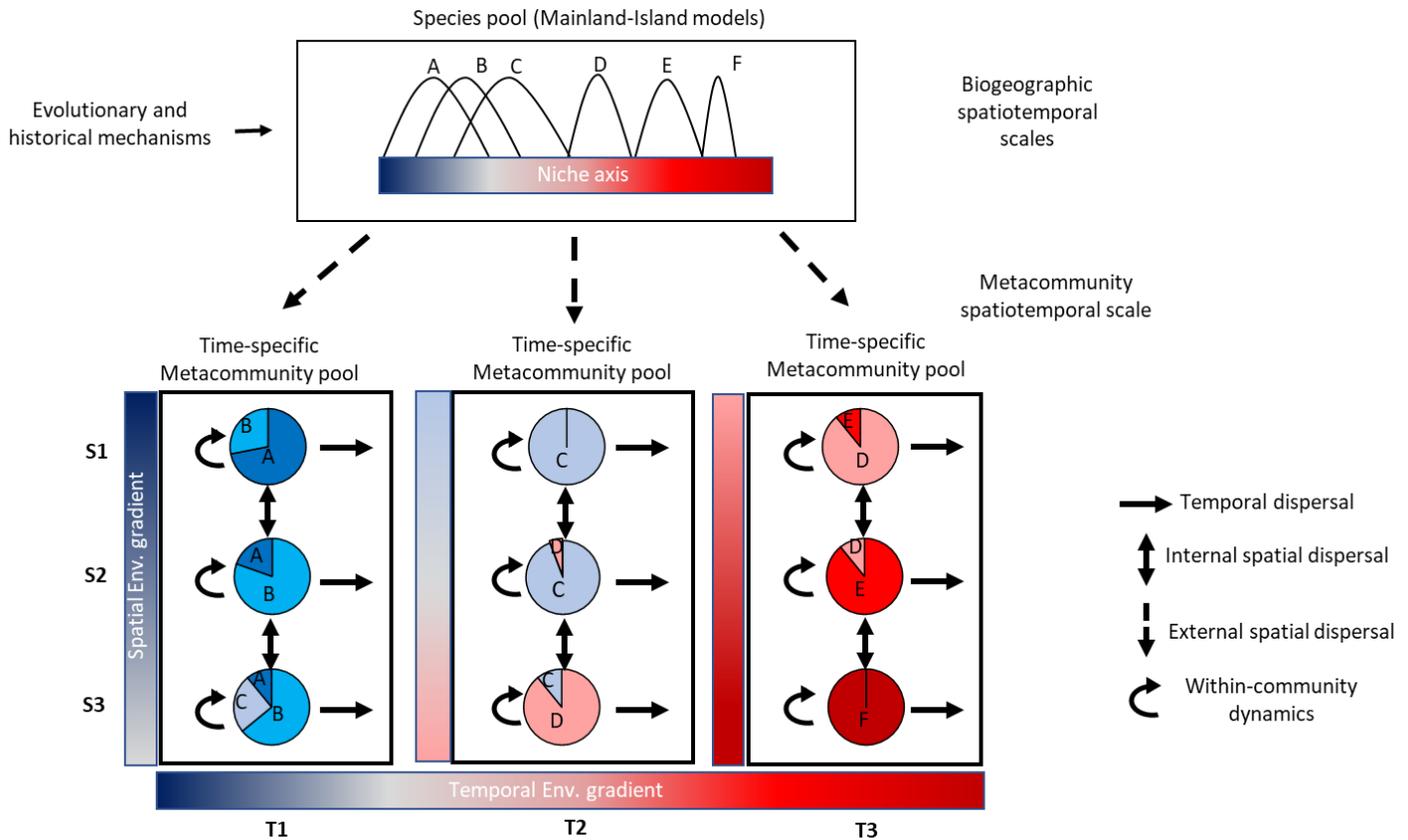
Figure 1.1: A conceptual framework for community assembly that integrates mainland-island and spatiotemporal metacommunity models. The local abundance of species (capital letters) in a specific locality (S1, S2, S3) at a specific moment in time (T1, T2, T3) is represented by different colors in the pie charts. The species pools of mainland-island and metacommunity models are linked by species dispersal in space and time (e.g., recovery from dormancy-like states). Following Fukami (2015), two types of spatial dispersal are represented: (i) external dispersal links the species pools of mainland-island models and the species pools of metacommunity models; and (ii) internal dispersal links communities in a metacommunity. Species pools of mainland-island models are fixed at fine ecological scales of community assembly but variable as a function of evolutionary and historical processes operating at biogeographic scales. Species pools of metacommunity models are temporally dynamic, changing as a function of within-patch mechanisms (species-habitat sorting, biotic interactions, demographic stochasticity) and internal (i.e., within metacommunity) dispersal. Although this conceptualization is shown here for illustrative purposes only, many of its aspects (e.g., different types of species pools, spatial and temporal dispersal, and within-community dynamics) are explicitly or implicitly considered in this thesis.

and evaluate causal relationships between species pool characteristics, landscape attributes, the relative importance of assembly mechanisms, and the overall structure of metacommunities.

Furthermore, while I acknowledge that speciation is a critical process to community assembly, investigating how bottom-up and top-down controls influence its relative importance on community composition is beyond the scope of this thesis (but see Webb et al. 2002; Hughes 2017; Rahbek et al. 2019,Leibold et al. 2022) Instead, I explore the effects of both types of controls on modulating the processes linking "contemporary" regional species pools to local diversity and vice-versa.

1.6 Chapters Overview, Impact, and Novelty of Research

1.6.1 Chapter 2: Determinism and stochasticity in the spatial–temporal continuum of ecological communities: the case of tropical mountains

> *"[…] no other experimental system 'designed' by nature beats the power of elevational gradients"* (Körner 2000)

Chapter 2 serves to introduce the perspective that well-informed predictions regarding the processes driving community variation in space and time should account for the mutual influences of top-down and bottom-up controls on community assembly (see graphical abstract in Figure 1.2). The narrative adopted in this chapter is based on the rationale that while similar processes may influence community composition in both space and time (here, ecological selection and drift), their relative influences may differ between these two dimensions (e.g., Stegen et al. 2013; Freestone and Inouye 2015; Van Allen et al. 2017). As such, in this chapter we ask: Under what ecological conditions should the mechanisms driving community assembly differ in space and time? To answer this question, we proposed a conceptual framework that predicts changes in the relative importance of ecological selection and drift in space and time when the following conditions are met: (i) species pools are dominated by ecological specialists that respond strongly to environmental variation; (ii) the landscape is characterized by a pronounced asymmetry in the steepness of environmental variation in space and time. Both conditions are commonly observed in tropical mountainous landscapes (Ghalambor 2006). Despite covering only about 25% of the land surface, mountains are home to one-third of global species diversity (Körner and Paulsen 2004). Thus, investigating the dynamics of ecological communities in tropical mountains fosters insights into community assembly under the ecological context experienced by a vast portion of current biodiversity.

To test these predictions, we started by introducing a general analytical framework that capitalizes upon previous work (Legendre and De Cáceres 2013, Legendre 2014) to partition the total variation of a species-by-site-by-time matrix (i.e., total beta-diversity; sites here representing patches or local communities in a given point in time) among its purely spatial (variation in space independent of time), purely temporal (variation in time independent of space) and spatiotemporal (i.e. variation across different sites across different moments in time) components. Through this framework, one can assess how beta-diversity is partitioned in space and time and estimate the relative importance of deterministic versus stochastic mechanisms across dimensions.
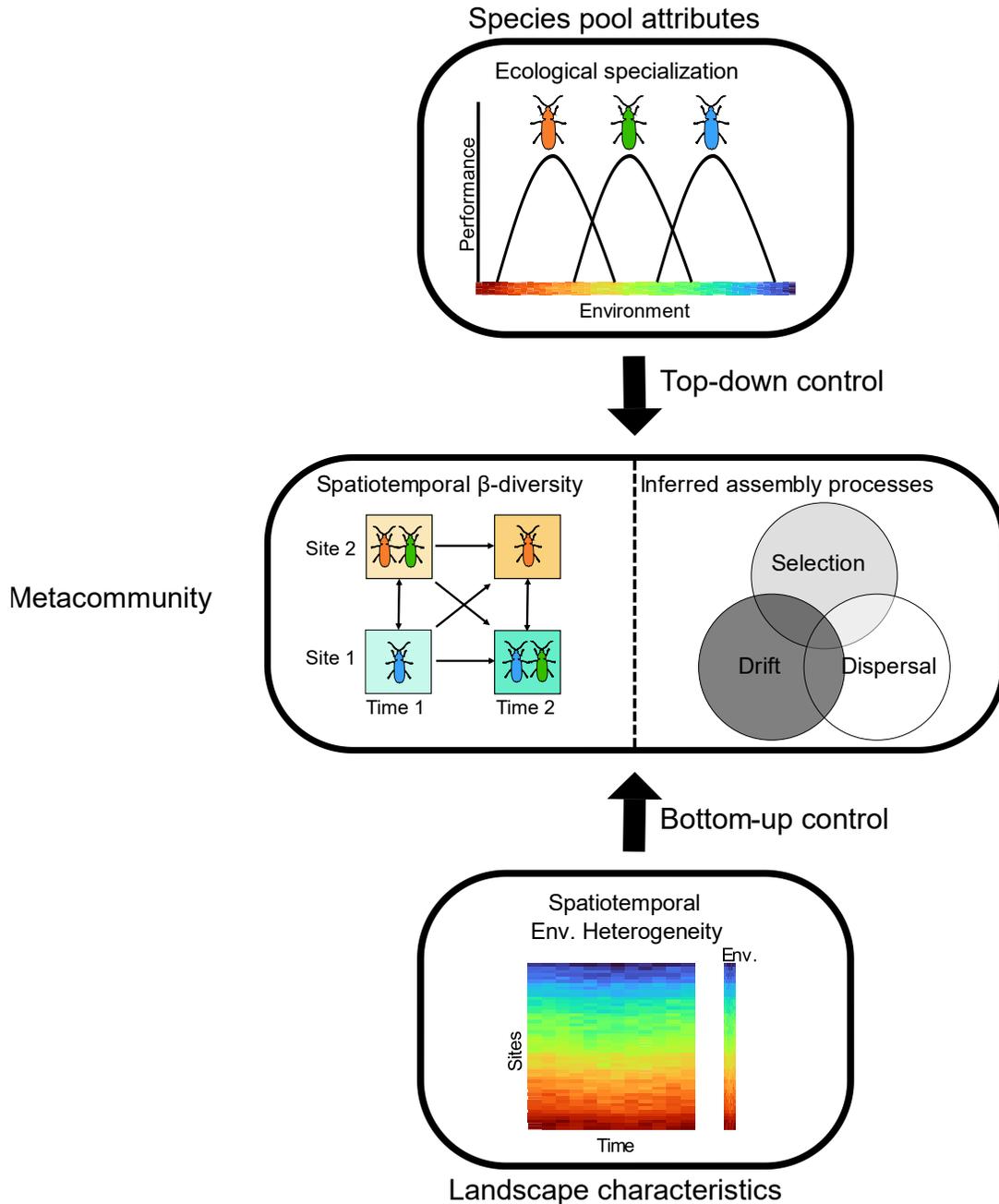
Figure 1.2: Graphical abstract for Chapter 2. In this chapter, we investigated how the typical spatiotemporal environmental heterogeneity found in tropical mountainous landscapes and the high degree of ecological specialization of its species pools determine: (i) spatial (middle panel, vertical arrows), temporal (horizontal arrows), and spatiotemporal (diagonal arrows) patterns in community variation (β-diversity); (ii) the relative importance of ecological selection and drift in shaping these patterns in space and time (Venn`s diagram).

Using simple simulation models that replicated in silico the conditions described above we provide theoretical support for our analytical framework is robust to assess these differences in the relative importance of assembly mechanisms across dimensions. Then, we provided

empirical support for our predictions by investigating patterns in the spatiotemporal distribution of distinct insect metacommunities in a tropical mountainous landscape.

Two important insights were gained from this study. Firstly, it provides compelling evidence that relying solely on unidimensional assessments of metacommunities may limit our ability to improve our understanding of their dynamics. This limitation arises from the fact that if different mechanisms govern community assembly in space and time, the classification of metacommunities into archetypal models based solely on spatial dynamics is insufficient to improve our understanding of community assembly (White et al. 2010; Wisnoski et al. 2019; Jabot et al. 2020). Secondly, our model and empirical results introduce the idea that the relative importance of assembly mechanisms hinges on where the assembly process occurs (i.e., the landscape) and the dominant traits (here, degree of climatic specialization) observed in the groups of species that coexist regionally. Framing the latter under the context of this thesis: community assembly is under a landscape-mediated bottom-up control and a species pool-mediated top-down control.

1.6.2 Chapter 3: The geography of metacommunities: landscape characteristics drive geographic variation in the assembly process through selecting species pool attributes

> *"[...] two exact pools of species may find different solutions for their coexistence based on different levels of landscape heterogeneity."*
> (Peres-Neto et al. 2012)

Chapter 3 expands some of the ideas elaborated in Chapter 2 by acknowledging the dominant life-history strategies observed in species pools that form metacommunities are inherently influenced by the characteristics of landscapes where the assembly process takes place (Büchi and Vuilleumier 2014; Fournier et al. 2020). We argue that acknowledging this non-random association between landscapes and species pools is relevant because it determines the geographic context of metacommunity dynamics: i.e., it drives predictable variation in the relative importance of mechanisms that assemble different metacommunities distributed along broad-scale ecological gradients or across (bio)geographic regions.

To illustrate that, we employed simulation models in which species pools with an identical initial distribution of niche breadths and dispersal abilities interacted within landscapes with contrasting features (Figure 1.3). At the end of each simulation interaction, we assessed which types of life history traits dominated each landscape type. Subsequently, we employed analytical approaches commonly used in the study of empirical metacommunities to investigate the resulting (simulated) metacommunities. By establishing the causal links between landscape attributes, species pool characteristics, and associated inferences about community assembly derived from statistical models, this chapter generates insights into why broad-scale empirical studies frequently observe (bio)geographic variation in metacommunity dynamics.
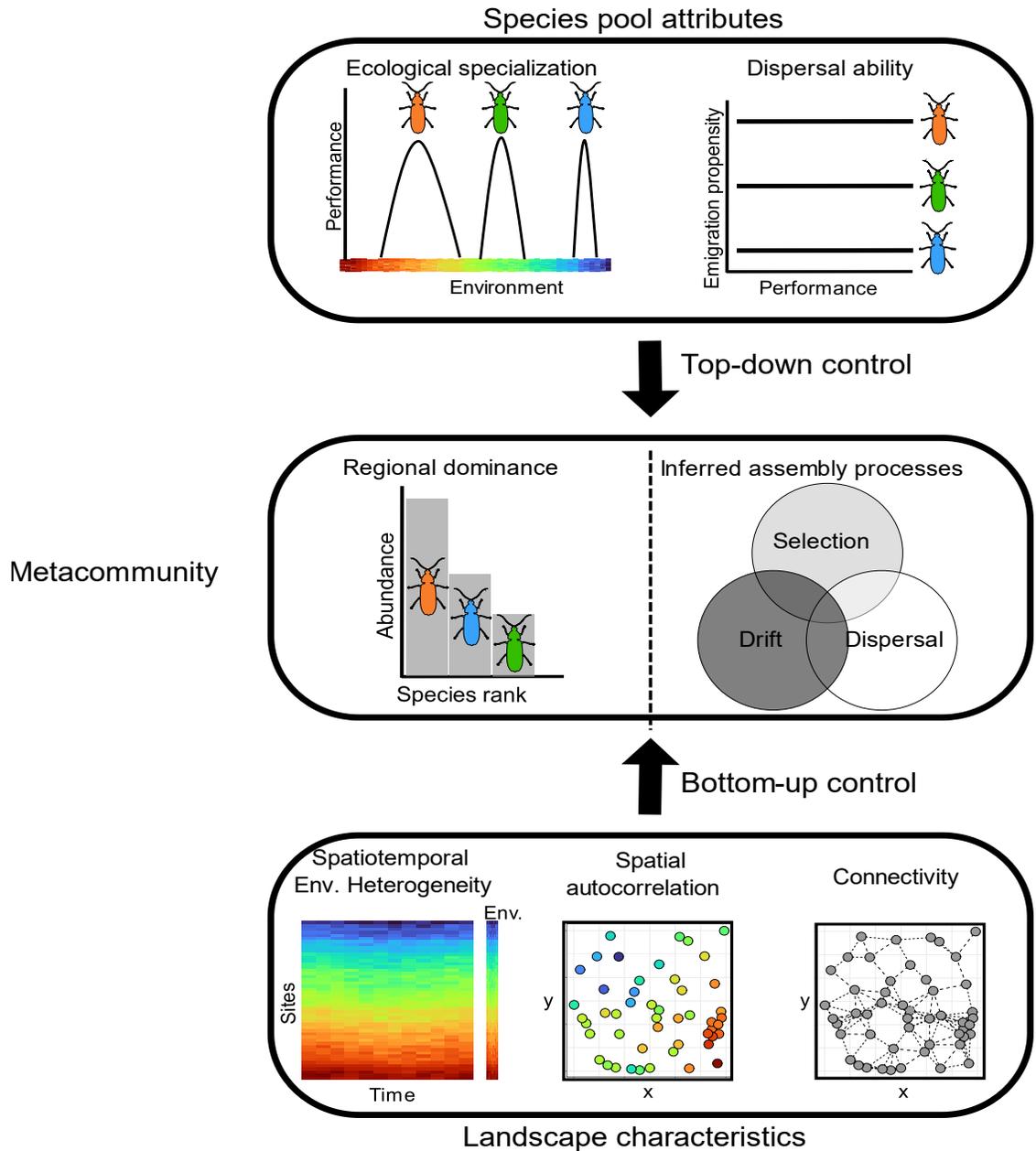
Figure 1.3 Graphical abstract for Chapter 3. In this chapter, we investigated the dominant life-histories (species ecological specialization and dispersal ability) observed in species pools that are ecologically selected (at the metacommunity scale, middle panel) by different landscape characteristics. We also investigated the causal links between landscape characteristics, species pool attributes, and our empirical inferences about the importance of mechanisms driving community assembly (Venn`s diagram). This chapter assumes that species dispersal abilities are species-specific but fixed (i.e., dispersal ability does not change depending on the surrounding abiotic and biotic conditions that determine species performances). More realistic assumptions about dispersal strategies are explored in Chapter 4.

We provided empirical support for some of the theoretical predictions derived from our simulation models by analyzing moth metacommunities in a tropical and a temperate mountainous

landscape that show distinct patterns of spatiotemporal environmental heterogeneity and, consequently, distinct species pool characteristics.

Chapter 3 makes two significant contributions. Firstly, it showcases the capacity of process-based simulations designed to study metacommunity dynamics at the landscape scale to replicate fundamental patterns underlying important hypotheses in Biogeography and Macroecology. Secondly, the chapter offers testable predictions that form the basis for understanding the variation in geographic patterns observed in community assembly. These predictions offer a valuable foundation for subsequent empirical studies, enabling researchers to understand the ecological causes of variation in the relative importance of assembly processes across different geographic contexts.

1.6.3 Chapter 4: Ecological selection of dispersal strategies in metacommunities: impact of landscape features and competitive dynamics

> *"[…] Dispersal can produce ecological patterns, but these patterns can again influence the selective pressures on dispersive traits. Only by closing this loop, that is, by realizing that dispersal can at the same time act as both a cause and an effect, will we get the full picture of the ecology and evolution of dispersal."* (Clobert et al. 2012)

Chapters 2 and 3 investigated how bottom-up and top-down controls determine the relative importance of assembly processes. In Chapter 4, our attention shifts to the intriguing question of how both types of controls can modulate the influence of ecological selection on species dispersal (see graphical abstract in Figure 1.4). The inspiration for this chapter comes from the growing body of empirical evidence indicating that dispersal can be simultaneously a cause and a consequence of metacommunity dynamics (e.g., De Meester et al. 2015; Fronhofer et al. 2018). More specifically, we tested whether metacommunity dynamics, a direct consequence of how species in regional pools interact with each other and the features of the surrounding landscape, select for different context-dependent dispersal strategies that can maximize species persistence and dominance in the metacommunity.

To investigate that, we employed simulation models that assumed species-specific context-dependent adjustments in decisions involving the timing to leave natal patches (i.e., emigration propensity), traveling distances (i.e., traversal), and the selection of a suitable new patch to settlement (i.e., habitat selection). This sequence of decisions determines the three stages of dispersal events: departure, movement, and settlement (*sensu* Clobert et al. 2009). We allowed species with distinct context-dependent dispersal strategies to reach coexistence at the metacommunity (regional) scale under different types of competition dynamics and under varying levels of spatial and temporal environmental variability. By analyzing the context-dependent dispersal strategies of the species that successfully persisted and dominated within these metacommunities, we could formulate well-informed predictions about how metacommunity dynamics influence the success of different dispersal strategies. We finish by contrasting the

theoretical predictions derived from our framework with existing empirical studies investigating
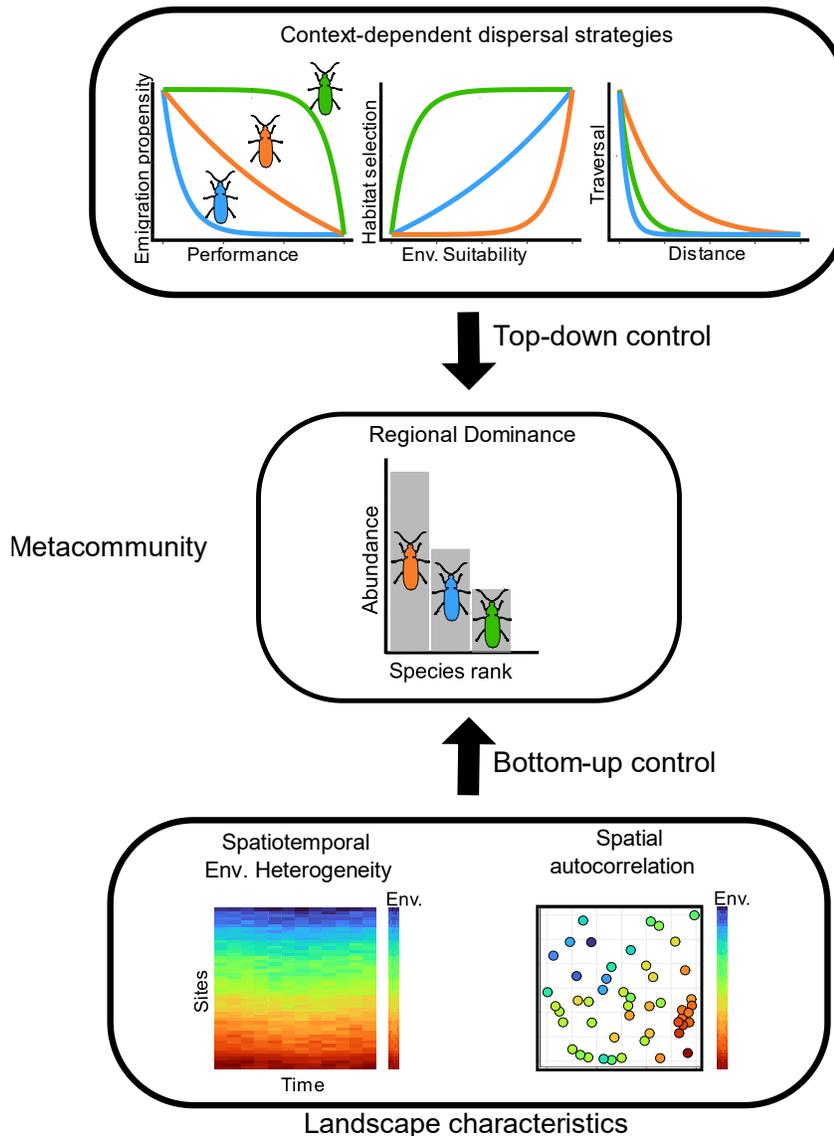
Figure 1.4 Graphical abstract for Chapter 4. In this chapter, we investigated how different landscape features and competitive dynamics within species pools (not shown) select for the success of different context-dependent dispersal strategies in metacommunities (success is measured as species dominance at the metacommunity scale, middle panel). There, dispersal strategies for emigration propensity, habitat selection, and traversal are species-specific and plastic (context-dependent). This chapter serves as a theoretical essay about the dual nature of dispersal in community assembly theory, i.e., it is simultaneously a cause and consequence of metacommunity dynamics.

the effects of species interactions and landscape attributes on dispersal strategies.

In addition to expanding upon the ideas developed in Chapter 3 regarding the influence of landscape attributes on the dominant life-history strategies in species pool, Chapter 4 integrates species-specific context-dependent dispersal strategies, a phenomenon extensively studied in movement and dispersal ecology, into the basis of community assembly theory. This integration

is particularly relevant because a greater understanding of the impacts of global change on biodiversity hinges on a better understanding of the forces that govern species dispersal (Urban et al. 2016).

1.6.4 Chapter 5: Uncovering the trajectories of metacommunities: insights gained from the keystone community concept

> *"[...] understanding the ecological mechanisms underpinning why change in beta diversity is more pronounced in some areas and less so (or absent) in others provides opportunities for ecologists and managers to identify specific spatial units that may serve as priorities for monitoring [...] and conservation interventions."*
> (Rolls et al, 2023)

Finally, Chapter 5 illustrates how considering the bottom-up and top-down controls of community assembly can aid in developing and implementing effective conservation and management strategies (see graphical abstract in Figure 1.5). In this chapter, we argue that effective plans to manage and protect metacommunities must acknowledge that their local communities play different roles in shaping their internal structure (Mouquet et al. 2013; Yang et al. 2020). When extirpated or disturbed, some local communities may trigger strong cascading secondary effects on extinction and colonization patterns, ultimately driving temporal changes in the structure of the remaining communities in a metacommunity. However, identifying "keystone communities" (*sensu* Mouquet et al. 2013) and understanding the conditions that make a community have (or not) a keystone role at the metacommunity level remain challenging.

In Chapter 5, we introduce a novel quantitative framework designed to estimate the importance of any given local community in maintaining the internal structure of their metacommunities over time (i.e., referred here as community "keystoneness"). The proposed framework can be implemented directly on empirical data on species distributions and generates estimates of keystoneness that are statistically independent of local diversity and community size (i.e., akin to the sampling effects of biodiversity in which richer local communities are expected, by chance, to have greater importance to metacommunity structure).

Through simulation models, we demonstrate that the characteristics of habitat patches (suitability and connectivity) determining the keystoneness of ecological communities change predictably with the dispersal ability and degree of ecological specialization of the species in the regional pools. Understanding the links between species dispersal ability, ecological specialization, and the features of habitat patches that support keystone communities should enable us to formulate more effective strategies for preserving and conserving entire metacommunities.

To showcase the added depth our proposed analytical framework brings to understanding the effects of anthropogenic stressors on metacommunities, we employed it to examine how light pollution impacts the structure of a moth metacommunity. This metacommunity is situated in a

protected mountainous landscape surrounded by urban settlements and, consequently, under the
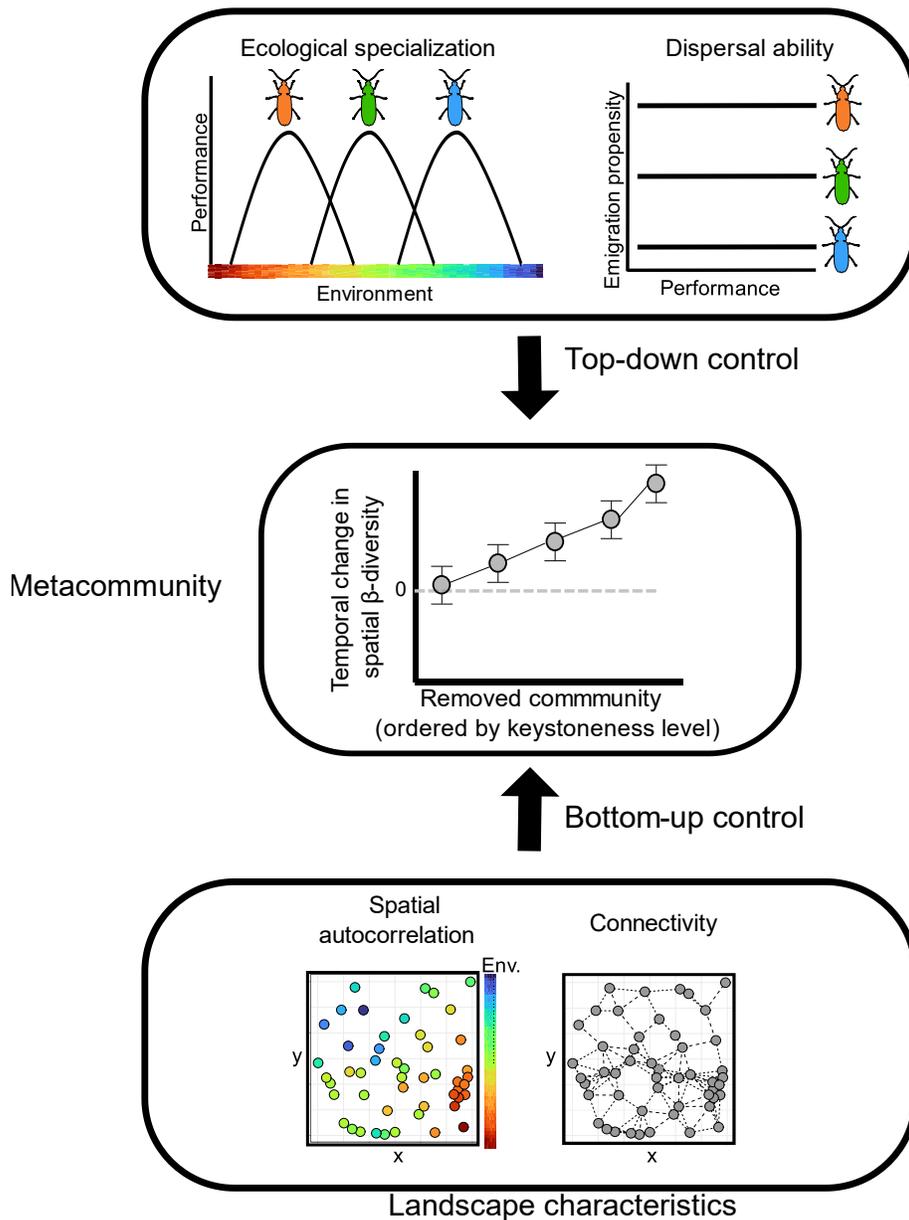
Figure 1.5: Graphical abstract for Chapter 5. In this chapter, we investigate how different life-histories in the regional pool (different combinations of ecological specialization and dispersal ability) and landscape features influence the attributes of local patches harbouring keystone communities. Keystone communities are those whose extirpation causes cascading secondary effects on extinction and colonization patterns, ultimately driving significant temporal changes in the spatial β-diversity (middle panel).

direct and indirect (in the form of artificial skyglow) influence of light pollution. We provide correlational evidence suggesting that, despite being in a protected area, the structure of this moth metacommunity is strongly affected by light pollution from surrounding sources. More specifically, our findings indicate a positive correlation between light pollution and local diversity,

yet a negative correlation with the keystoneness degree of local communities. This finding exemplifies how the proposed framework can provide valuable insights into the conservation value of local communities in situations where the effects of anthropogenic stressors on dispersal diminish the reliability of local diversity as an indicator of habitat suitability. Furthermore, these results emphasize the importance of considering metrics beyond local richness to assess the conservation value of local communities.

Chapter 2 : Determinism and stochasticity in the spatial-temporal continuum of ecological communities: the case of tropical mountains[45]

2.1 Abstract

Ecological communities are assembled in a spatial-temporal continuum. However, we still have a poor understanding of the relative importance of different mechanisms structuring community composition (i.e., beta-diversity) in space and time. In this study, we start by introducing a conceptual model that capitalizes upon the core-occasional species concept to predict that the assembly process in tropical mountains is driven by the deterministic turnover of core species in space via habitat sorting, but the turnover of occasional species through time via stochastic events of colonization and local extinctions. We then propose a general analytical framework that allows assessing these predictions by partitioning the total variance of a species-by-site-by-time matrix (i.e., total beta-diversity) among its purely spatial (variation in space independent of time), purely temporal (variation in time independent of space), and spatiotemporal (i.e., variation across different sites across different moments in time) components. Through simulation models, we provided theoretical support that the proposed analytical framework is suitable to test the predictions derived from our conceptual model. We then used this framework to identify general patterns and quantify the relative importance of processes underlying the spatial and temporal organization of ten distinct insect metacommunities along a tropical elevational gradient. As predicted, we found that, across taxa, spatial beta-diversity was mainly explained by environmental variation alone: a pattern that indicates the spatial turnover of core species. In contrast, temporal beta-diversity could not be distinguished from the expectation of null models where communities are simply represented by random draws from species pools: a pattern that indicates a temporal turnover of occasional species within communities. Taken together, our findings illustrate how our conceptual model and quantitative framework can articulate a better understanding of community assembly in space and time.

2.2 Introduction

Community assembly theory focuses on the processes through which species from a large regional pool come together to form assemblages at smaller scales (Fukami 2015; Mittelbach and Schemske 2015; Vellend 2018). Early conceptual models argued that the assembly process in space and time should be underpinned by analogous ecological mechanisms (Preston 1960). This

---

[4] Adopting the terminology of Chase and Myers (2011), here we refer to the influences of selection and ecological drift in community assembly as deterministic and stochastic mechanisms, respectively.

[5] Khattar, G., Macedo, M., Monteiro, R., Peres-Neto, P.R. (2021) Determinism and stochasticity in the spatial–temporal continuum of ecological communities: the case of tropical mountains. Ecography, 44(9), 1391-1402

argument aligns with recent attempts to synthesize the "overwhelming" number of factors (*sensu* Lawton 1999) that may determine the size and composition of ecological communities (e.g., Vellend 2010, 2018, Chase and Myers 2011). Despite the many possible mechanisms underlying spatiotemporal changes in the composition of contemporary communities (i.e., not considering speciation), they can be sorted into two broad classes (*sensu* Chase and Myers 2011). Deterministic mechanisms drive dynamics in community structure due to differences in species fitness often related to differences in traits and environmental niches (e.g., species-environment sorting and biotic interactions). Stochastic mechanisms are those that dictate variation in community structure that occur at random with respect to species identities, niche components, and life history characteristics (Vellend et al. 2014). Common examples of stochastic mechanisms are demographic stochasticity (Adler and Levine 2007; Orrock and Watling 2010; Shoemaker et al. 2020*a*) and the stochastic recruitment of individuals from regional pools into local communities (i.e., sampling effects, Kraft et al. 2011, Tucker et al. 2016). Along with dispersal, deterministic and stochastic mechanisms simultaneously shape community structure in space and time (Jabot et al. 2020).

While the assumption that deterministic and stochastic mechanisms underly the structure of communities is conceptually reasonable, we should not readily assume that their relative importance within metacommunities is the same in space and time. Strong statistical interactions between spatial and temporal variation in species abundances/occurrences suggest that the relative importance of mechanisms driving community assembly in one dimension (e.g., space) may differ from their relative importance in the other (i.e., time; e.g., Collins and Glenn 1991, Legendre et al. 2010, Ward et al. 2015). However, very few studies have assessed and contrasted the relative importance of deterministic and stochastic mechanisms in underlying the assembly process in both dimensions (e.g., Stegen et al. 2013, Freestone and Inouye 2015, Van Allen et al. 2017). Consequently, we still poorly understand: (i) the conditions under which the mechanisms driving community assembly within metacommunities should differ in space and time, and; (ii) how these differences influence the way in which changes in community composition are partitioned in space and time. Such understanding is relevant because it paves the way for synthesis that aims at extending the framework of community assembly theory beyond its traditional spatially-oriented scope (e.g., White et al. 2010, Wisnoski et al. 2019).

In this context, the main contribution of the present study is twofold. First, we introduce a conceptual model that describes the conditions under which the relative importance of deterministic and stochastic mechanisms should differ in space and time within metacommunities. Second, we introduce a general analytical framework that capitalizes upon previous work (see Legendre and De Cáceres 2013, Legendre 2014) to partition the total variance of a species-by-site-by-time matrix (i.e., total beta-diversity) among its purely spatial (variation in space independent of time), purely temporal (variation in time independent of space), and spatiotemporal (i.e., variation across different sites across different moments in time) components. Through this framework the relative importance of deterministic versus stochastic mechanisms can be estimated

in space and time (see Methods). The conceptual model and the analytical framework are summarized in Figure 2.1.

Our conceptual model describes the assembly process at spatial and temporal scales where the influence of speciation and dispersal limitation on beta-diversity is not relevant (e.g., along short spatial and temporal gradients). It posits that the relative importance of deterministic and stochastic mechanisms will differ in space and time when two primary conditions are met. The first condition is that not all species found in a given community at a specific moment in time are in equilibrium with local habitat conditions. As such, local communities are composed of two different types of species: core and occasional (*sensu* Magurran and Henderson 2003, Snell Taylor et al. 2018, 2020). Core species are those whose environmental requirements match with local habitat conditions and, consequently, can sustain viable local populations over time even in the absence of immigration (Coyle et al. 2013). Occasional species are then those poorly suited to local habitat conditions and whose local occurrence/permanence depends more on the random immigration of individuals from neighboring communities than on environmental matching (White et al. 2010; Umaña et al. 2017). See an extended discussion on the core-occasional species framework in Supp. Information, Box I, and references within.

Acknowledging the existence of core and occasional species within metacommunities is relevant because their distributions in space and time are driven by distinct assembly mechanisms (Belmaker 2009; Snell Taylor et al. 2018). The diversity of core species across communities is governed by deterministic mechanisms such as trait and/or environmental selection (Umaña et al. 2017). In a variation partitioning framework (*sensu* Borcard et al. 1992, Peres-Neto et al. 2006) the turnover of core species is represented by the amount of variation in the response community matrix explained by environmental variation alone (see Supp. Information and Box I for proof of concept). Conversely, the diversity of occasional species across communities result from stochastic events of colonization and local extinctions. Under a null model framework, the turnover of occasional species generates beta-diversity patterns that equate to patterns underlying the stochastic sampling of individuals from the regional species pool into local communities (i.e., see Supp. Information and Box I for proof of concept).

$$\text{BD}_{total} \qquad \text{BD}_{time} \qquad \text{BD}_{space} \qquad \text{BD}_{SpxT}$$

$$\frac{1}{n(n-1)}\sum D_{time,space,SpxT} = \frac{1}{n(n-1)}\sum D_{time} + \frac{1}{n(n-1)}\sum D_{space} + \frac{1}{n(n-1)}\sum D_{SpxT}$$

Figure 2.1: A summary of our conceptual model and quantitative framework. The heatmap in the left represents the spatiotemporal structure of environment in the study system (i.e., scores of elevation-by-month samples in the first principal component of the correlation matrix of climate variables). It is highly stratified elevationaly but not seasonally, which is typical for mountain systems (i.e., high vertical but low horizontal color variation in the heatmap). The panel in the middle represents four fictional communities varying in time (T1; T2) and space (S1; S2). The resulting Sørensen's dissimilarity matrix (right panel) is then used to partition the total variance of the species-by-site-by-time matrix (total beta-diversity, BD_{total}, Legendre and De Cáceres 2013) into its purely spatial (i.e., BD_{space}, red arrows), purely temporal (i.e., BD_{time}, green arrows), and spatiotemporal (i.e., $\text{BD}_{SpxT,}$ purple arrows) components based on the equation at the bottom. Here we consider two types of species within each community (adapted from White et al. 2010). Core species (large font size) are well adapted to local environmental conditions and, consequently, maintain long-term viable local populations even in the absence of immigration. Occasional species (small font size) are not well adapted to local environmental conditions and, consequently, their local occurrence depends on the random immigration of individuals from neighboring communities. In this example, BD_{space} is given by the turnover of core species caused by the high stratification of environmental conditions in space (i.e., deterministic species-environment sorting mechanisms). The low stratification of environmental conditions in time allows the maintenance of the local population of core species. As such, BD_{time} represents the turnover of occasional species driven by stochastic events of colonization and local extinctions that cannot be distinguished from the random allocation of individuals from the regional species pool into local communities (i.e., stochastic sampling effects). BD_{SpxT} is then the outcome of the deterministic turnover of core species in space and stochastic turnover of occasional species in time.

The second condition in which the relative importance of deterministic and stochastic mechanisms will differ in space and time is that the steepness of environmental gradients must differ between these two dimensions. This condition is relevant because it ensures that spatial and

temporal beta-diversities will estimate the turnover of either type of species across communities and, consequently, will be explained by distinct assembly mechanisms. For instance, consider the case of tropical mountainous landscapes, a ubiquitous and highly diverse system where environmental conditions change faster in space than in time (see Figure 2.1 Ghalambor 2006, Zuloaga and Kerr 2017). In these landscapes, our conceptual model predicts that beta-diversity along the steep spatial environmental gradient will be mainly explained by environmental variation: a pattern that indicates the high spatial turnover of core species (i.e., elevational specialists) across communities via species-environment sorting. In opposition, the shallow environmental gradient in time allows the local persistence of core species. As such, temporal beta-diversity will estimate the turnover of occasional species caused by stochastic colonization and extinction that generate diversity patterns similar to the ones expected under ecological null models.

To test the predictions derived from our conceptual model, we used the proposed analytical framework to assess and contrast the patterns underlying the spatial and temporal distribution of 10 functionally distinct groups of beetles and wasps along an elevational gradient in the Atlantic Rainforest, South America. Through mechanistic simulations that reproduced *in silico* the conditions proposed in our conceptual model, we provide theoretical support for our assumptions that different mechanisms drive the turnover of core and occasional species and that our analytical framework is robust to assess these differences (see Box I and Supp. Information). Even though our simulations were not parametrized on the basis of empirical metacommunities, they yielded results qualitatively similar to what was observed in the empirical data (see more Supp. Information). These similarities indicate that our conceptual model serves as a proper approximation of the complex dynamics dictating the spatiotemporal organization of real metacommunities.

2.2 Methods

Sampling was carried out in the Serra dos Orgãos National Park (22°27′49″S; 43°01′50″W), Rio de Janeiro, Brazil. The park is one of the most preserved remaining fragments of the Atlantic Rainforest (Castro 2018), one of the hottest hotspots of global biodiversity (Myers et al. 2000). Along a complete elevational transect that ranges from 100 to 2130 meters above sea level (MASL), two flight interception Malaise traps were placed every interval of approximately 150 MASL, summing up to 15 sampled elevations. The collecting bottle of each Malaise trap was replaced every month from December 2014 to November 2015, summing up to 360 samples (2 traps x 15 elevations x 12 months). A data logger was placed next to each pair of traps to record hourly variation in climatic conditions (air temperature and relative humidity) over the entire sampling period. Insects captured in each sample were sorted at the species level by our team and external collaborators (see acknowledgments). We focused our study on ten relatively well-known families and subfamilies of beetles (Coleoptera) and wasps (Hymenoptera). Refer to Results and Supp. Information for more information about the diversity and functional role of the selected taxonomic groups.

Box I: Distinct mechanisms dictate the distribution of core and occasional species within metacommunities

Here we make use of mechanistic simulation models to provide theoretical support for our assumptions that the turnover of core and occasional species is driven by, respectively, deterministic and stochastic mechanisms (as shown in empirical data in Belmaker 2009, Coyle et al. 2013, Umaña et al. 2017). We used simulated rather than observational data to validate these assumptions because we wanted to ensure that species were correctly classified as either core or occasional community members. This assessment in empirical communities is a daunting task and is prone to errors when there is a lack of knowledge about the temporal occurrence, traits, and life history of species under consideration (as is the case of most tropical insects). It is also prone to errors when it is made based on observational data sampled in highly heterogeneous landscapes at relatively fine temporal scales (i.e., as is the case of our data, see extended discussions in Snell et al. 2018, 2020). As such, instead of simply proposing a conjecture in the discussion of how communities are structured by differences in turnover dynamics of these two types of species, we provide a demonstration based on a theoretical model (details in Supp. Information).

The data used to validate our assumptions was simulated using a spatially implicit metacommunity model where population dynamics are discrete in time and were modeled according to a competition form of the Beverton-Holt model of population growth (Beverton and Holt 1957) . Dispersal is global (i.e., all sites have equal probabilities of receiving immigrants), and community assembly at the time "t" is the outcome of three sequential steps: (i) within patch dynamics (i.e., demographic stochasticity, intra and interspecific competition, and habitat selection); (ii) emigration and; (iii) immigration (Shoemaker and Melbourne 2016). Metacommunity dynamics were modeled in a virtual landscape where the spatiotemporal structure of environmental conditions resembles the one observed in tropical slopes (see Figure 2.1 and Supp. Information Fig. 1). Population dynamics were carried across 15 different sites over 1200 time steps (each time step representing one month, i.e., 100 years in total). To ensure that analyses were performed on stable rather than transient communities, only the last 12 time steps (i.e., the last year) were considered in the final species-by-site-by-time matrix. By keeping track of each species' growth rate in each site over time, we could discriminate which species were core and occasional members of each community. More specifically, species whose average long-term growth rate through time was higher than or equal to 1 were considered core members of the focal site, while the others were considered as "occasional" members.

At the end of each of the 50 simulation rounds, we applied the analytical framework described in this study (see Methods) to analyze the final species-by-site-by-time matrix (i.e., hereafter full matrix). On average, these analyzes yielded results that were qualitatively similar to the results obtained in our empirical data (e.g., compare the first panel of Box I Figure I and Figure 2. 4 in the main text. Also, see Supp. Information). This similarity indicates that our theoretical model serves as a proper approximation of the complex dynamics dictating the spatiotemporal organization of the empirical metacommunities. Finally, when analyzing simulated matrices in which either core or occasional species had been removed from communities, we showed that deterministic species-environment sorting mechanisms mainly explained the beta-diversity of core species, and the beta-diversity of occasional species was better explained by the stochastic allocation of species from the regional pool into local communities. These results support our initial assumption that the relative importance of distinct mechanisms dictating community assembly is related to the turnover of different types of species within metacommunities. They also indicate that our analytical framework can assess these differences properly.

Box I: **Distinct mechanisms dictate the distribution of core and occasional species within metacommunities**
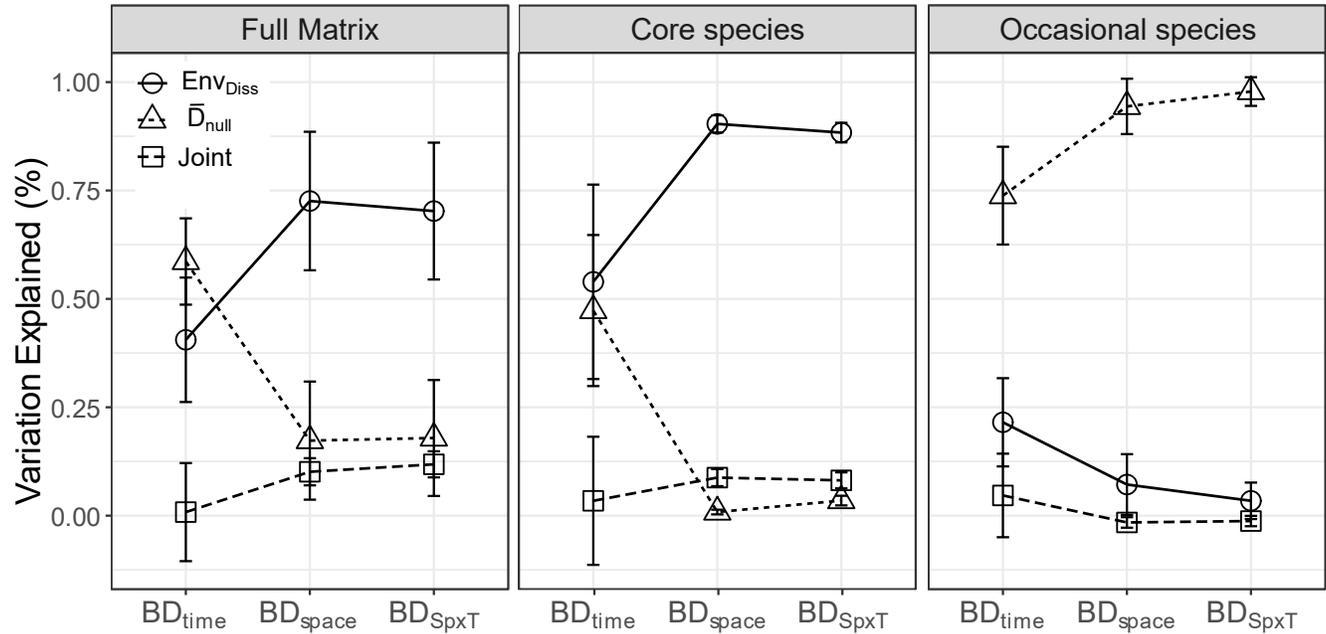
Figure Box I Interaction plots showing the percentage of the total variation in the spatial (BD$_{Space}$), temporal (BD$_{Time}$), and spatiotemporal (BD$_{SpxT}$) dimensions of the compositional dissimilarity matrix explained by environmental dissimilarities (Env$_{diss}$), stochastic sampling effects ($\bar{\textbf{D}}_{\textbf{null}}$) and their joint contribution (Joint). Symbols represent the mean values obtained across all 50 simulation rounds. The panels represent the results obtained when either core and occasional (i.e., Full matrix), only core species (middle), or only occasional (right panel) members of communities were considered in the proposed analytical framework. Distinct mechanisms explain the spatiotemporal distribution of core and occasional species in the metacommunity. Whiskers = Standard error of the mean.

We considered only the adults of the focal taxonomic groups in the final species-by-time-by-site matrices (i.e., a total of 6996 individuals from 549 species across taxa). This is because flight-interception traps such as Malaises, by definition, do not effectively capture individuals at early ontogenetic stages (i.e., apterous larvae and pupae). Given that the age structure of many insect populations in the region changes fast within short periods (Flinte et al. 2009, 2015), there is an inevitable decrease in the capacity of Malaise traps in detecting species in months when populations are mainly composed of juvenile individuals. As such, we decided to consider only months with the highest activity of adults across all taxonomic groups (i.e., from December to February and June to August). This allowed reducing the cases of false absences but did not affect the completeness of sampling effort (as can be inferred by the asymptotic individual-based-rarefaction curves estimated for most taxonomic groups, see Supp. Information). Additionally, this decision did not alter the qualitative component of results as our analytical framework is robust in

controlling for unequal sampling efforts in space and time when estimating beta-diversity across dimensions (see below and results of simulations presented in Supp. Information).

2.2.1 Partitioning beta-diversity into its spatial, temporal and spatiotemporal dimensions

Beta-diversity can be estimated as the total variance in a community compositional matrix (hereafter BD_{total}, see Legendre and De Cáceres 2013). BD_{total} can be calculated from a matrix of pairwise dissimilarities **D** as follows:

$$BD_{total} = \sum_{j=1}^{n-1} \sum_{i=j+1}^{n} D_{ji} \, / [n(n-1)] \quad (2.1)$$

where n is the number of samples (local communities) and D_{ij} is the compositional dissimilarity between the ith and j^{th} communities measured using either (abundance-based) Sørensen's or Jaccard's dissimilarity coefficients (see more in Legendre and De Cáceres 2013, Legendre 2014). We used both coefficients, but we chose for no particular reason to report on Sørensen's as both yielded very similar results.

Local communities were represented by the individuals of all species of a given taxonomic group sampled at a specific site (i.e., elevation) at a particular time (i.e., month). Given that each local community has a location in space and time, we can further determine whether the compositional dissimilarity between any pair of local communities in **D** is either purely spatial, purely temporal, or spatiotemporal (Figure 2.1). For instance, consider a focal community i located at elevation S_1 in month T_1 (represented as $|S_1T_1|$ as in Figure 2.1). The dissimilarity between communities i and j (D_{ij}) estimates spatial beta-diversity if j is located at $|S_2T_1|$; temporal beta diversity if j is at $|S_1T_2|$; finally, it estimates spatiotemporal beta-diversity if j is at $|S_2T_2|$.

We decomposed **D** into subsets of pairwise dissimilarities representing either compositional differences in space alone (D_{space}), time alone (D_{time}), or in space and time simultaneously (D_{SpxT}). From equation 2.1, we can further partition BD_{total} into its purely spatial ($BD_{spatial}$), purely temporal ($BD_{temporal}$), and spatiotemporal (BD_{SpxT}) additive components as follows:

$$BD_{total} = BD_{space} + BD_{time} + BD_{SpxT} \qquad (eq.\ 2.2)$$

where

$$BD_{space} = \sum D_{space} \, / [n(n-1)] \qquad (eq.\ 2.3)$$

$$BD_{time} = \sum D_{time} \, / [n(n-1)] \qquad (eq.\ 2.4)$$

$$BD_{SpxT} = \sum D_{SpxT} \, / [n(n-1)] \qquad (eq.\ 2.5)$$

In Figure 2.1, we depict an example where the number of entries in matrix **D** representing D_{space}, D_{time}, and D_{SpxT} is the same (i.e., two entries representing each dimension), representing a balanced design. In this case, differences in BD_{space}, BD_{time}, and BD_{SpxT} can be directly contrasted and will only reflect differences in the amount of variation in the species matrix represented by

compositional dissimilarities in each dimension. However, if another site (e.g., S$_3$) had also been surveyed in T$_1$ and T$_2$, the number of entries in matrix **D** representing D$_{space}$, D$_{time}$, and D$_{SpxT}$ would have been six, three, and six, respectively. In this case, rather than representing only differences in compositional variation in each dimension, differences in BD$_{space}$, BD$_{time}$, and BD$_{SpxT}$ may also result from unbalances in the number of entries in **D** representing each dimension. As such, to fairly compare the contribution of each dimension to BD$_{total}$ in unbalanced designs, we must divide the values obtained for BD$_{space}$, BD$_{time,}$ and BD$_{SpxT}$ by the number of entries in **D** representing, D$_{space}$, D$_{time}$, and D$_{SpxT}$, respectively. Even though these fractions do not sum up to BD$_{total}$, they are directly comparable and can be understood as the average contribution of each dimension to BD$_{total}$. We used simple simulations to demonstrate that averaging effectively controls for differences among BD$_{space}$, BD$_{time}$, and BD$_{SpxT}$ that result solely from unbalances in **D** (see Supp. Information for details). As such, our framework is suitable to analyze community data sampled in different sites at different moments in time, and it allows one to contrast the contribution of each dimension to total beta-diversity even if there are differences in sampling effort across dimensions.

2.2.2 Decomposing beta-diversity into its Turnover and Nestedness components

Comparing the importance of nestedness (i.e., ordered loss of species in a way that species poor sites are a subset of species rich sites) and turnover (i.e., the replacement of species among communities) components across dimensions may provide additional insights into how communities respond to environmental gradients (Baselga 2010). We can also adapt our framework to decompose the pairwise dissimilarities in **D** into their additive turnover and nestedness components (see Legendre 2014 for detailed calculations) as:

$$BD_{total} = BD_{turn} + BD_{nest} \text{ (eq. 2.6)}$$

where

$$BD_{turn} = \frac{\sum D_{turn\ space}}{[n(n-1)]} + \frac{\sum D_{turn\ time}}{[n(n-1)]} + \frac{\sum D_{turn\ SxT}}{[n(n-1)]} \quad \text{(eq. 2.7)}$$

$$BD_{nest} = \frac{\sum D_{nest\ space}}{[n(n-1)]} + \frac{\sum D_{nest\ time}}{[n(n-1)]} + \frac{\sum D_{nest\ SxT}}{[n(n-1)]} \quad \text{(eq. 2.8)}$$

Using equations 2.7 and 2.8, one can assess the turnover and nestedness beta-diversity components in each dimension separately.

2.2.3 Assessing the relative importance of environmental filtering and stochastic sampling effects across dimensions

We estimated the relative importance of deterministic and stochastic mechanisms in structuring beta-diversity in each dimension in two different ways. In the first approach, we used null models to determine the pairwise dissimilarities expected values if community composition were generated by random draws of individuals from the regional species pool (Kraft et al. 2011). By fixing the relative abundance of species in the regional pool and the local abundance of communities across randomizations, these null models allowed generating communities in the absence of spatiotemporal intraspecific aggregation as expected by deterministic species-

environment relationships (Myers et al. 2013; Engel et al. 2020). We calculated a dissimilarity matrix for each of 1000 random generated metacommunities ($\mathbf{D_{null}}$) that are used below to estimate standardized effect sizes of pairs of entries in D_{space}, D_{time}, and D_{SpxT}. To infer the role of random sampling in generating dissimilarity patterns in each dimension, we further estimated the standardized effect size (SES_{ij}) of each pairwise dissimilarity. The SES_{ij} measures, in standard deviation units, how much the observed pairwise dissimilarities between communities i and j deviates from the null expectation. The SES_{ij} can be estimated as follows:

$$SES_{ij} = (D_{ij} - \overline{D}_{ij\,null})/\sigma_{D_{ij\,null}} \qquad \text{(eq. 2.9)}$$

where $\overline{D}_{ij\,null}$ and σ are the average and standard deviation of $D_{ij\,null}$, respectively. By estimating the mean SES_{ij} values considering all pairwise dissimilarities representing D_{space}, D_{time}, and D_{SpxT} separately, we can infer and contrast the relative importance of deterministic versus stochastic mechanisms underlying community assembly within each dimension. A mean SES_{ij} value close to 0 indicates that the observed pairwise dissimilarities representing beta-diversity in the focal dimension are, on average, not different from what would be expected under stochastic sampling alone. A mean SES_{ij} that deviates from 0 indicates that pairs of communities in the focal dimension are, on average, either more (positive values) or less (negative values) dissimilar than expected by chance. Our simulations (Supp. Information) also showed that the calculation of mean values of SES_{ij} across dimensions is not affected by unbalances in our sampling design.

The second approach assessed the relative importance of environmental sorting (i.e., deterministic mechanisms) and stochastic sampling effects (i.e., stochastic mechanisms) in each dimension by fitting the observed values of D_{space}, $D_{time,}$ and D_{SpXT} separately as a function of their respective environmental dissimilarities (i.e., Env_{diss}) and corresponding pairwise dissimilarities due to stochastic sampling estimated in $\overline{\mathbf{D}}_{\mathbf{null}}$. The spatiotemporal coordinates (i.e., the geographic coordinates and month) of samples were included as covariates in the models. These models were fitted using generalized dissimilarity models (GDMs), which, contrary to other matrix-based regressions (e.g., Mantel correlations), account for the non-linear relationship between community dissimilarity and ecological gradients when species turnover is high (Ferrier et al. 2007; Fitzpatrick et al. 2013). Variation partitioning (Borcard et al. 1992) was used to partition the null deviance explained by the fitted GDMs among each variable alone and their shared contribution. Given our scope (i.e., contrast the importance of deterministic and stochastic assembling mechanism across dimensions), we discuss only the proportion (%) of the total explained variation (i.e., relative importance) attributable to Env_{diss} and $\overline{\mathbf{D}}_{\mathbf{null}}$ alone and by their joint contribution across dimensions. Details regarding the proportion of explained variation accounted by all variables in the full model fitted for each taxonomic group can be found in Supp. Information. GDMs were fitted using the R pacakge *gdm* (Manion et al. 2018).

2.2.4 Contrasting patterns across taxonomic groups

We used the proposed framework to calculate for each of the ten taxonomic groups separately: (a) The average contribution of each dimension (BD_{space}, BD_{time}, and BD_{SpxT}) and beta-

diversity component (BD_{turn} and BD_{nest}) to BD_{total} ; (b) The mean SES_{ij} for each dimension; (c) The proportion of variation in D_{time}, D_{space}, D_{SpxT} attributable to the isolated and joint contributions of Env_{diss} and D_{null} in the fitted GDMs.

In this study, we aimed at identifying repeatable patterns across taxonomic groups concerning the following questions: (Q1) How is beta-diversity partitioned in space and time ? (Q2) How do the nestedness and turnover beta diversity components vary in space and time? (Q3) Do the relative importance of stochastic and deterministic assembly mechanisms significantly differ in space and time? To tackle these three questions, we used mixed-effects ANOVAs (summarized in Table III of Supp. Information II) where "Taxonomic Group" was considered as a random factor (intercepts). By doing so, we could control for hierarchical structures in our data caused by differences in life histories within and among taxonomic groups that are beyond the scope of this study. The fixed factors considered in the mixed-effects ANOVAs were called "Dimension" (levels: BD_{space}, $BD_{time,}$ and BD_{SpxT}), "Components" (levels: BD_{turn}, BD_{nest}), and "Variables" (levels: Env_{diss}, $\mathbf{\overline{D}_{null}}$, and their Joint contribution).

To answer Q1 and Q2, we fitted the average contribution to BD_{total} as a function of Dimensions and Components. These models determined which dimension contributed the most to BD_{total}, and whether nestedness and turnover components varied significantly among dimensions (i.e., a significant interaction between factors Dimensions and Components). To answer Q3, we fitted two different models. In the first, we fitted SES_{ij} as a function of Dimensions. We also used one-sample t-tests to evaluate whether the mean SES_{ij} values considering all taxonomic groups combined across dimensions were significantly different from zero. By doing so, we inferred whether beta-diversity in each dimension was significantly different from the null expectation. In the second model, we assessed whether the proportion (%) of the total variance explained by fitted GDMs across taxonomic groups changed as a function of Dimension and Variables. Taken together, these models informed whether the relative importance of assembly mechanisms differed among dimensions (e.g., significant interaction between Dimension and Variables). When necessary, we rank transformed the response variables to meet assumptions regarding residual normality and homoscedasticity. The mixed models were fitted using the R package *nlme* (Pinheiro et al. 2019).

2.3 Results

(Q1) How is beta-diversity partitioned in space and time? We observed that community composition varied more in space than in time consistently across taxonomic groups. More specifically, we found that the average contribution of BD_{SPxT} and BD_{space} to BD_{total} did not differ significantly but was higher than the average contribution of BD_{time} (Figure 2.2, df=2/18, F-value= 23.56, p-value < 0.001, Supp. Information).

(Q2) How do the nestedness and turnover beta diversity components vary in space and time? Our conceptual model predicts an increase in the importance of nestedness in temporal beta-diversity due to the persistence of core species through time along with the stochastic loss of

occasional species. In contrast, the replacement of core species due to species-environment sorting should increase the importance of turnover in spatial beta-diversity. As predicted, our results show that although turnover was the main component of beta-diversity across all three dimensions, nestedness was relatively (and significantly) higher in time than for the other components (Figure 2.2, interaction between Dimensions and Components, df=2/45, F-value=30.58, p-value < 0.001, Supp. Information).

(Q3) Do the relative importance of stochastic and deterministic assembly mechanisms significantly differ in space and time? Our conceptual model predicts that spatial beta-diversity will be mainly driven by deterministic species-environment sorting, while temporal beta-diversity results from stochastic sampling effects. Considering all taxonomic groups combined, the mean SES_{ij} was significantly lower in BD_{time} (Figure 2.3, df=2/18, F-value 21.84, p < 0.0001, Supp. Information II, Table III) and not significantly different from zero (one-sample t-test, df=9, t=0.981, p=0.345). When contrasting the results of the variation partitioning approach across taxonomic groups, we observed that BD_{time} was mainly explained by $\overline{D}_{null}$ (ca. 32% on average), though the differences between the variance explained by $\overline{D}_{null}$, Env_{diss}, and their Joint contribution were not significant (post hoc Tukey test, Env_{diss} vs. D_{null} p =0.92, Env_{diss} vs. Joint p= 0.73, D_{null} vs. Joint p= 0.64, Figure 2.4). These results corroborate our prediction that temporal beta-diversity is strongly driven by the replacement of occasional species caused by stochastic events of colonization and extinction. In contrast, BD_{space} was significantly more explained by Env_{diss} than by $\overline{D}_{null}$ and their Joint contribution (Figure 2.4). These results conform with our prediction that beta-diversity in space is driven by the turnover of core species associated with species sorting dynamics. Finally, variation in D_{SpxT} was equally explained by the isolated contribution of $\overline{D}_{null}$ and Env_{diss}. These findings therefore suggest that spatiotemporal changes in community composition are simultaneously driven by deterministic mechanisms in space and stochastic mechanisms in time.

Figure 2.2: Panel A summarizes (i.e., bar heights = mean, whiskers = Standard error of the mean) the average contribution of spatial (BD_{space}), temporal (BD_{time}) and spatiotemporal (BD_{SpxT}) beta-diversity to BD_{total} observed across the 10 taxonomic groups (seen in Panel B). Across taxa, the average contribution of BD_{time} to BD_{total} is significantly lower than the contributions of BD_{space} and BD_{SpxT}. Additionally, BD_{time} is significantly more nested than BD_{space} and BD_{SpxT}. Panel B, first row (top): Phengodidae (glow-worms), Lampyridae (fireflies), Carabidae (ground beetles). Second row: Eumolpinae (Leaf beetles), Antrhibidae (fungus weevils), and Cerambycidae (Longhorn beetles). Third row: Metopiinae, Pimplinae, Meteorus. Fourth row: Mesosotoinae.

2.4 Discussion

In this study, we investigated how compositional dissimilarities among communities (i.e., beta-diversity) are partitioned in space and time, and what is the dominant mechanism underlying compositional changes in each dimension. In the studied tropical slope, where environmental gradients are steeper in space than in time, spatial ($BD_{spatial}$) and spatiotemporal (BD_{SpxT}) compositional dissimilarities contributed more to the total variance of the species-by-site-by-time matrix (BD_{total}) than temporal changes alone (BD_{time}) (**Q1,** Figure 2.2). We also observed that while beta-diversity across all dimensions was mainly represented by species turnover, nestedness was significantly higher in time (**Q2**, Figure 2.2). These results were consistent across taxonomic groups and supported the predictions derived from the proposed conceptual model. The relatively low temporal beta-diversity observed is likely to result from the fact that core species, by definition, can persist locally when local environmental conditions do not change substantially over time. As such, temporal beta-diversity is consistently low and significantly more nested across taxonomic groups because the occurrence of occasional species fluctuates over time, but the occurrence of core species remains relatively constant (Snell Taylor et al. 2018). The higher contribution of spatial beta-diversity to total beta-diversity should be then due to the turnover of

core species associated with spatial variation in environmental conditions (Coyle et al. 2013). Finally, we observed that spatiotemporal changes in community composition contributed the most to the total variation of the species occurrence matrix, albeit this contribution was not significantly higher than the contribution of spatial beta-diversity alone (Figure 2.2). This pattern indicates that when comparing the composition of two sites in different moments in time, beta-diversity is high because it represents the turnover of both types of species: spatial changes in the composition of core species coupled with temporal fluctuation in the composition of occasional species within communities.
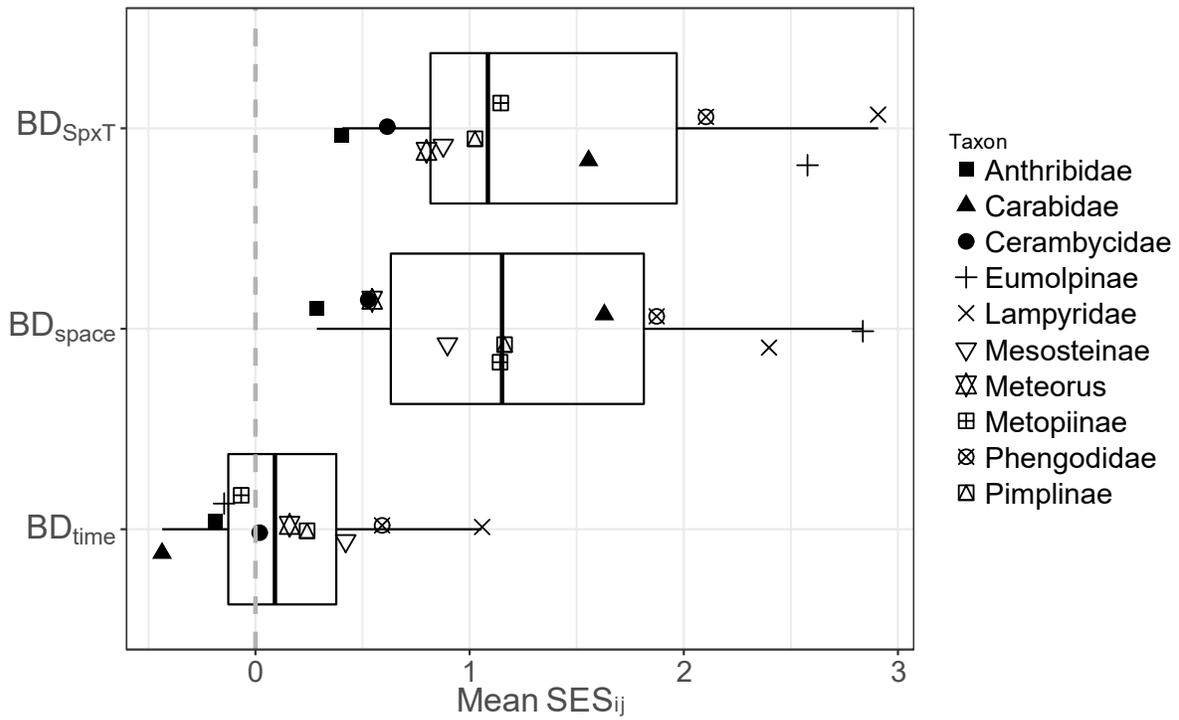
Figure 2.3: Considering all taxonomic groups combined (n = 10) , dissimilarities in temporal beta-diversity (BD_{time}) are, on average, not different from null model expectation (mean standardized effect size -mean SES_{ij}- overlaps with dashed line in 0). Conversely, spatial (BD_{space}) and spatiotemporal (BD_{SpxT}) beta-diversities were higher than the null expectation (i.e., mean $SES_{ij} > 0$ in both cases) and significantly higher than in BD_{time}. A jitter effect was used to place taxonomic groups according to their mean SES_{ij} values across dimensions.

Our results conform with our prediction that community assembly in tropical mountains is represented by spatially deterministic but temporally stochastic changes in community composition (**Q3**, Figures 2.3 and 2.4). The high importance of species-sorting mechanisms in determining the distribution of species along tropical slopes is a ubiquitous pattern in nature (Jankowski et al. 2009). Indeed, many studies discuss this pattern through the lens of Janzen's seasonality hypothesis (1967) which proposes that the low seasonality of tropical latitudes would favor the origin (i.e., speciation) and persistence of species with narrow environmental tolerances in the regional species pool (e.g., McCain 2009, Zuloaga and Kerr 2017). As such, local

communities at different portions of tropical elevational gradients would be only composed of elevational specialists (i.e., core species) that are in equilibrium with local habitat conditions. This rather "Clementsian" perspective over the nature of communities is often invoked to explain the high importance of environmental heterogeneity in the assembly of communities along tropical elevational gradients (Presley et al. 2012; Willig and Presley 2016).

However, a Clementsian perspective implicitly assumed under Janzen's hypothesis falls short on explaining the temporal beta-diversity patterns observed here and in other studies (e.g., da Silva et al. 2018, Wardhaugh et al. 2018). Suppose communities in tropical slopes were only composed of elevational specialists (i.e., core species). In this case, the low temporal heterogeneity of environmental conditions would allow all species found locally to maintain stable populations through time. Consequently, temporal changes in community composition should be negligible (BD_{time} closer to zero, Snell Taylor et al. 2018) in the absence of disturbance events that promote historical contingences in the assembly process (*sensu* Fukami 2015). Instead, our empirical data indicates that the average contribution of temporal beta-diversity to the total variance of species matrices is significantly different from zero (Figure 2.2) and is strongly explained by the stochastic allocation of individuals from the regional species pools into local communities (Figures 2.3 and 2.4). We also observed a high relative importance of the Joint component of the variation partitioning procedure in explaining temporal beta-diversity. The underlying cause is likely due to the fact that unsuitable local conditions keep populations of occasional species at low abundances and, therefore, more prone to local extinction through time via demographic stochasticity (Siqueira et al. 2020). Together, these results suggest that the composition of communities in tropical mountains is temporally dynamic because of random colonization-extinctions events that promote the turnover of species that are not in equilibrium with local habitat conditions (i.e., occasional species). Therefore, we argue that the capacity of Janzen's hypothesis in predicting temporal diversity patterns in tropical slopes can be improved once we reframe its framework to accommodate a "Gleasonian" perspective based on the existence of core and occasional species within communities.

Here we proposed a conceptual model that associates the relative importance of deterministic species-environment sorting with the turnover of core species across communities. However, it is important to highlight that this association is likely to be stronger at fine spatiotemporal scales such as the ones encompassed by our empirical data (i.e., a relatively short elevational transect over a year). This is because, at broader scales, dispersal limitation may also shape the distribution of core species in space. For instance, recent studies investigating the mechanisms driving variation in beetle communities along broad spatial scales concluded that dispersal limitation, and not deterministic habitat selection, explains the absence of many taxa in northern communities since the Last Glacial Maximum (Gómez-Rodríguez and Baselga 2018). These results suggest that high-latitude biomes may harbor communities depauperate in core species because dispersal limitation may keep them from recolonizing habitats with suitable contemporary climate.

Nevertheless, our conceptual model fosters two important insights into the reasons why the relative importance of deterministic and stochastic mechanisms within metacommunities may differ in space and time. First, it provides strong evidence that there are clear limits to how much we can learn from metacommunities based on unidimensional assessments alone. This is because
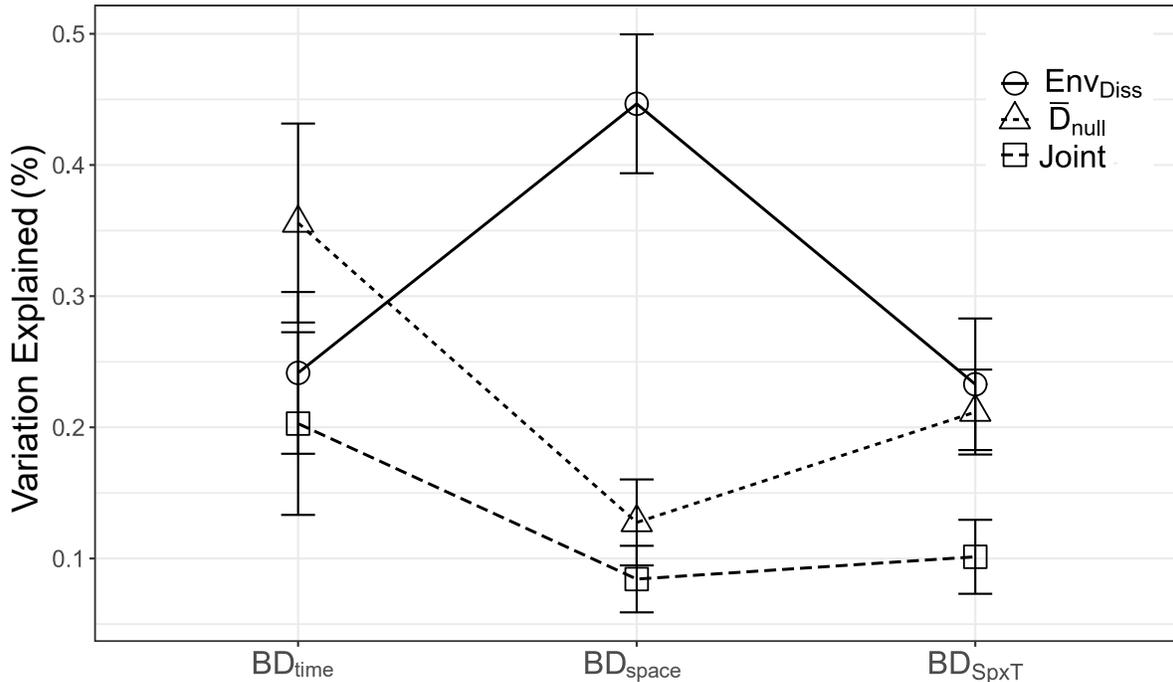
Figure 2.4: Interaction plot showing the percentage of the total variation in the spatial (BD$_{space}$), temporal (BD$_{time}$) and spatiotemporal (BD$_{SpxT}$) dimensions of the compositional dissimilarity matrix explained by environmental dissimilarities (Env$_{diss}$), stochastic sampling effects ($\overline{D}_{null}$) and their joint contribution (Joint). Symbols show mean values obtained considering all 10 taxonomic groups. Whiskers = Standard Error of the mean.

if different mechanisms structure metacommunities in space and time, sorting metacommunities into archetypes (or paradigms *sensu* Leibold et al. 2004) based only on spatial dynamics can only yield an incomplete understanding of the assembly process (e.g., Wisnoski et al. 2019, Jabot et al. 2020). Second, our model and empirical results indicate that well-informed predictions on the mechanisms driving community assembly should consider the complex spatiotemporal structure of environmental conditions in landscapes (e.g., Bar-Massada et al. 2014). For instance, in temperate mountains, an ecosystem where environmental gradients are steeper in time than in space, community assembly should be a product of the changes in composition that are stochastic in space but deterministic in time (i.e., opposite to those in tropical mountains). If corroborated, this pattern may indicate a latitudinal gradient in the relative importance of deterministic and stochastic mechanisms underpinned by differences in the spatiotemporal structure of environmental conditions under which lineages have evolved (Kozak and Wiens 2010; Cadena et al. 2011). Future research that replicates our framework across mountain ranges (or any other

system where environmental variation differs in time and) in different latitudes and biogeographic regions may test these additional predictions, thus providing a broader understanding of the assembly process in space and time.

2.5 Supplementary Information

2.5.1 The existence of core, satellite, and occasional species in metacommunities

Hanski (1982) proposed that species in a regional pool could be divided into two types, core, and satellite, depending on their frequency (spatial occupancy) and abundance within communities. In Hanski's framework, core species were those with high regional occupancy and high abundance within local patches, while satellite species were those with low occupancy and rare. Magurran and Henderson (2003) proposed a temporal analog of Hanski's spatially oriented classification by showing that locally abundant species are often those with higher temporal occupancy within local communities. In this temporally oriented classification, core species (also named "frequent," sensu Ulrich and Ollik 2004) are those that can maintain self-sustaining local populations for relatively long periods of time. Conversely, species that are poorly suited to local habitat conditions and, consequently, are unable to keep viable local populations in the long term have been named in different ways: occasional, accidental, vagrant, tourist, or transient (see Snell Taylor et al. 2018 and references within). These names refer to the probabilistic (i.e., stochastic) occurrence of this second type of species (hereafter referred to as occasional) in a given community at a specific moment in time. An important distinction between the spatially and temporally oriented classifications is that while the former is species-specific (i.e., a species is either core or satellite), the latter is community-specific (i.e., a species can be occasional in one community but core in another). As such, satellite species are not necessarily the same as occasional species (Umaña et al. 2017). For instance, species with narrow spatial distributions (i.e., satellite) may be a persistent component of local communities over time (i.e., core), while species widely distributed in the landscape may be occasional members of local communities (occasional). Even though both types of classifications are likely to be somewhat correlated, here we only considered the site-specific classification (i.e., core-occasional) in our proposed conceptual model.

The core-occasional species framework may help us broaden our understanding of the assembly process because: (i) it has clear predictions on the mechanisms driving the diversity and distribution of core and occasional species in metacommunities (see main text). These predictions have found empirical support across many different ecosystems and taxa (e.g., Belmaker 2009, Supp et al. 2015, Umaña et al. 2017); (ii) recent meta-analyses showed that the removal of occasional species from ecological analyses affects well-known patterns such as species-area relationships, species-energy relationships, and species abundance distributions (Magurran and Henderson 2003, Snell Taylor et al. 2018).

However, identifying the core and occasional members of empirical communities is not straightforward. Over the past few decades, several studies have proposed different ways to do so

by using detailed information on species life-history (Belmaker 2009, Supp et al. 2015) and traits (Ulrich and Zalewski 2006, Umaña et al. 2017). When such data is not available, identifying core and occasional species within communities is based on somewhat arbitrary decisions. For instance, many studies use thresholds to classify species based on their temporal occupancy in a given site (e.g., core species are those that occur in the focal community over more than X% of the total years surveyed). In addition, recent simulations have shown that the probability of misclassifying species based on their temporal occupancy is particularly high in observational studies carried out in highly heterogeneous landscapes at relatively fine temporal scales. This is because mass-effect dynamics can allow species that fail to maintain positive population growth rates to persist locally through time due to repeated immigration of individuals from neighboring communities (Snell Taylor et al. 2020). Given the lack of knowledge about the distribution and taxonomy of tropical insects (i.e., the Wallacean and Linnean shortfalls, respectively), and the relatively fine temporal scale of our observational data (i.e., one year long), we decided to use simulated rather than observational data to validate the assumptions of our conceptual framework (see main text).

2.5.2 Replicating in silico the proposed conceptual model and validating its assumptions
In this model, population size of species *j* in site *s* at time *t* is given by:

$$N_{js,t} = WPD_{js,t} - E_{js,t} + I_{jx-s,t} \text{ eq. (SI 2.1)}$$

where $WPD_{js,t}$ represents the number of individuals after within-patch dynamics, $E_{js,t}$ represents the number of individuals emigrating from *s*, and $I_{jx,t}$ is the number of individuals immigrating from site x to site s in time t. $WPD_{js,t}$ was modeled using a modified form of the Beverton-Holt competition model (Beverton and Holt 1957, Tucker et al. 2016) given by:

$$WPD_{js,t} = Poisson(N_{js,t-1} \, GR_{js,t}) \text{ eq. (SI 2.2)}$$

$N_{js,t-1}$ is the population size in t-1 while $GR_{Js,t}$ represents the growth rate of *j* when conditioned by within-patch abiotic and biotic conditions, and is calculated as:

$$GR_{js,t} = R_{js,t} \frac{1}{1 + \alpha_{jj} \sum N_{js,t} + \alpha_{kj} \sum N_{ks,t}} \text{ eq. (SI 2.3)}$$

where R_{jst} is the effect of environmental selection in population dynamics and is given by:

$$R_{js,t} = R_{max} \exp\left(\frac{-(Env_{s,t} - \mu_j)^2}{2 \, \sigma_j^2}\right) \quad \text{eq. (SI 2.4)}$$

where R_{max} represents the maximum intrinsic growth rate. $Env_{s,t}$ is the local abiotic conditions, μ_j is the environmental optimum of species *j* and σ_j is its niche breadth (i.e., environmental tolerance). The right term of equation SI. 2.4 represents species-specific Gaussian responses to environmental variation and range between 0 (unsuitable environmental conditions) and 1 (perfectly suitable environmental conditions).

The term $1/1 + \alpha_{jj} \sum N_{js,t} + \alpha_{kj} \sum N_{ks,t}$ of equation SI. 2.3 represents the effects of intra and interspecific competition experienced by species j in community s. The coefficient α_{jj} is the per capita effect of species j on itself while α_{kj} is the per capita effect of species k on the growth of j. In this model, we assumed a stabilizing competition, i.e., species compete more strongly with themselves than with each other (i.e., $\alpha_{jj} > \alpha_{kj}$). Note that stabilizing competition increases the chances of dominant and rare species to coexist in a community over time (Thompson et al. 2020). When considering equation SI. 2.3, suboptimal growth ($GR_{js,t} < 1$) results from the negative impact of local competition on population growth and/or a mismatch between local environmental conditions and species environmental preferences. Finally, stochasticity in birth and survival was incorporated by drawing the number of individuals of species j in community s in time t from a Poisson distribution whose mean is equal to the deterministic part of equation SI. 2.3 (i.e., terms within parenthesis) (Shoemaker and Melbourne 2016; Thompson et al. 2020)

The argument $E_{js,t}$ of equation 2 represents the number of individuals of species j that emigrated from patch s at time t. This number was determined by the successful number of $WPD_{js,t}$ draws from a binomial distribution with different probabilities *(p)* depending on whether local conditions were suitable (p_{suit}) or not (p_{unsuit}). That is, if $GR \geq 1$ then $p = p_{suit}$, but if $GR < 1$ then $p = p_{unsuit,}$. Here we defined that $p_{suit} < p_{unsuit}$ so individuals are more likely to remain if habitat conditions are benign, but more likely to escape if habitat conditions are unsuitable.

The argument $I_{jx-s,t}$ in equation SI. 2.1 represents the number of individuals of species j that emigrated from site x and arrived in community s at time t. All individuals that emigrated across communities have an equal probability of immigrating to all patches but the one from which they emigrated. The destination of immigrants was computed as random sampling with replacement and equal probabilities repeated. To incorporate the costs of migration in our simulations, a number of immigrants is selected to die during the process. This number is given by a binomial distribution whose number of trials is equal to $E_{jx,t}$, and the success probability is 0.1.

Population dynamics were carried across 15 different sites over 1200 time steps (each time step representing one month, i.e., 100 years in total). To ensure that analyses were performed on stable rather than transient communities, only the last 12 time steps (i.e., the last year) were considered in the final species-by-site-by-time matrix. As such, the final species matrix has similar dimensions to those of our empirical data. In t=1, all 30 species were set with the same initial regional abundance (ca., 100), and were placed across the 15 modeled patches according to their environmental preferences. Environmental conditions in the landscape were manually set to depict a similar environmental profile as in our empirical data (compare Figure 2.1 and Figure SI 2.I). That is to say that the simulated landscape mimicked the high elevational but low seasonal variation in environmental conditions often observed along a tropical elevational gradient (see Figure 2.1 in the main text). Environment is represented by a single variable that ranges from 0 to 1. Environmental optimum μ of each species was equally spaced and ranged from 0 to 1, while

species niche breadth σ was set at 1 for all species. This value for σ decreased regional extinctions by ensuring that all species were able to persist over time in at least one of the 15 sites modeled. In addition, the immigration pool was rescaled to 500 individuals in every timestep to minimize temporal fluctuations in the size of the regional pool caused by regional extinctions. Core species in a given community had an average *GR* over the last 120 time steps higher equal or greater than one, whereas occasional species were those with average *GR* smaller than one.

We repeated the simulation 50 times. At the end of each simulation round, we applied our partitioning framework described (see Methods) to analyze the final species-by-site-by-time matrices (i.e., full matrix). We also repeated our analyses after removing each type of species from the final species matrices. This is the standard protocol of studies that aim at understanding the mechanisms governing the diversity of core and occasional species in a metacommunity (e.g., Belmaker 2009, Supp et al. 2015, Umaña et al. 2017). The analyses performed considering the full matrix yielded results that were qualitatively similar to the results observed when analyzing the observational data (Compare panel "Full Matrix" in the SI figures below and figures in the main text). The goal of our model was to demonstrate a link between the importance of deterministic and stochastic mechanisms and the turnover of different types of species among communities. Indeed, when analyzing matrices in which either core or occasional species were removed, we showed that deterministic species-environment sorting mechanisms mainly explained the spatiotemporal turnover of core species. In contrast, the turnover of occasional species was better predicted by the stochastic allocation of species from the regional pool into local communities (Figure SI. II, III and Figure Box I in main text).

The values of all parameters considered here are listed in Table SI 2.I. The chosen value for each parameter was set in a way that the following conditions were met:

> 1-At least 50% of the species in the regional pool remained in the metacommunity in the last 12 time steps;
>
> 2-All sites had at least one core member;
>
> 3-Core and occasional species had similar regional abundances in the last 12 steps.

These conditions were critical to ensure that our analyses based on the final matrix was able to detect the signal of both types of species in the metacommunity. It is important to highlight that these simulations were used to validate our conceptual framework's assumptions and should not be interpreted as an attempt to estimate the parameters of any real metacommunity. As such, while the results obtained here were robust to changes in the size of the metacommunity (number of species and their initial abundances), changes in other simulation parameters' values may change the qualitative patterns observed. Different combinations of values may also allow the final species matrix to meet the conditions listed above but exploring the different results obtained along the parameter space is beyond the scope of this simple simulation framework.
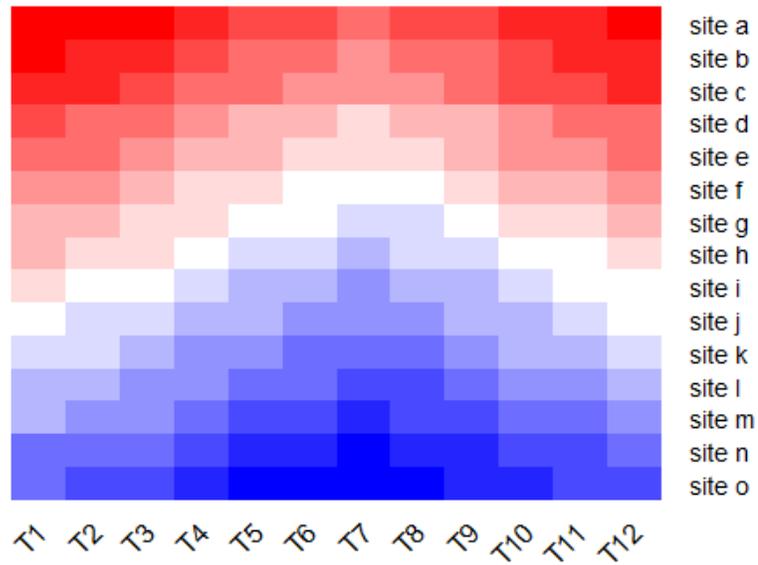
Figure SI. 2.I Virtual landscape in which metacommunity models were set to run (only last 12 time steps are shown). The spatiotemporal structure of environmental conditions is similar to the one observed in the tropical slope where insect communities were sampled (see Fig. 2.1 in the main text)

Table SI 2.I: Simulation parameters: Values and underlying distribution of related process

Parameter	Description	Min	Max	mean	Value	Distribution
Nsp	Number of species	-	-	-	30	-
Nt1	Regional Abundances T1	-	-	-	2000	-
Rmax	Max growth rate in optimum conditions	-	-	-	1.45	-
N_patch	Number of patches	-	-	-	15	-
N_month	Number of time steps	-	-	-	1200	-
E	Environment in a patch at specific moment in time	0	1	0.5	-	-
α_{jj}	Coeff. Intra-specific Competition	-	-	-	1/200	-
α_{kj}	Coeff. Inter-specific Competition	-	-	-	1/800	-
m	Mortality rate	-	-	-	0.1	binomial
p_{unsuit}	Chance to disperse when GR<1	-	-	-	0.8	binomial
p_{suit}	Chance to disperse when GR>1	-	-	-	0.3	binomial
μ	Niche optima	0	1	0.5	-	-
σ	Niche breadth	-	-	-	1	-

Figure SI. 2.II: Average contribution of spatial (BD$_{space}$), temporal (BD$_{time}$), and spatiotemporal (BD$_{SpxT}$) changes in community composition to the total variance of species-by-site-by-time matrix (BD$_{total}$). In the first panel analyses were carried out considering both types of species across communities (Full matrix) and results were qualitatively similar to the results observed across insect communities sampled in this study (see Fig. 2.2 main text). The other two panels show the results when only occasional (middle) or core species (right) were kept within communities. As we can see, the composition of core species varies relatively little over time, while the composition of occasional species varies in all dimensions.

Figure SI. 2.III: Mean Standardized effect sizes (SES) observed across dimensions calculated considering full communities and after the removal of either core or occasional species (see caption of Figure SI 2. II). Patterns observed in the full matrix were qualitatively similar to the patterns observed in analyses considering real communities (see Fig. 2.3). When considering only occasional species (middle), there is a higher similarity between the null expectation and the final simulated matrix (SES closer to zero). This result indicates that stochastic events of colonization extinction are the main drivers of beta-diversity when only occasional species are taken into consideration in our analyses. Conversely, beta-diversity was very different from the null expectation (i.e., above 0 in BD_{space} and BD_{SpxT}, but below zero 0 in BD_{time}), which indicates that other mechanisms drive the spatiotemporal distribution of this type of species across communities

2.5.3 Empirical data and analyses

Table SI 2.II: Characterization of taxonomic groups

Family (Subfamily/Genus)	Order	Common name (English)	Occurrence	Global Richness	Sampled Richness	Sampled Abundance	Functional role
Lampyridae	Coleoptera	Fireflies	Global	ca.2000	62	1042	Specialist Predators of worms, gastropods and other fireflies
Phengodidae	Coleoptera	Glow-worm Beetles	New World	ca.270	31	912	Predator of Millipedes and other Insects living in the litter
Carabidae	Coleoptera	Ground Beetles	Global	ca. 40000	55	517	Generalist Predators
Eumolpinae	Coleoptera	Leaf Beetles	Global	ca. 7000	67	1237	Specialist Herbivores
Cerambycidae	Coleoptera	Longhorn Beetles	Global	ca. 24000	59	344	Generalist Herbivores
Anthribidae	Coleoptera	Fungus weevils	Global	ca. 3900	63	214	Feed on Fungus and decaying matter
Ichneumonidae (Metopiinae)	Hymenoptera	-	Global	ca. 830	67	782	Endoparasitoids of Lepidoptera
Ichneumonidae (Pimplinae)	Hymenoptera	-	Global	ca. 1740	97	1514	Parasitoids of Lepidoptera and Spiders
Braconidae (Mesostoinae)	Hymenoptera	-	Neotropical and Australasian	ca. 32	32	317	Ectoparasitoids of Lepdoptera
Braconidae (Meteorini/Meteorus)	Hymenoptera	-	Global	ca. 250	22	316	Endoparasitoids of Lepidoptera
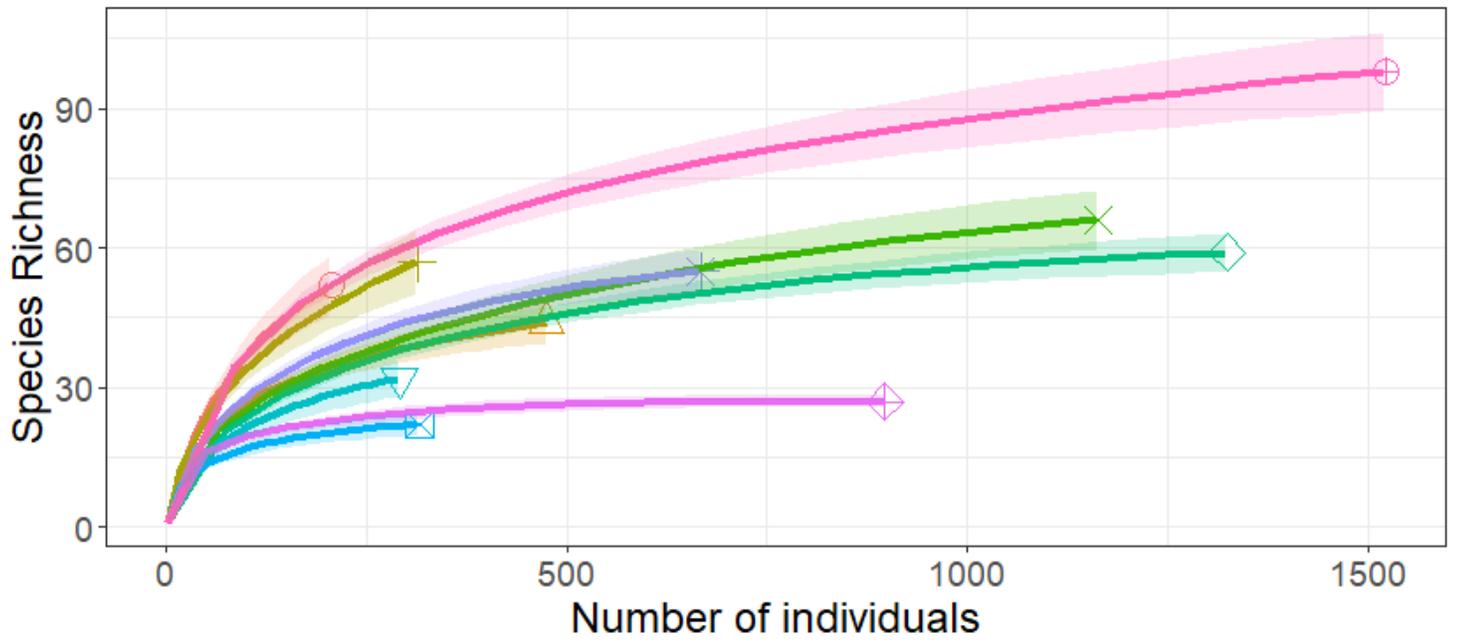
Figure SI. 2.IV; Individual based rarefaction per taxonomic group

Table SI 2.III: Variation Partitioning across taxonomic groups. Values represent the % of the total variation explained in fitted GDMs accounted by Env_{diss} alone, D_{null} alone, their joint contribution (Joint). The values assigned in the column Coord represent the % explained by Coord (i.e. elevation per month) alone + the joint contribution of Coord and Env_{diss} + the joint contribution of Coord and D_{null} + the joint contribution of Coord, Env_{diss} and D_{null}. * Negative values were considered equal to 0 because they represent scenarios where the independent variables explain less variation than a random variable following a normal distribution

Taxa	Dimension	% Env_{Diss}	% D_{null}	% Joint	% Coord
Phengodidae	D_{time}	23.5	23.3	14	39
	D_{space}	39.6	2.3	4.8	53
	D_{SpxT}	23.3	14.5	11	53
Lampyridae	D_{time}	46.6	20.0	23.3	9.6
	D_{space}	60.7	11.3	8.2	19.6
	D_{SpxT}	59.7	13.3	15.4	11.4
Carabidae	D_{time}	27.1	15.4	11.5	45.9
	D_{space}	57.3	7.7	11.6	23.2
	D_{SpxT}	19.2	22.2	5.2	53.6
Pimplinae	D_{time}	4.5	65.8	21.6	7.8
	D_{space}	16.1	24.3	3	56.4
	D_{SpxT}	11.9	37.9	7.8	42.2
Eumolpinae	D_{time}	8.5	27.3	13.1	50.9
	D_{space}	44.4	2.1	1.8	51.5
	D_{SpxT}	12.1	6.4	*0	83.2
Metopiinae	D_{time}	3.4	66.8	19.4	10.1
	D_{space}	36	22.1	16	25.7
	D_{SpxT}	29.3	25.8	20.4	24.2
Mesostoinae	D_{time}	58.7	38.9	*0	11.5
	D_{space}	21.8	29.9	*0	49
	D_{SpxT}	8.2	34	*0	59.4
Antrhibidae	D_{time}	10.3	20.8	25.2	43.5
	D_{space}	65.9	4.1	27.5	2.3
	D_{SpxT}	41.1	9.1	26.1	23.5
Cerambycidae	D_{time}	57.2	18.5	19.4	4.7
	D_{space}	48.5	3.5	10.2	37.6
	D_{SpxT}	13.7	29.1	10.5	46.4
Meteorus	D_{time}	26.5	17.8	19.3	4.8
	D_{space}	56.1	3.7	10	37.6
	D_{SpxT}	13.48	28.6	10.1	46.8

Table SI 2.IV: Results of mixed models ANOVAs. In all models, "Taxonomic Group" was set as a random factor. DV= Dependent Variable; Levels of the fixed factor (FF) "Dimensions" = BD_{Space}, BD_{Time}, and BD_{SpxT}; "Components" = BD_{nest} and BD_{turn}; "Variables"= Env_{diss}, D_{null}, and Joint. * Significant at alpha = 0.05

Question	DV	FF	Dfnum/DFden	F-statistic	p-value	Figures
Q1	Avg.Cont to BDtotal	Dimensions	2/18	23.56	<0.0001*	
Q2	Avg.Cont to BDtotal	Dimensions	2/45	0.86	0.46	2.2
		Components	1/45	56.67	<0.0001*	
		Dimensions: Components	2/45	30.58	<0.0001*	
	SES_{ij}	Dimensions	2/18	21.84	<0.0001*	2.3
Q3	Explained variation (%)	Dimensions	2/72	1.765	0.17	
		Variables	2/72	14.71	<0.0001*	2.4
		Dimension: Variables	4/72	4.95	0.0014*	

2.5.4 Exploring the effects of unbalances in the dissimilarity matrices

We make use of simple simulations to demonstrate that averaging effectively controls for differences among BD_{space}, BD_{time}, and BD_{SpxT} that result solely from unbalances in the dissimilarity matrix **D**. We started by simulating 150 regional pools using three different species abundance distributions (Log-normal, Poisson-lognormal, and negative binomial distributions, 50 regional pools each). Each regional pool was composed of 1000 individuals distributed across 50 species. Each individual within any given pool was then randomly allocated across 90 samples with specific spatial and temporal coordinates. Then we calculated the pairwise dissimilarity among simulated communities and obtained the matrix $\mathbf{D_{sim}}$. Note that the number of entries in $\mathbf{D_{sim}}$ representing D_{time}, D_{space}, and D_{SpXT} was set to be, respectively, 225, 630, and 3150; the same unbalance observed in our real **D** matrices. Since the composition of simulated communities is generated by chance alone, all dimensions should equally contribute to BD_{total}. However, without proper correction, we observe that $BD_{SpxT} > BD_{space} > BD_{time}$; a pattern that is caused by the observed unbalance in $\mathbf{D_{sim}}$ (Figure SI. 2. V). Only after dividing BD_{space}, BD_{time}, and BD_{SpxT} by the number of entries in $\mathbf{D_{sim}}$ representing D_{space}, D_{time}, and D_{SpxT}, we observe that all dimensions, on average, contribute equally to BD_{total} (Figure SI. 2. VI). Based on the same simulations, we also show that the results of the SES procedure (see main text) were not influenced by unbalances in **D**. More specifically, in these simulations where species were randomly distributed across communities, the mean SES_{ij} for each dimension is not significantly different from 0 (Figure SI. 2.V).

Figure SI. 2.VI Partitioning BD_{total} into its dimensions (BD_{space}, BD_{time}, BD_{SpxT}) without accounting for unbalances in the dissimilarity matrix $\mathbf{D_{sim}}$. Even though our simulations were set in a way that community composition would change equally across dimensions ($BD_{space} = BD_{time} = BD_{SpxT}$), the contribution of each dimension to BD_{total} differed due to differences in the number of entries in $\mathbf{D_{sim}}$ representing beta-diversity in each dimension
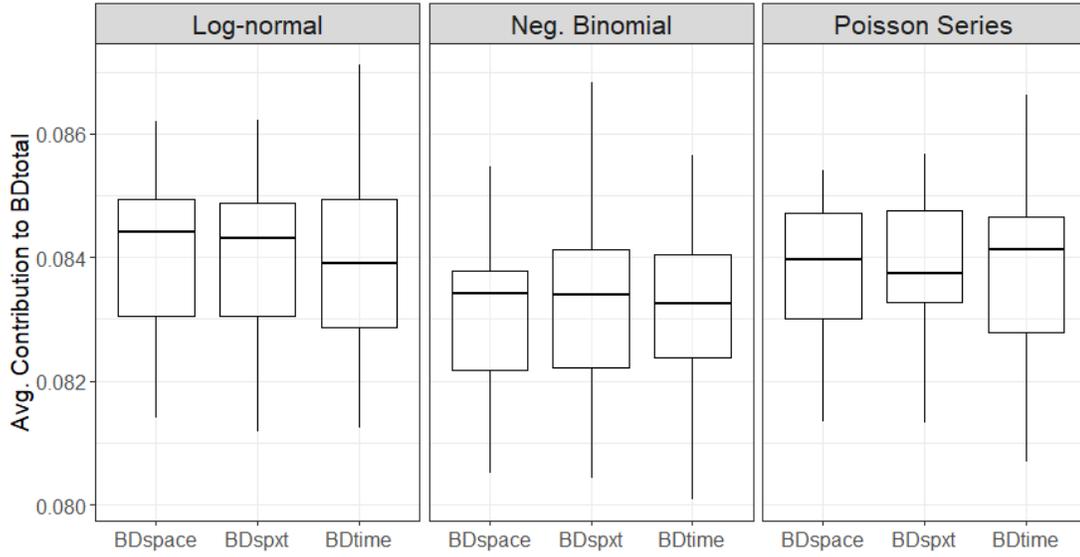
Figure SI. 2.VII: Same as in figure SI. 2.V but now we averaged the contribution of each dimension to BD_{total} by the number of entries in $\mathbf{D}_{sim}$ representing beta-diversity in each dimension. After this correction, the contribution of each dimension to BD_{total} was not significantly different.
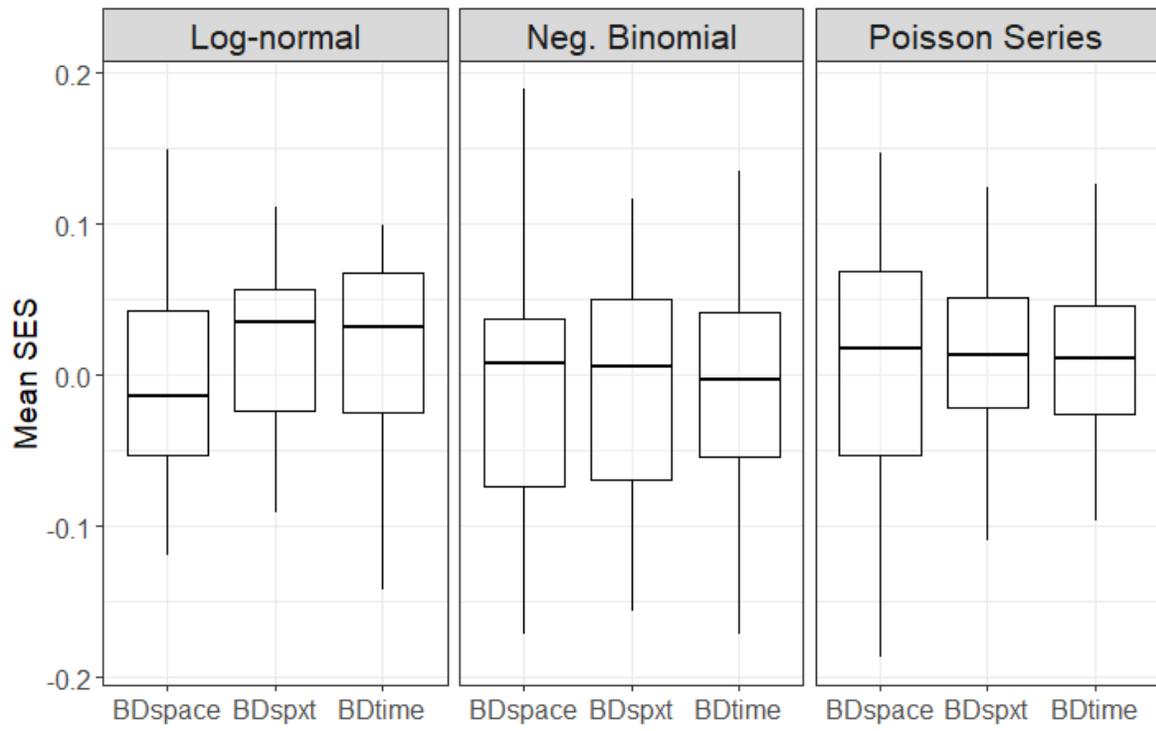
Figure SI. 2.VIII: Standardized Effect Sizes (SES) are similar across dimensions in metacommunities despite unbalances in D_{sim}. Mean pairwise SES values across dimensions revolve around zero because communities were set to represent random samples taken from the regional species pool

Chapter 3 : The geography of metacommunities: landscape characteristics drive geographic variation in the assembly process through selecting species pool attributes

3.1 Abstract

The non-random association between landscape characteristics and the dominant life-history strategies observed in species pools is a typical pattern in nature. Here, we argue that these associations determine predictable changes in the relative importance of assembly mechanisms along broad-scale geographic gradients (i.e., the geographic context of metacommunity dynamics). To demonstrate that, we employed simulation models in which groups of species with the same initial distribution of niche breadths and dispersal abilities interacted across a wide range of landscapes with contrasting characteristics. By assessing the traits of dominant species in the species pool in each landscape type, we determined how different landscape characteristics select for different life-history strategies at the metacommunity level. We analyzed the simulated data using the same analytical approaches used in the study of empirical metacommunities to derive predictions about the causal relationships between landscape characteristics, dominant life-histories in species pools, and their reciprocal influence on empirical inferences regarding the assembly process. We provide empirical support for these predictions by contrasting the assembly of moth metacommunities in a tropical versus a temperate mountainous landscape. Collectively, our model framework and empirical analyses demonstrate how the geographic context of metacommunities influences our understanding of community assembly across broad-scale ecological gradients.

3.2 Introduction

Community assembly theory studies the mechanisms by which species from a broader pool of potential colonizers assemble into local communities at finer scales (Hillerislambers et al. 2012). Metacommunity theory advances our understanding of a wide range of biodiversity patterns by extending community assembly theory to incorporate mechanisms such as dispersal limitation, environmental selection, and ecological drift (Mouquet and Loreau 2003; Vellend et al. 2014; Fournier et al. 2017; Koffel et al. 2022). Theoretical models have predominantly advanced our knowledge about the importance and links among these mechanisms by systematically manipulating parameters governing two distinct metacommunity components: (1) the attributes of species pools that form metacommunities (e.g., degree of species ecological specialization and dispersal ability); and (2) the characteristics of landscapes (e.g., environmental heterogeneity, connectivity) where metacommunity dynamics occur. For instance, by independently manipulating species pool attributes and the environmental heterogeneity and connectivity of the landscape, one can simulate assembly dynamics corresponding to distinct metacommunity archetypes (e.g., Thompson et al. 2020; Suzuki and Economo 2021). Such theoretical frameworks

produce insights into how distinct combinations of species pool attributes and landscape characteristics can generate the multitude of diversity patterns frequently observed in empirical metacommunities (e.g., Ovaskainen et al. 2019, Guzman et al. 2022).

However, the dominant life-history strategies observed in species pools forming metacommunities are selected by the characteristics of landscapes where the assembly process occurs (Büchi and Vuilleumier 2014; Fournier et al. 2020). Indeed, this is a fundamental idea in spatial ecology (e.g., Peres-Neto *et al.* 2012) that also underlies well-established ecogeographical rules and macroecological hypotheses. For instance, Janzen's seasonality hypothesis states that latitudinal variation in the degree of spatial and temporal variation in landscape environmental conditions explains latitudinal clines in the degree of ecological specialization of species in the regional pools (Janzen 1967; Ghalambor 2006; Sheldon et al. 2018). Similarly, Rapoport`s rule (i.e., the increase in species geographic ranges with latitude, Stevens 1989, Ruggiero and Werenkraut 2007) is assumed to be a consequence of the dominance of strong dispersers in temperate landscapes where temporal variability in habitat conditions is high. While the non-random association between species pool attributes and landscape characteristics is a common pattern in nature (e.g., Sunday et al. 2011; Sheard et al. 2020), we have yet to determine its influence on our empirical understanding of metacommunity patterns and the relative importance of underlying assembly mechanisms. This understanding should be particularly relevant for generating insights into why broad-scale empirical studies frequently report (bio)geographic variation in metacommunity dynamics (e.g., Qian and Ricklefs 2012; Myers et al. 2013; Henriques-Silva et al. 2015; Nishizawa et al. 2022).

In this study, we set out to determine how the dependence of species pool characteristics on landscape attributes influences the geographic context of metacommunity dynamics, i.e., how it drives predictable variation in the relative importance of mechanisms that assemble different metacommunities distributed along broad-scale ecological gradients, across biogeographic regions, or even at the global scale. Our conceptual framework can be described as a partial mediation model (Figure 3.1) in which landscape attributes (i.e., exogenous variables) determine the degree of specialization and dispersal ability of species that dominate species pools at the metacommunity scale (i.e., mediator variables). These two (model) compartments jointly dictate the relative importance of different assembly mechanisms (i.e., endogenous variable). Putting in ecological terms, landscape attributes that vary across large-scale gradients (e.g., seasonality) should determine large-scale geographic changes in the dominant traits and life-history strategies observed in species pools that form metacommunities (Peres-Neto et al. 2012, Henriques-Silva et al. 2015). As demonstrated in this study, these non-random associations between landscape characteristics and species pool attributes underpin geographic shifts in the relative importance of assembly mechanisms.

To provide theoretical validation and illustrate the utility of our conceptual framework, we built a process-based (simulation) metacommunity model wherein groups of species with the same initial distribution of continuous traits (here, ecological specialization and dispersal ability) were

allowed to colonize and reach coexistence in landscapes with varying of levels spatiotemporal environmental heterogeneity, physical connectivity, and spatial structure (autocorrelated) of environmental (habitat) conditions. By evaluating the degree of ecological specialization and dispersal ability of the species that could persist and dominate the metacommunity (defined here as the metacommunity species pools, *sensu* Fukami 2015) across various types of landscapes, we were able to understand how distinct landscape characteristics select for different dominant life-history strategies. This modeling approach allowed us to understand how geographic clines in life-history strategies, often attributed to broad-scale variation in evolutionary and historical mechanisms (e.g., trait evolution/conservatism and speciation, Hua 2016), can also arise from ecological dynamics operating at the finer spatial and temporal scales of metacommunity dynamics (Henriques-Silva et al. 2015, Mittelbach and Schemske 2015).

To determine how the interdependences between species pool and landscapes influence inferences about metacommunity dynamics, we analyzed the resulting (simulated) metacommunities employing analytical approaches commonly used to infer the relative importance of assembly mechanisms in empirical metacommunities (discussed in Methods and see relevant conceptual and statistical limitations in Gilbert and Bennett 2010, Peres-Neto and Legendre 2010). We then explored the causal links between landscape attributes, dominant life-history strategies in species pools, and related inferences about community assembly through statistical models.

To provide empirical support for some of the theoretical predictions derived from our simulation framework, we analyzed empirical data on moth metacommunities in a tropical and temperate mountainous landscape. Tropical and temperate mountains are known to exhibit distinct patterns of spatial and temporal environmental heterogeneity (Zuloaga and Kerr 2017), making them suitable "natural experiments" for testing our theoretical predictions. For instance, a strong prediction derived from our simulation models (see below) posits that in landscapes where environmental variation is stronger in space than in time (e.g., tropical mountains), environmental specialists will predominate in species pools. Hence, empirical studies in these areas are likely to conclude that species-environment sorting is the primary driver of spatiotemporal variation in community composition. On the other hand, in regions where environmental variation is stronger in time than in space (e.g., temperate mountains), generalists should dominate species pools. As a result, mechanisms other than environmental selection (e.g., dispersal limitation or autocorrelation of demographic processes) are expected to play a greater role in affecting spatiotemporal variation in community composition. By contrasting the predictions generated by our conceptual model with the outcomes observed in these moth metacommunities, we demonstrated how our conceptual framework can serve as an inferential tool for investigating the geography of metacommunity dynamics.
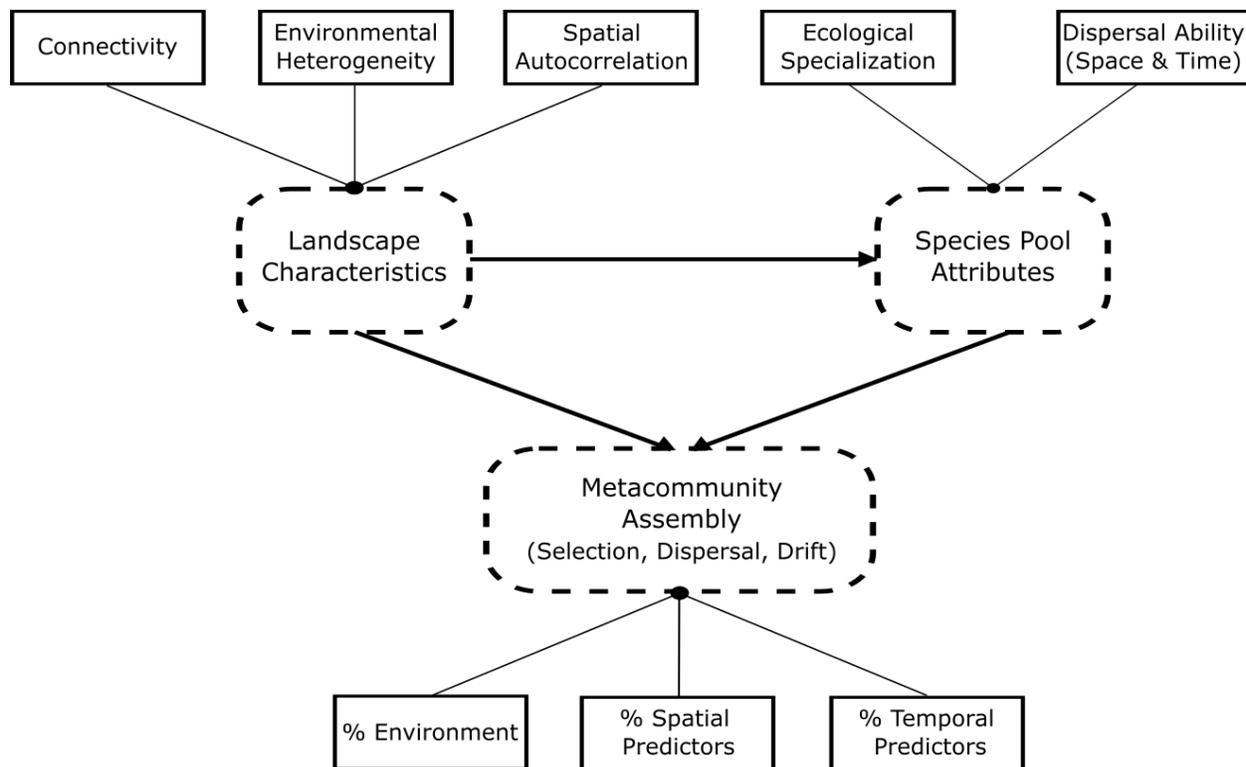
Figure 3.1: A mediation model for the geography of metacommunity assembly. It incorporates the effects of both landscape (exogenous variables) and species pool (mediator variables) attributes on the relative importance of selection, dispersal, and drift (i.e., Endogenous variable). Dashed round-edged boxes represent theoretical constructs, i.e., components of the metacommunity theory that are inferred from measurable variables and patterns observed in empirical metacommunities (solid rectangles). "%" represents the amount of variation in community composition explained by environmental variables and spatial and temporal predictors. The variation explained by their covariation (i.e., joint contribution) is omitted.

3.3 Methods

3.3.1 Simulated landscapes

For the sake of brevity, we only briefly describe how we simulated landscapes here. An extended description is found in Supp. Information. We generated a total of 216 types of landscapes considering a wide range of spatiotemporal heterogeneity levels (8), physical connectivity (9 levels), and spatial distribution of environmental conditions (3 levels) (Figure 3.2). These landscape attributes have been shown to modulate the mechanisms underlying species coexistence, which, in turn, influence metacommunity dynamics (Büchi et al. 2009; Moritz et al. 2013; Fournier et al. 2017).

We randomly distributed 60 habitat-patches in a geographic space defined by x and y coordinates ranging from 0 to 60. The environmental conditions in the landscape were set to range within the interval [0,5]. Three types of spatial distribution of environmental conditions were considered: random, autocorrelated, or linear gradient. Temporal variation in environmental

conditions followed a sinusoidal function (plus a random error $\sim N(0,0.1)$) with 100 periods (e.g.,
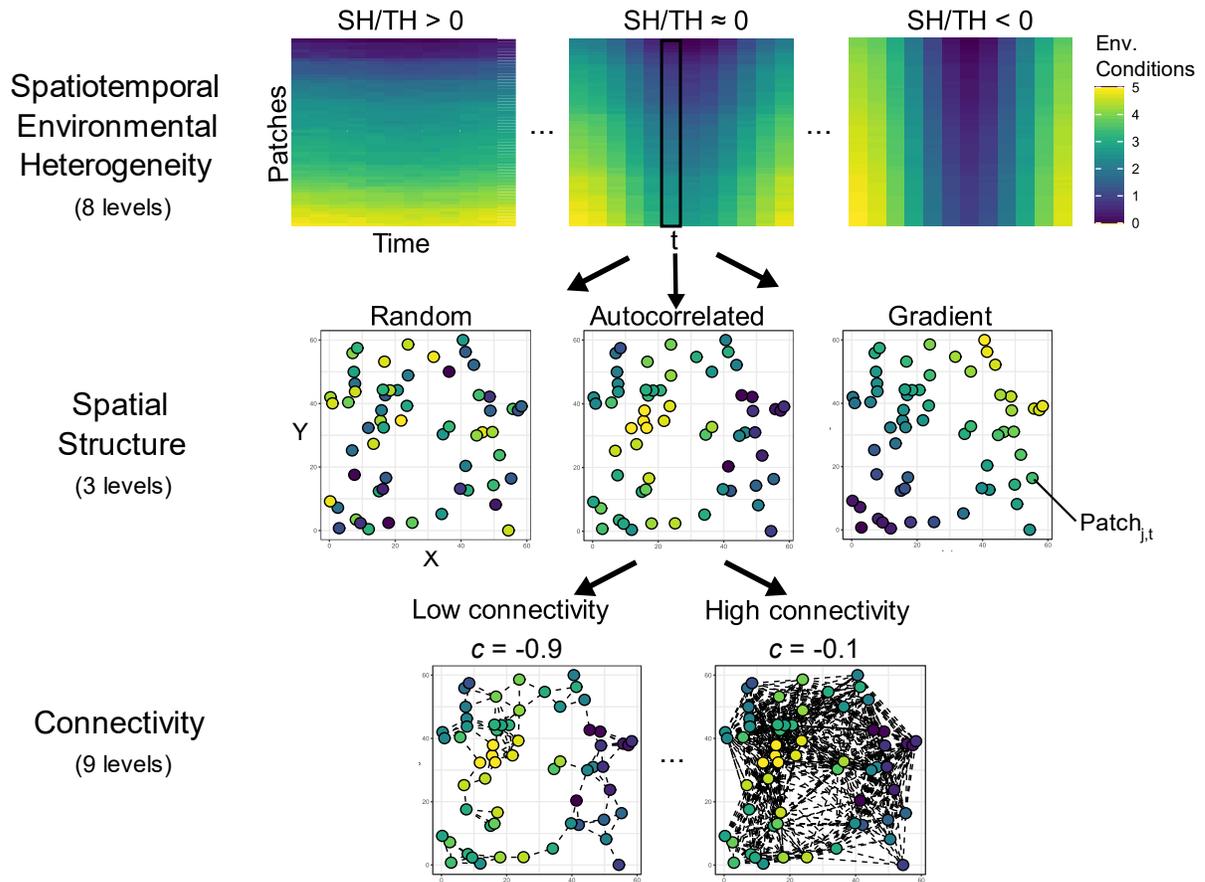
Figure 3.2: Schematic representation of simulated landscape characteristics. Spatiotemporal environmental heterogeneity SH/TH is calculated as the log of the ratio between the average variance of environmental conditions in space (SH) and the average variance of environmental conditions in time (TH). In the top heatmaps, patches are ordered based on environmental characteristics to aid in the visual comparison between spatial (vertical color variation) and seasonal (horizontal color variation) environmental heterogeneity. Spatial structure represents the type of spatial distribution of environmental conditions considered in the simulations - from totally random, through autocorrelated landscapes, to a linear gradient. Connectivity decayed exponentially with geographic distance between patches at rate c and values below a fixed threshold were truncated to 0.

100 years), each consisting of 12-time steps, with distinct amplitudes to simulate different levels of landscape seasonality.

Landscape spatiotemporal environmental heterogeneity (SH/TH) was calculated as the log of the ratio between the average variance of the environment in space (SH) and the average variance of the environment through time (TH). SH/TH > 0 indicated spatially heterogeneous but aseasonal landscapes; SH/TH ≈ 0 indicates similar levels of environmental heterogeneity in space and time; SH/TH < 0 spatially homogenous but highly seasonal landscapes.

The degree of physical connectivity (*Connectivity)* between pairs of patches was set as a negative exponential function of their distance (see equation SI-3.I). *Connectivity* values below a

threshold of 10^{-4} were truncated to 0 to generate truly disconnected pairs of patches (as in Fournier et al. 2017). By varying the degree of exponential decay in connectivity but keeping this threshold constant, we generated landscapes with contrasting degrees of average connectivity among patches.

3.3.2 Species pools and metacommunity dynamics

At the beginning of each simulation run (time step = 1), we generated 100 species with distinct environmental optima (μ), environmental tolerance (σ), and dispersal ability (η; i.e., here defined as emigration propensity). μ, σ, and η were randomly drawn from continuous uniform distributions with ranges [0, 5], [0.1, 2], and [0.01, 0.5], respectively. This ensured that: (1) all simulation runs were seeded with groups of species with the same initial trait value distributions; (2) different combinations of σ and η (i.e., different life-history strategies) were equally likely across all simulation runs (e.g., specialists and poor dispersers, specialists and strong dispersers, generalists and poor dispersers, and generalists and strong dispersers).

Species were allowed to colonize and reach stable coexistence in landscapes with distinct attributes (described above). The set of species that persisted in the metacommunity at the end of each simulation run (i.e., after reaching stable coexistence; see below) was the operational definition of "species pool" in this study. This operational definition aligns with the definition used in empirical studies in metacommunity ecology (Cornell and Harrison 2014)., i.e., it refers to the set of all species sampled across local communities in a metacommunity. It also implies the assumption that changes in local communities driven by mechanisms operating at fine spatiotemporal scales (here, environmental selection, dispersal, and demographic stochasticity) scale up to impact the size and composition of species pools directly (Fukami 2015). Refer to Supplementary Information I for a detailed discussion on the assumptions associated with this operational definition of species pool.

Our model, largely inspired by Büchi and Vuilleumier (2014), Shoemaker and Melbourne (2016), and Thompson et al. (2020), generates metacommunity dynamics through a combination of density-dependent (intra and interspecific competition) and density-independent (species-environment sorting) selection, spatial and temporal dispersal, and ecological drift (see schematic representation in Supp. Information Figure SI 3.I).

Considering that $N_{i,j,t}$ is the abundance of species i in patch j in time t, population dynamics was governed by:

$$N_{i,j,t} = Poisson\left(N_{i,j,t-1} * P_{i,j,t}\right) - \left(E_{i,j,t\ total}\right) + \left(I_{i,j,t\ total}\right) \qquad (3.1)$$

The first term of equation 3.1 is a modified version of a Beverton-Holt model that equates discrete population growth as a function of selection and ecological drift (i.e., demographic stochasticity). $P_{i,j,t}$ is the local performance (i.e., growth rate) of species i when conditioned to competition and habitat selection in patch j and time t, and is modeled as:

$$P_{i,j,t} = R_{i,j,t} * \frac{1}{(1 + \alpha_{intra} N_{i,j,t} + \alpha_{inter} \sum_{k \neq i}^{Sp} N_{k,j,t})} \quad (3.2)$$

where $R_{i,j,t}$ is the influence of local environmental conditions on species performance given by a Gaussian response:

$$R_{i,j,t} = u * \frac{1}{\sigma_i \sqrt{2\pi}} * \exp\left(\frac{-(Env_{j,t} - \mu_i)^2}{2\sigma_i^2}\right) \quad (3.3)$$

where Env_{jt} represents local abiotic conditions. The term $1/(\sigma_i \sqrt{2\pi})$ in equation 3.3 scales species responses to the environmental gradient, ensuring that, in the absence of competition, all species that share the same environmental optima have identical cumulative growth rates along the environmental gradient regardless of their niche breadth (i.e., same areas below the performance-environment curves, see Supp. Information and Büchi and Vuilleumier 2014). As such, any artificial advantages that may have influenced the persistence and dominance of either specialists or generalists in different landscapes were removed. u (set at 10 after pre trials that showed it allows the persistence of a larger number of species over time) is a scaling factor that ensures that all species were able to reach positive growth (i.e., $P_{i,j,t} > 1$) when local abiotic and biotic conditions were suitable.

The term on the right of equation 3.2 models the effects of density-dependent competition on population size at the intraspecific and interspecific levels. α_{intra} represents the per capita effects of species i on itself, whereas α_{inter} is the per capita effect of all other species on the local performance of i. Here, we assumed stabilizing competition in which $\alpha_{intra} > \alpha_{inter}$. This assumption is relevant because stabilizing competition favors coexistence by increasing the chances of locally rare species to keep positive population growth when locally dominant species have reached equilibrium at high abundances (i.e., the so-called "invasibility criterium" for coexistence; Chesson 2000, Grainger et al. 2019). By assuming stabilizing competition, we increased the chances of species with different life-history strategies to coexist in suitable habitats and persist in the metacommunity (Thompson et al. 2020). We acknowledge that competition types other than stabilizing (e.g., equalizing: $\alpha_{intra} = \alpha_{inter}$, destabilizing: $\alpha_{intra} < \alpha_{inter}$) may be important to metacommunity dynamics, but evaluating their influence on the way landscapes and species pools are related is beyond our goals here (but see Thompson et al. 2020, Wisnoski and Shoemaker 2022). Across all simulations α_{intra} and α_{inter} were set to 1/400 and 1/800 (minimum values that allowed for species regional coexistence at high abundances based on pre trials), respectively.

We added ecological drift (demographic stochasticity) into local birth and survival by drawing the final local abundance of species i from a Poisson distribution (equation 3.1). This distribution's mean was determined by the deterministic influence of biotic density-dependent (here, competition) and abiotic density-independent (environment sorting) selection on population dynamics (following Shoemaker and Melbourne 2016; Shoemaker et al. 2022).

Individuals able to persist in any given local community after within-patch selection and local demographic stochasticity at time t could then disperse. To align our framework with recent developments in metacommunity ecology (e.g., Wisnoski et al. 2019, Wisnoski and Shoemaker 2022), we modeled two types of dispersal: spatial and temporal (Buoro and Carlson 2014). Here, we define temporal dispersal as any physiological (e.g., diapause, dormancy) or behavioral strategies (e.g., hiding in refugia) that buffer local extinctions. These strategies enable individuals to escape from short-term unfavorable conditions by avoiding costs related to reproduction and resource consumption. This was operationalized by temporally removing individuals from local communities and allowing them to return to the same patch in the future (see below). Temporal dispersal is relevant because, akin to spatial dispersal, it promotes local and regional coexistence when local abiotic and biotic conditions favor competing species in different periods (i.e., via temporal storage effects, Chesson 2000, Wisnoski and Shoemaker 2022). Therefore, dispersal in space and time can be understood as alternative risk-spreading strategies that can maximize species persistence in metacommunities under varying levels of spatial and temporal environmental heterogeneity (Buoro and Carlson 2014; Holyoak et al. 2020).

The total number of emigrants of species i leaving patch j in time t ($E_{i,j,t,total}$) is determined by binomial trials with a size equal to the outcomes of within-patch dynamics (first term of equation 3.1), and the probability of success defined as the species-specific dispersal ability (η). Species with higher η were more propense to emigrate than species having lower η. To further explore the effects of spatial and temporal dispersal on the model outcomes, we created different scenarios wherein species would be more or less likely to undergo either type of dispersal. This was achieved by adjusting the values of the parameter *Dispersal Strategy*, which represents the probability of success in binomial trials used to determine the number of emigrants in $E_{i,j,t,total}$ that would undergo temporal dispersal ($E_{i,j,t,time}$). It follows that the number of spatial emigrants ($E_{i,j,t,space}$) is then given by $E_{ijt,total}$- $E_{ijt,time}$. We considered three different scenarios: in the "Equal Scenario", species had an equal probability of emigrating through either spatial or temporal dispersal (*Dispersal Strategy =0.5*); whereas in the "Mainly temporal dispersal" and "Mainly spatial dispersal" scenarios, *Dispersal Strategy* was set as very high (0.99) and very low (0.01), respectively, for all species.

The total number of immigrants of species i arriving at patch j in time t ($I_{i,j,t,total}$) was given by the sum of spatial ($I_{i,j,t,space}$) and temporal ($I_{i,j,t,time}$) immigrants. Spatial immigration was spatially explicit, meaning that individuals were more likely to immigrate to closer patches than distant ones. This was operationalized as follows. Consider the total number of spatial emigrants departing from patch h at time t *($E_{i,h,t,space}$)*. Let D$_{space}$ be the set of potential destination patches (e.g., j) of each one of these individuals, and let P$_{space}$ be the corresponding set of unequal sampling probabilities (scaled to sum to 1) of drawing any element in D$_{space}$ in a random sampling process. These unequal probabilities were given by the degree of *Connectivity* between h and neighboring patches (which decayed exponentially with distance following equation SI-3.I). Based on the probabilities in P$_{space}$, a random sampling process with replacement was repeated $E_{i,k,t,space}$

times to define the destination of spatial emigrants (see mathematical definition in Supp. Information). As such, the number of individuals of species i that left patch h and immigrated to patch j was then given by the number of times patch j was randomly drawn from D_{space}. It follows that $I_{i,j,t,space}$ is given by the sum of all individuals of species i that immigrated to patch j at time t coming from all other patches connected to j.

A similar procedure was used to determine temporal immigration. Consider $E_{i,j,t-x,time}$ as the total number of individuals of species i that underwent temporal dispersal (e.g., entered dormancy) in patch j at time t-x. Let D_{time} be the set of potential moments in the future (e.g., t) when individuals can recover from dormancy, and P_{time} be the set of probabilities of drawing each element in D_{time} in a random sampling process. Contrary to previous studies that assumed a constant recovery rate from "dormancy" over time (e.g., Wisnoski et al. 2019), we considered a more realistic temporal decay in recovery rates. For instance, individuals that underwent dormancy at time t-x were more likely to recover from dormancy at time t, if t is in the imminent future. This was operationalized by making the probabilities in P_{time} to be *exp(-dt*Δt)*, where Δt is the difference in time between t *and* t-x (Δt, min = 1, max =11), and dt is the rate of temporal decay. After pretrials where we tested different values for dt (not shown), we fixed it at 0.5 because it was the lowest value that allowed species persistence in highly seasonal and disconnected landscapes. Considering D_{time} and P_{time}, a sampling process with replacement was repeated $E_{i,j,t-x,time}$ times. The total number of individuals of species i that underwent dormancy in patch j at time t-x and recovered from dormancy in the same patch at time t was given by the total number of times t was drawn from D_{time}. It follows that $I_{i,j,t,time}$ is the sum of all individuals of species i that underwent in dormancy in patch j at a given moment in the past and recovered from dormancy at time t.

3.3.3 Simulation iterations

For each parameter combination and dispersal scenario, we ran 20 independent replicates, yielding 12960 simulations runs in total (20 replicates x 8 SH/TH levels x 9 connectivity levels x 3 types of the spatial structure of environment x 3 dispersal scenarios). Population dynamics of the 100 initial species in the regional pool were carried across all 60 patches over 1200-time steps (i.e., 100 complete seasonal cycles). Between time-steps 1 to 120, all patches were simultaneously seeded with species populations randomly drawn from a Poisson distribution (λ = 0.5). This allowed an opportunity for establishment and population growth for all species, provided that local abiotic and biotic conditions were suitable. The random placement of species populations across patches allowed those with similar habitat conditions to develop communities with dissimilar compositions due to priority effects (Thompson et al. 2020). To ensure that model summaries were carried out in stable rather than transient metacommunities, only communities in the last seasonal cycle (last 12 time-steps) were analyzed. This decision was supported by sensitivity analyses (not shown) that demonstrated stabilization of species pools (rate of regional extinctions close to zero) after approximately 700 time-steps on average.

3.3.4 Analyzing simulated metacommunities

We determined the dominant life-history of species in the regional pool (i.e., all species that persisted in the metacommunity in the last seasonal cycle) by calculating the regional-relative-abundance weighted mean of niche breadth (hereafter metacommunity-weighted niche breadth) and dispersal ability (hereafter metacommunity-weighted dispersal ability) of each of the 12960 simulations. By doing so, we could derive theoretical predictions underlying the life-history strategies that maximized species persistence and dominance across different landscape types.

Our model was designed to generate insights into how landscape attributes and species pool characteristics influence inferences about the relative importance of different assembly mechanisms based on analytics commonly used to infer processes in empirical metacommunities (e.g., Cottenie 2005; Soininen 2014; Gálvez et al. 2022). To do so, we used variation partitioning (Borcard et al. 1992; Peres-Neto et al. 2006) to estimate the contribution of different groups of variables to the variation in community composition across simulated local communities. Since we used a simulation model that incorporates known processes and lacks missing predictors (such as unmeasured spatiotemporal environmental variables that influence species distribution), variation partitioning can draw direct inferences from the observed patterns, which may be challenging when using empirical data. Our simulations reproduced data commonly collected in metacommunity studies and were analyzed using the same inferential approach, enabling comparison and contextualization of our theoretical results with empirical findings (e.g., Nishizawa et al. 2022 and see "Empirical support" below).

Variation partitioning was applied to the final patch-by-species-by-time matrix. We started by calculating pairwise compositional dissimilarities matrices and then using generalized dissimilarity models (GDMs, Ferrier et al. 2007) to fit these as a function of environment, positive spatial autocorrelation (here represented by positive Moran`s eigenvector maps, MEMs, calculated based on the patch geographic xy coordinates, Dray et al. 2006), and positive temporal autocorrelation (here represented by Asymmetric eigenvector maps, AEMs, Blanchet et al. 2008). Pairwise dissimilarities were calculated using the abundance-based Bray-Curtis index, which is widely used and underlies the link and variance functions of GDMs (see Ferrier et al. 2007). Traditionally, the amount of variation in pairwise compositional dissimilarity matrices explained by environmental variables alone is considered a measure of the strength of species-environment sorting; the variation explained by spatial MEMs alone represents spatial autocorrelation in species distributions caused by the spatial signature of demographic events such as dispersal (Cottenie 2005, Beisner et al. 2006); the variation explained by temporal AEMs alone represents temporal autocorrelation in species dynamics associated with demographic events that are not related to extrinsic environmental factors (Legendre and Gauthier 2014). The variation explained by the covariation among variables (i.e., their joint contribution) was also estimated, though its association with specific ecological mechanisms is less clear, particularly in empirical data that often contain missing environmental variables in the model (but see Peres-Neto et al. 2012).

Finally, we ranked the relative importance of each component (in ascending order) to facilitate comparisons across our 12960 independent simulation rounds.

All simulations and statistical analyses described above and below were conducted using R (v.4.1.0) (R Core Team 2023). AEMs and MEMs were generated using the *adespatial* package (Dray et al. 2022). We used the *vegan* (Oksanen et al. 2020) package to calculate compositional dissimilarities and the *gdm* package (Manion et al. 2018) to fit generalized dissimilarity models.

3.3.5 Identifying interdependencies between landscape characteristics, species pool attributes, and assembly mechanisms

We used path analysis to identify the causal interdependencies (pathways) between landscape attributes (i.e., exogenous variables: SH/TH, connectivity, and spatial structure of environment [ordinal; 1= Random, 2= Autocorrelated, 3 = Gradient]), species pool characteristics (i.e., mediators: metacommunity-weighted niche breadth and metacommunity-weighted dispersal ability), and the variation partitioning components (i.e., endogenous variables) across the 12960 simulation rounds. All predictors were standardized (mean = 0 and standard deviation =1) prior to model fit to allow comparing fitted relationships. Pathways' direction (i.e., positive/negative) and magnitude (i.e., standardized estimates) represented the general theoretical predictions derived from our simulations. Given that we used simulated data, the p-values of parameter estimates in the path models were not used to assess pathway significance because large simulation replications can yield low p-values even with negligible effect sizes (see White et al. 2014). Refer to Results and Discussion section for how we selected which pathways to interpret.

To assess whether theoretical predictions vary in direction (i.e., quantitively) and/or magnitude (i.e., qualitatively) across dispersal scenarios (i.e., "Equal dispersal," "Mainly temporal dispersal," and "Mainly spatial dispersal"), we contrasted the outcomes of a path analysis that combined all dispersal scenarios (global model) with the results of separate path analyses for each dispersal scenario. We used AIC_c to evaluate the fit of path models that considered different combinations of the linear and quadratic terms of predictors. The models that considered only linear terms (simplest) were identified as the best-fitting path models in most cases. We used the *piecewiseSEM* R package (Lefcheck 2020) to fit the path models across dispersal scenarios.

3.3.6 Empirical support: the assembly of moth metacommunities in tropical and temperate mountainous landscapes

To generate empirical support for the core theoretical predictions derived from our conceptual and simulation model (see below), we analyzed published data on moth metacommunities in two mountainous landscapes: The tropical Mount Cameroon (hereafter MTC, Maicher et al. 2019) and the temperate H.J. Andrews experimental forest (hereafter AEF, Miller and Jones 2005; Highland et al. 2013).

Moths in both datasets were collected using light traps along an elevational gradient (MTC: from 35 to 2000 meters above sea level; AEF: from 400 to 1400 meters above sea level). Sampling was carried out at different moments of the reproductive season in each region (AEF: we used data from May 2004 to October 2004; MTC: different moments of the dry season and the transition between dry to wet and wet to dry seasons, see more details in references and Supp Information). Because moths are generally good spatial dispersers when adults and can persist in the landscape through prolonged juvenile diapause (Lees and Zilli 2019), both spatial and temporal dispersal are likely to influence the structure of moth metacommunities.

To estimate and contrast the degree of spatiotemporal environmental heterogeneity (SH/TH), we used sample coordinates to extract monthly temperature (mean, max, minimum monthly values) and precipitation data at 1 km × 1 km resolution (CHELSA data, Karger et al. 2017). We then performed a principal component analysis on the temperature variables and log-transformed precipitation (standardized to mean = 0 and standard deviation = 1) and used the sample scores on the first two PC axes as a proxy of the climatic conditions of each site across different time periods of the year. In both the MTC and AEF, PC1 explained a substantial portion of the climate data variance at 76.2% and 74.2%, respectively, while PC2 accounted for 22.1% and 20.7%, resulting in cumulative proportions of 98.3% and 94.9%, respectively.

We estimated the climatic tolerance of each species through the tolerance index of Dolédec (2000) using the package "ade4" (Thioulouse et al. 2018). This index estimates species-specific climatic tolerance (i.e., niche breadth) based on the dispersal of samples that contain the target species in the multivariate climatic space. We pooled together data on moths and climate variables of both mountainous landscapes to estimate climatic tolerance in the same multivariate space. By doing so, we could directly contrast the degree of ecological specialization of species observed in both datasets. Lastly, we inferred the relative importance of different assembly mechanisms in both landscapes using variation partitioning (following the same steps described in *Analyzing simulated metacommunities*). This was done by estimating the variation in the community composition data explained by climate (PC1 and PC2), space (spatial MEMs), and time (temporal AEMs).

To consider how differences in sample design (e.g., the length of elevational range sampled in each mountain) of both datasets could influence our inferences, we rerun the analyses described above after removing samples in the MTC dataset so that it would span the same elevational range as the AEF dataset (i.e., approximately 1000 m.a.s.l.). Given that the results remained qualitatively the same (see Supp. Information), here we only report the results considering the complete elevational gradient in the MTC.

3.4 Results and Discussion

Due to qualitative similarities in the results of path models fitted considering each dispersal scenario separately (see tables SI 3.I -IV, Supp. Information), we only report the results considering data on all dispersal scenarios pooled together. However, we also highlight and discuss

cases in which there were differences in the direction of pathways across dispersal scenarios. For purposes of tractability and synthesis, we focused our discussion on the pathways with the highest importance in the fitted path models. That is, only pathways (minimum of two per mediator and endogenous variables) with partial coefficients higher than the median coefficient across all relationships among exogenous (landscape characteristics), mediators (species pools attributes), and endogenous variables (i.e., the isolated contribution of environment, spatial MEMs, and temporal AEMs on variation partitioning) are discussed and reported here. Nonetheless, the complete set of numerical relationships estimated by path analyses across all dispersal scenarios and considering the full set of variation partitioning components can be found in Supp. Information.

3.4.1 Theoretical Predictions: landscape attributes influence the degree of ecological specialization and dispersal ability of dominant species in the regional pool

Our simulation clearly showed that seasonality (measured as the ratio between spatial and temporal heterogeneity, SH/TH) was the most important factor determining the degree of ecological specialization of the dominant species in the regional pool (Figures 3.3 and 3.4). Ecological specialization was favored in aseasonal landscapes where environmental heterogeneity was higher in space than in time (SH/TH > 0). Conversely, ecological generalization was favored in highly seasonal landscapes where environmental heterogeneity is higher in time than in space (SH/TH < 0). Notably, we observed an increase in the persistence of ecological specialists in seasonal landscapes when we considered the "Mainly temporal dispersal" scenario (see Figure SI 3.II). These findings highlight the importance of temporal dispersal to the coexistence of specialists and generalists in (temporally) fluctuating environments (Chesson 2000; Wisnoski and Shoemaker 2022).

The spatial structure of the environment also influenced the overall niche breadth of species pools, but this relationship was relatively weak (Figures 3.3 and 3.4). When environmental conditions were randomly distributed across the landscape, an increase in the dominance of generalists in the regional pool was observed. Reduced spatial structure (autocorrelation) in habitat conditions increased the chances of environmental specialists being isolated in patches surrounded by unsuitable habitat conditions (Büchi and Vuilleumier 2014; Fournier et al. 2017). Since isolation increases populations' chances of becoming locally extinct due to demographic stochasticity, the lack of spatial structure in environmental conditions should increase isolation and, consequently, local, and regional extinction of ecological specialists.

Dispersal ability was influenced by seasonality and the level of connectivity in landscapes (Figures 3.3 and 3.4), though the strength of these relationships varied across dispersal scenarios (see Supp. Information, Tables SI 3.I-IV). When metacommunity dynamics were primarily driven by spatial dispersal (i.e., the "Mostly spatial" scenario), dispersal ability increased at intermediate levels of seasonality but increased linearly with physical connectivity. This finding suggests that highly connected landscapes reduce the risks associated with spatial dispersal by increasing the

likelihood of species successfully tracking suitable patches when environmental heterogeneity is equally strong in space and time (see Kubisch et al. 2014 and references within). In contrast, when
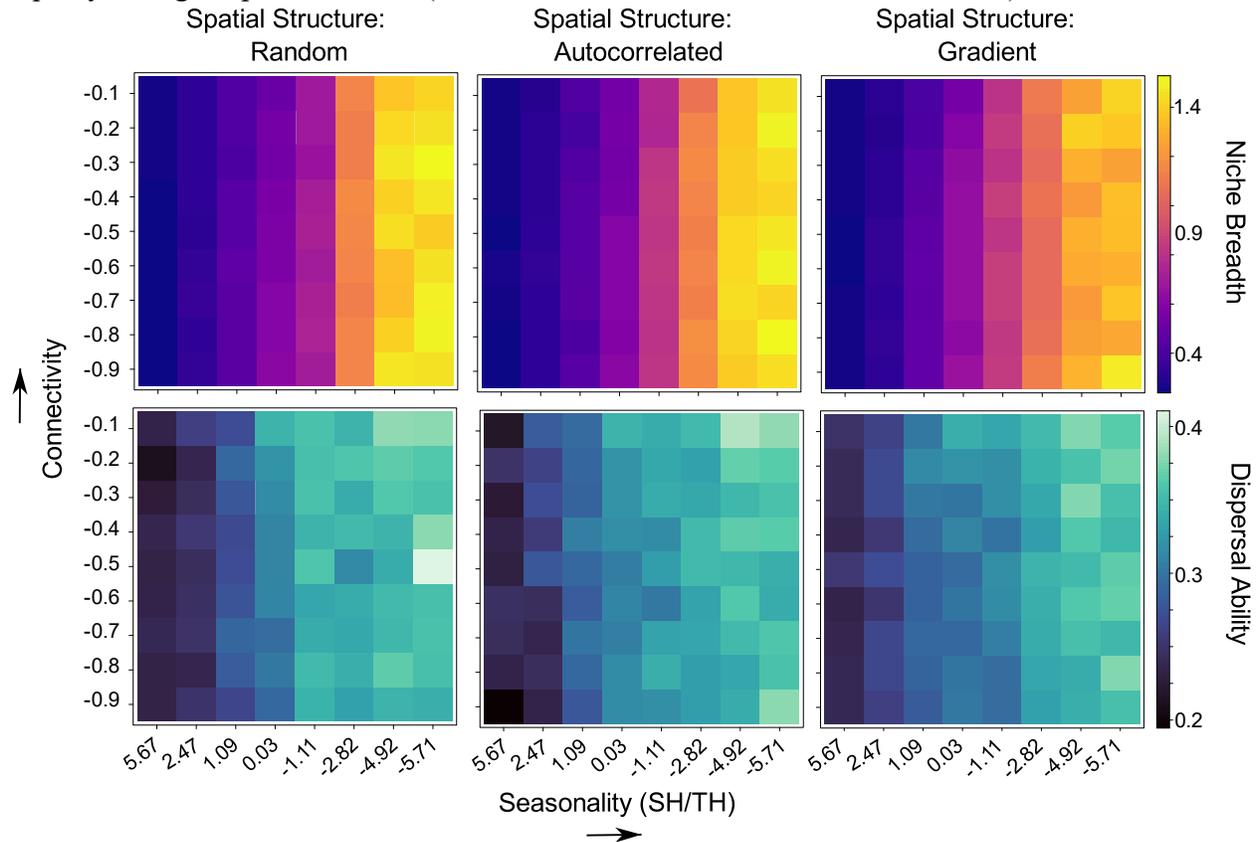
Figure 3.3: Landscape attributes determine the dominant life-history strategies in species pools. Among landscape attributes, variation in landscape seasonality was the main driver of variation on metacommunity-weighted niche breadth and dispersal ability (also see Fig.4). Aseasonal (SH/TH > 0) landscapes selected for environmental specialists (i.e., narrow niche breadth) that were also weak dispersers (i.e., low dispersal ability). Seasonal (SH/TH < 0) favored the dominance of environmental generalists that were also strong dispersers. These are the results reported for the "Equal" dispersal scenario where species were equally likely to disperse spatially and temporally. The results for the "Mostly Spatial" and "Mostly Temporal" dispersal scenarios are reported in Supp. Information Figures SI 3.III-IV.

species mainly dispersed over time (the "Mostly Temporal" scenario) or had equal chances of dispersing in space and time (the "Equal" scenario), landscape's spatiotemporal environmental heterogeneity (which is negatively correlated with seasonality) emerged as the most important landscape attribute selecting for species with weak dispersal abilities. This implies that species' ability to disperse in time is critical to their persistence in highly seasonal landscapes. Additionally, these results illustrate that spatial and temporal dispersal are risk-spreading strategies favored by different levels of spatial environmental heterogeneity (Buoro and Carlson 2014).

Collectively, our findings provide theoretical support for macroecological hypotheses and ecogeographic rules invoked to explain latitudinal clines on species' ecological specialization and

dispersal ability. For instance, Janzen`s seasonality hypothesis posits that the high elevational stratification of climate and the low seasonality of tropical mountainous landscapes should favor the dominance of environmental specialists whose spatial distributions are restricted to different types of climate (Janzen 1967). Conversely, strong seasonality in temperate regions should favor the dominance of species that have broad physiological tolerances and are less sensitive to spatial variation in climate (Sheldon and Tewksbury 2014). Previous studies have demonstrated that niche evolution through a mutation-selection process is critical to the patterns predicted by Janzen's hypothesis (e.g., Hua 2016). Our study expands on this understanding by demonstrating that latitudinal clines in niche breadth can arise due to metacommunity dynamics at fine temporal scales where speciation and trait evolution is expected to play a minimal role in community assembly.

Our model successfully replicated the expected relationship between temporal variability in the environment and the optimal level of dispersal ability in the regional pools that shape metacommunities (e.g., Jocque et al. 2010; Sheard et al. 2020). We found that weak dispersers that are also highly specialized in local conditions dominate local communities and increase their persistence in the regional pool when the environment is temporally homogenous. In contrast, high temporal variability of environmental conditions favored species with increased dispersal ability that can escape from temporally unsuitable local conditions (Figure 3.3). Given that: (i) dispersal ability can contribute substantially to geographic ranges (Alzate and Onstein 2022, but see Lester et al. 2007); and (ii) the strength of seasonality (particularly in temperature) increases from the equator to the poles, our model was able to recreate the underlying conditions that lead to an increase in range size as a function of latitude, as predicted by the Rapoport`s rule (Stevens 1989).

3.4.2 Theoretical Predictions: landscape and species pool attributes influence inferences about the relative importance of assembly mechanisms

Our theoretical framework allowed us to generate a mechanistic understanding of how landscape characteristics and species pools can influence empirical inferences about the relative importance of assembly mechanisms in metacommunities (Figure 3.4). The unique contribution of the environment (via variation partitioning) captures the importance of species-environment sorting in community assembly (Cottenie 2005, Ovaskainen et al. 2019). Path analyses applied to the combined results of all dispersal scenarios indicate that the strength of species-environment sorting on community composition is reduced when landscapes are composed of large clusters of suitable habitat conditions. They also indicate that species-environment sorting increases when species pools are dominated by environmental specialists that are weak disperses (i.e., the species sorting paradigm).

However, the direction of the relationship between dispersal ability and the contribution of environmental variables in the variation partitioning was not constant across dispersal scenarios

(see Supp. Information, tables SI 3.I-IV). When spatial dispersal occurs as frequently as, or more frequently than, temporal dispersal (i.e., the Equal and Mostly Spatial scenarios), dispersal ability increased the relative importance of environmental selection in community assembly. This suggests that the influence of the environment on community composition can intensify with spatial dispersal when it increases the likelihood of specialists reaching and persisting in large numbers of suitable patches. Conversely, this relationship becomes negative when dispersal is constrained to be mostly temporal (i.e., under the Mostly Temporal scenario). This pattern suggests that "seed banks" buffer the extinction of populations in unsuitable local conditions, decreasing the strength of the match between community composition and environment (Wisnoski et al. 2019).
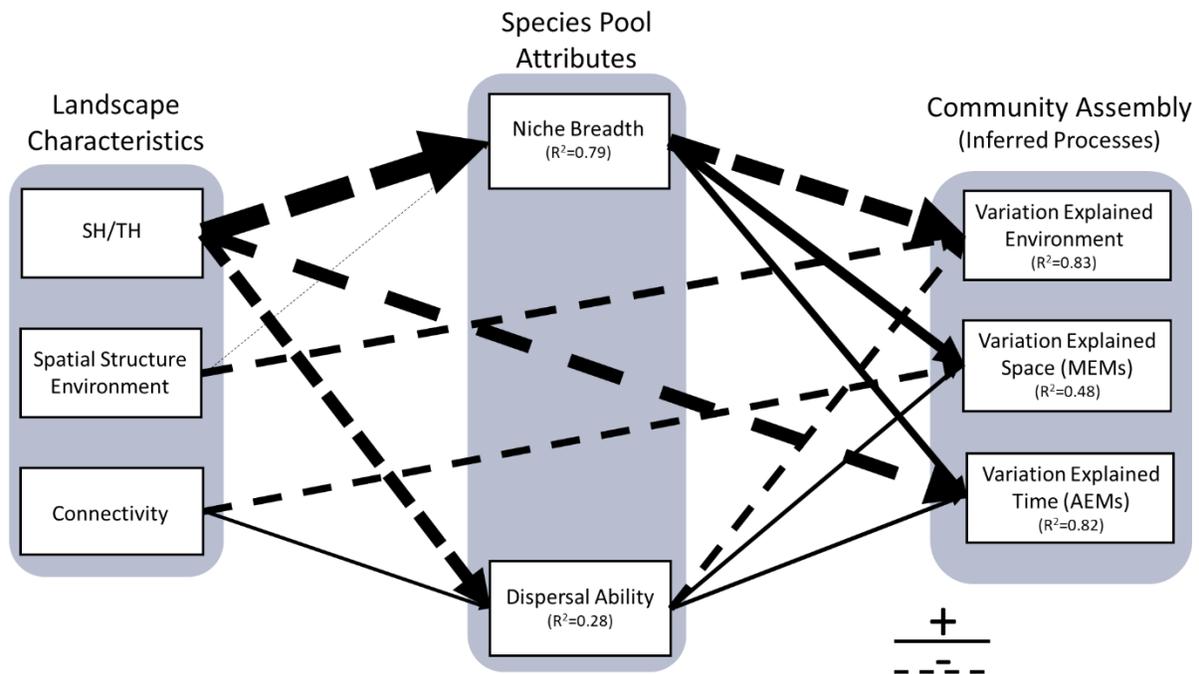
Figure 3.4: Theoretical predictions derived from path analysis considering the relationships between landscape characteristics (exogenous variables), the dominant life-history strategies in species pools (mediators), and the variation partitioning components (endogenous variables). For purposes of tractability and synthesis, only pathways with effect sizes higher than the median absolute effect sizes across all relationships among exogenous (landscape characteristics), mediators (species pool attributes), and endogenous (variation partitioning assembly) are reported here. Arrow widths are proportional to the effect sizes estimated. The SH/TH index is given by log of the ratio between spatial and temporal environmental heterogeneity. It has positive values in landscapes where spatial environmental variation is stronger than seasonal variation but negative values in landscapes where spatial environmental variation is weaker than seasonal variation. Results reported considering all dispersal scenarios pooled together. The numerical results obtained from path analyses considering each dispersal scenario pooled together and separately are reported in Supp. Information (Tables SI 3.I-IV).

The importance of the unique contribution of space (spatial MEMs) is typically associated with the influence of dispersal limitation in community assembly (Cottenie 2005, Beisner et al. 2006). We observed that community composition became more spatially structured as niche breadth and dispersal propensity increased but decreased with landscape connectivity (Figure 3.4). This suggests that spatial autocorrelation in community composition unrelated to the spatial structure of the environment arises when species with weak responses to environmental gradients are constrained to dispersing to neighboring patches.

The proportion of variation in the community matrix explained by the unique contribution of temporal variation (AEMs) is usually linked to temporal autocorrelation in population dynamics unrelated to environmental variation (Legendre and Gauthier 2014). Our simulations indicate that this type of autocorrelative pattern tends to increase when temporal environmental variation is weak (i.e., aseasonal landscapes) and generalists with strong dispersal capacity dominate metacommunities (Figure 3.4). Under these conditions, stochastic events of colonization and local extinctions outweigh the influence of species-environment sorting in generating temporal autocorrelation in population dynamics. These results are aligned with previous empirical studies demonstrating that the stochastic signature of temporal changes in community composition increases in aseasonal landscapes where environmental heterogeneity is stronger in space than in time (e.g., Khattar et al. 2021).

In summary, our model demonstrates that landscape attributes and species pool characteristics are strongly associated and should not be considered as independent axes in the assembly process. It also demonstrates that this link can lead to variation in the relative importance of assembly mechanisms along broad-scale gradients that encompass variation in key landscape attributes

3.4.3 Empirical Support

While our model should not be interpreted as an attempt to scale directly with the dynamics of any given real metacommunity, it generated testable predictions on empirical data (Figure 3.5). For instance, a strong prediction derived from our model is that in landscapes where environmental heterogeneity is relatively greater in space than in time (aseasonal landscapes, SH/TH > 0), species pools should be dominated by environmental specialists (Figures 3.3 and 3.4). Consequently, environmental selection should be the primary mechanism driving community assembly in these landscapes. Conversely, generalists should dominate species pools in landscapes where environmental conditions change relatively more in time than in space (seasonal landscapes, SH/TH < 0). As such, it is reasonable to infer that mechanisms beyond environmental selection alone likely play a significant role in driving community assembly.

As typically observed in tropical mountains (Figure 3.5 panel A), climate (PC1 scores) varied more across elevations than over time in MTC (SH/TH = 2.12). Conversely, climate varied more in time than across elevations in the temperate AEF (SH/TH= -2.56). In Figure 3.5 (panel B), we contrasted the degree of climatic tolerance of the dominant species in the regional pool of

both landscapes. As anticipated, the aseasonal MTC exhibited a species pool dominated by climate specialists (i.e., metacommunity-weighted mean climatic tolerance = 0.21), while the seasonal AEF favored the prevalence of climate generalists in its pool (metacommunity-weighted mean climatic tolerance = 1.23).

As predicted (Figure 3.5, panel C), variation in community composition in the aseasonal MTC (where specialists dominated the species pool) was mostly explained by climate variation alone. This pattern suggests a strong influence of species-environment sorting in community assembly in aseasonal landscapes. In contrast, variation in community composition in the highly
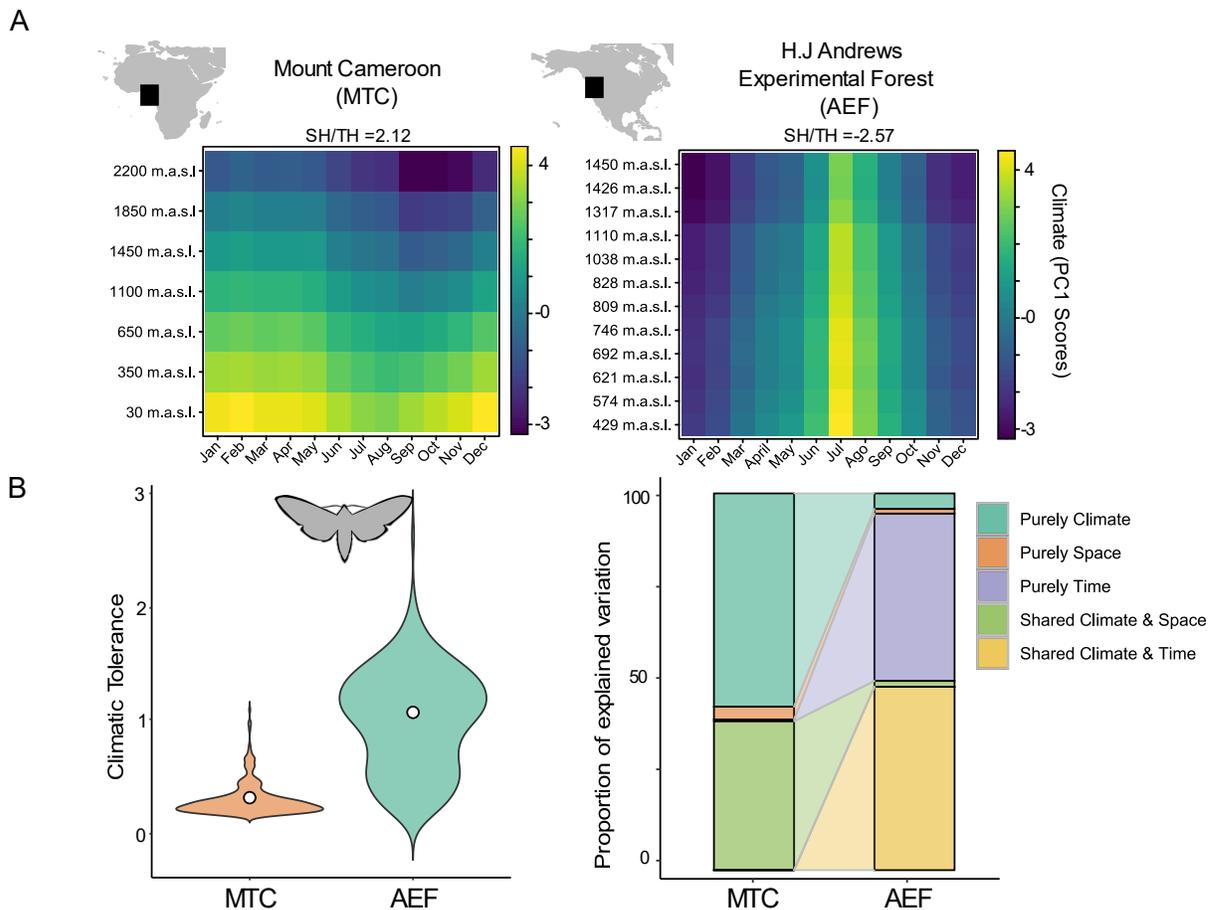
Figure 3.5: We analyzed the assembly of moth metacommunities in two different mountainous landscapes: the tropical and relatively aseasonal Mount Cameroon (MTC, SH/TH >0) and the temperate and relatively seasonal H.J Andrews experimental forest (AEF, SH/TH <0) (panel A). In the MTC, the regional pool is dominated by climate specialists, while climate generalists dominate the regional pool in the AEF (panel B). As such, deterministic species-environment sorting is the primary driver of community assembly in the MTC, whereas temporal autocorrelation on population dynamics and the temporal structure of climate are the main drivers of variation in community composition in the AEF (panel C). White dots in panel B represent estimated metacommunity-weight climate tolerances. Shared contributions of climate, space, time, and time and space were extremely small in both metacommunities (< 0.1 %) and, therefore, were omitted in the plot in panel C

seasonal AEF (where generalists dominate the species pool) was mainly explained by temporal autocorrelation in community composition underpinned by endogenous demographic mechanisms and their association with climate (Legendre and Gauthier 2014).

3.5 Conclusions, assumptions, and future directions

In this study, we proposed a conceptual framework for metacommunity assembly that acknowledges the dependency of species pool attributes on landscape characteristics and elucidates how their combined and individual contributions determine the relative importance of different assembly mechanisms. By doing so, we derived testable predictions underlying geographical patterns of metacommunity assembly when inferred from empirical data.

While we recognize that our conceptual framework and theoretical model did not consider other aspects of landscapes that are known to influence the coexistence of specialists and generalists, these could be incorporated in future model versions. For instance, recent empirical studies have shown that the spatial frequency of climate conditions at large scales and patch heterogeneity are relevant factor determining the degree of ecological specialization of species pools (Fournier et al. 2020).

Lastly, our framework for the geography of metacommunities assumes that species pool dynamics are primarily influenced by mechanisms operating at the landscape scale while intentionally disregarding the effects of evolutionary and historical mechanisms operating at biogeographic scales. Nevertheless, our proposed framework proves valuable in advancing syntheses to explore the substantial variation in the relative importance of mechanisms observed in empirical metacommunities across different parts of large-scale ecological gradients. Future studies could explore how evolutionary processes mediate the relationships between dominant life-history strategies, landscape attributes, and assembly mechanisms at the metacommunity level (e.g., Mittelbach and Schemske 2015).

3.6 Supplementary Information

3.6.1 Simulated landscapes: Extended description

We started by randomly distributing 60 patches in a geographic space defined by x and y coordinates ranging from 0 to 60. The degree into which any given two patches are physically connected decays exponentially with Euclidean distance (ΔS) according to the following kernel function:

$$Connectivity_{i,j} = \exp{(-c * \Delta S_{ij})} \qquad \text{(SI 3.I)}$$

where the term c is the rate at which connectivity decays with spatial distance. *Connectivity* values below a threshold of 10^{-4} were truncated to 0 so that individuals could not move between the focal pair of patches, thus creating patches that are truly disconnected (as in Fournier et al. 2017). By

varying *c* (here from 0.1 to 0.9) but keeping the threshold constant, we could generate landscapes with contrasting degrees of average connectivity among patches. The degree of connectivity between any given pair of patches defines the weighted probabilities of spatial dispersal between these patches (see "Species pools and metacommunity dynamics" below). The range of values for parameter c [-0.1, -0.9],, which defined the *Connectivity* of landscapes were defined to ensure that generated landscapes were fully connected, i.e., all patches were connected to at least one neighboring patch. This is relevant because a landscape with disconnected patches could bias the outcomes of metacommunity dynamics by creating fully isolated communities. Whenever a landscape is generated, our code tests if the landscape is fully connected. If a disconnected landscape is generated, there is an automatic reattempt to recreate a fully connected one by redrawing the coordinates (xy) of patches in the 60x60 geographic space (max 1000 attempts). Our pretrials have shown that our function could not create fully connected landscapes when values < -0.9 are assigned to c even after 10000 attempts.

The environmental conditions in each landscape were set to range in the interval [0,5] to scale with species environmental optima (see main text) and varied in space across three different spatial types: random, autocorrelated, and linear gradient. In random landscapes, the initial environmental value of each patch was randomly drawn from a continuous uniform distribution *U(0,5)*. In autocorrelated landscapes, we modeled environmental conditions using a multivariate normal distribution (*mu*=0) with a covariance matrix defined as the exponential decay of environmental similarity between pairs of patches with distance ΔS and the constant phi (here, set at 0.15). In gradient landscapes, we modeled the initial environment of each site as a linear function of their x and y coordinates.

To simulate seasonal environmental variation, we set local environmental conditions to follow a sinusoidal function with 100 periods, each composed of 12-time steps (e.g., 100 years) plus a random error $N(0,.1)$. The amplitude of the sinusoidal variation in the environment over time was modulated by a multiplicative factor *s* (constant across all patches). As such, the higher the value assigned to *s*, the higher the seasonal variation in environmental conditions. The final 60 (patches) x 1200 (times) matrix containing the environmental values was used to calculate an index of spatiotemporal environmental heterogeneity, hereafter SH/TH. SH/TH was calculated as the log of the ratio between the average variance of the environment in space (i.e., SH - average variance across columns) and the average variance of the environment through time (i.e., TH - average variance across rows). SH/TH values higher than 0 are observed in landscapes environmental heterogeneity is stronger in space than in time (i.e., spatially heterogenous but aseasonal landscapes); values close to 0 indicate that the level of environmental heterogeneity is similar in space and time; values lower than 0 indicate that environmental heterogeneity is stronger in time than in space (spatially homogenous but highly seasonal landscapes). The range of values for parameter s that defined the strength of seasonal patterns in landscapes was chosen so that the maximum and minimum levels of the SH/TH index were equally distant from 0 [5.67, -5.71].

3.6.2. Species pools: Operational definitions and associated assumptions

Species pools are broadly defined as the set of all species in a region that are available to colonize a given local site (Cornell and Harrison 2014). Species pools are a core theoretical construct of community assembly theory, i.e., they are an abstract ecological entity inferred from empirical observations such as large-biodiversity surveys and museum data. Consequently, operational definitions of species pools (i.e., the set of species considered in analyses of community assembly) vary broadly among empirical studies (Cornell and Harrison 2014). The first reason for such variation is related to the idiosyncrasies of sampling designs (e.g., differences in the definitions of a "region" and a "locality" across studies). The second, which we will focus our arguments upon, stems from study-specific (implied) assumptions about the nature of the hierarchical relationships between local and regional diversity (Fukami 2005, 2015; Cornell and Harrison 2014). For sake of example, let's consider different, but not mutually exclusive, archetypical representations of species pools frequently observed across two widely known conceptual models for community assembly theory: The classic mainland-island models (Macarthur and Wilson 1967; Keddy 1992) and metacommunity models (e.g., Leibold 2004). Below we provide a brief description of the main differences between these archetypical models for community assembly. But see a detailed description in Fukami (2005, 2015)

Mainland-island models for community assembly represent species pools as being decoupled from local communities (e.g., Species pool (Mainland) in the Figure SI 3.I). Depicting species pools as being external to local communities is more than a stylish decision; it implies the assumption that species pools are little influenced by within-community mechanisms operating at the fine spatiotemporal scales, serving as an external reservoir of species that is shaped by evolutionary and historical mechanisms operating at broad-spatiotemporal scales. This definition of species pools is aligned with the perspective that local species composition is under strong regional control (Ricklefs 2008). This definition also underlies the "species pool" hypothesis (Taylor et al. 1990) and its more recent analogs (e.g., Lessard et al. 2012), which propose that the idiosyncratic evolutionary history of various species pools explains broad-scale variation in community patterns and process.

In contrast, metacommunity archetypal models define species pools as the collection of all species across local communities forming the metacommunity (e.g., Species pool (metacommunity) in Figure SI 3.I). Because they are formed by the combined diversity of local communities, these models imply that species pools dynamically change at fine spatiotemporal scales as a consequence of changes in local communities (Cornell & Harrison 2014; Fukami 2015). Therefore, changes in local communities driven by mechanisms operating at fine spatiotemporal scales directly impact the size and composition of "species pools".

In our study, our operational definition of species pools is more aligned with the definition in metacommunity models. By doing so we make explicit our assumption that the non-random association between species pools traits and landscapes can emerge a result of mechanisms operating at fine spatiotemporal scales (e.g., dispersal, drift, selection). This assumption is relevant

in the context of our simulation framework because we deliberately did not consider broad-scale evolutionary mechanisms (trait evolution, speciation, etc). We acknowledge that broad-scale mechanisms are of extreme importance to the genesis and long-term dynamics of species pools in natural systems. However, even without considering these mechanisms, our simulation framework was still able to recapitulate well-known patterns in trait distribution that underly relevant hypotheses in biogeography and macroecology (Janzen 1967; Stevens 1989).

It is important to highlight that both archetypical species pools are not mutually exclusive: "metacommunity pools" are nested within "Mainland species pools" (Cornell & Harrison 2014; Fukami 2015). The integration of these two pools imply a feedback loop wherein dynamics within and between communities scale up to drive changes in the metacommunity pool, while the mainland regional pool, trickles down to influence community dynamics (Mittelbach and Schemske 2015). In our simulation framework (partially illustrated in Figure SI 3.I) we operationalized this integration by seeding landscapes with groups of species with the same initial distribution of continuous traits in the beginning of each simulation iteration (akin Mainland-Island models), to then investigate patterns in the distribution of traits in the subgroup of species that were able to persist and dominate the landscape in the long run (the metacommunity species pool).
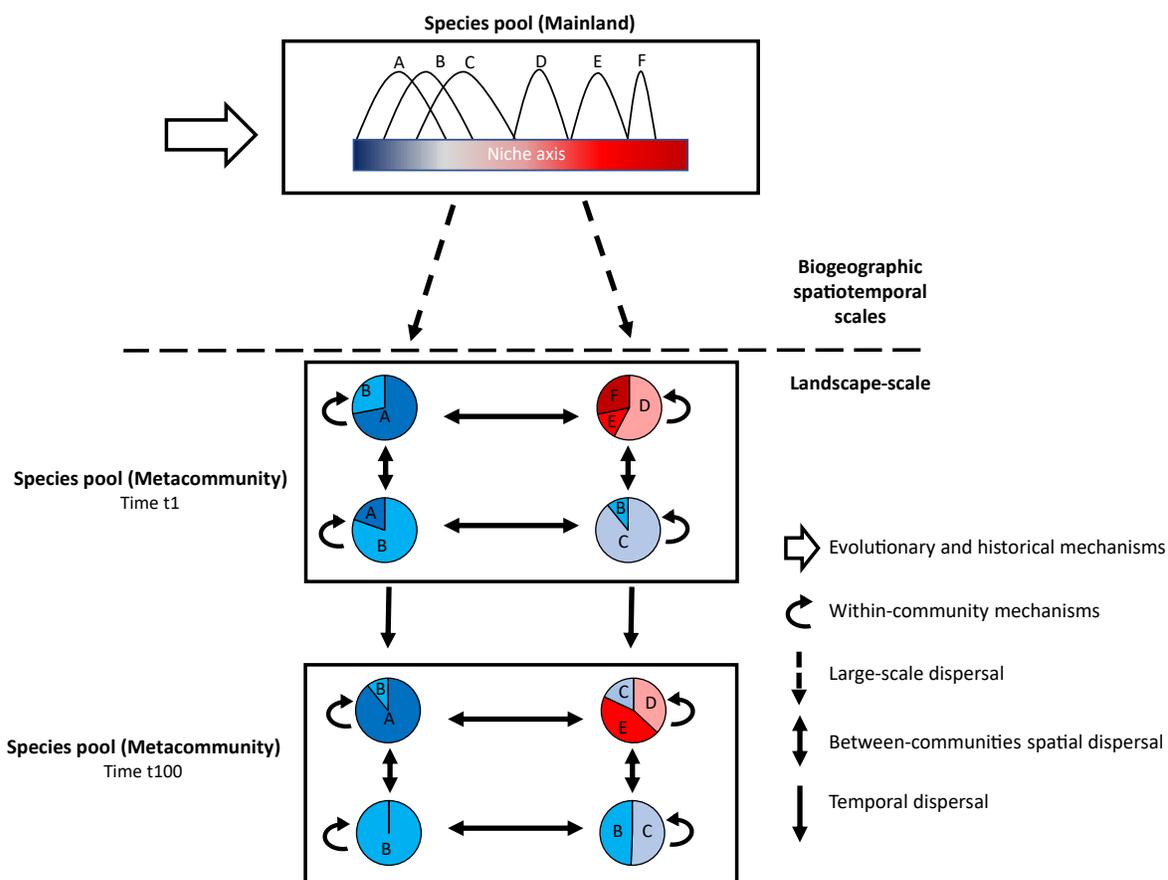
Figure SI. 3.I: A simple conceptual framework for community assembly that integrates different operational definitions of species pools, different types of dispersal, within-community assembly mechanisms (ecological selection and drift), and mechanisms operating at broader spatiotemporal scales. Pie charts represent different communities. The relative abundances of each species (capital letters) are represented by different colors in the pie charts. Adapted from Fukami (2015).

3.6.3 Scaling species performances based on their ecological tolerance.

Species performances peak at their environmental optima (μ), but the height and shape of performance curves depend on their niche breadth (σ). In Figure SI 3.II we depict estimated relationships using 5 species with different μ and σ as examples (following eq. 3 in main text). In this example, Sp1 has the highest level of specialization ($\sigma=0.1$) while Sp5 is the most generalist

(σ=1.5). All species have the same cumulative growth rate along the environmental gradient (i.e., the same areas under the performance curves) regardless of their species-specific σ.
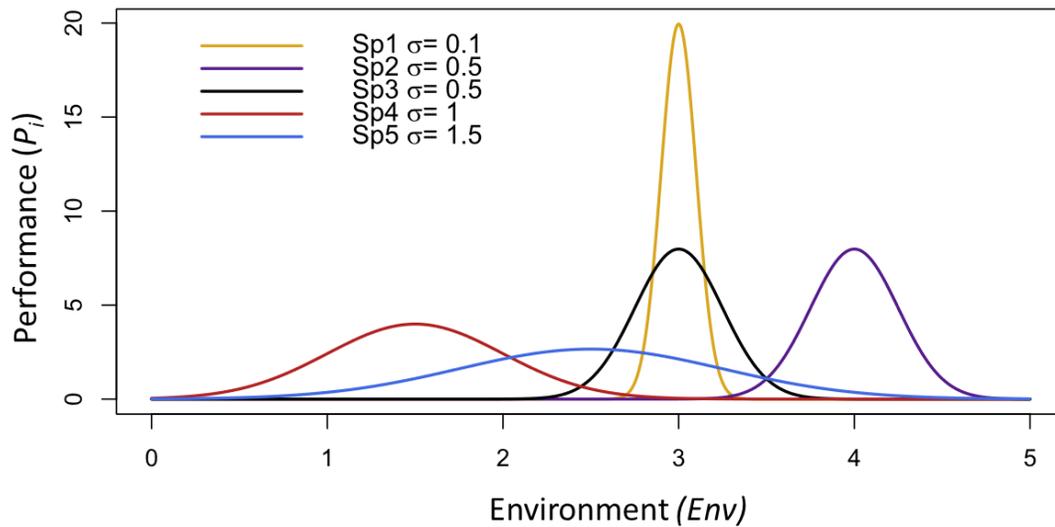
Figure SI. 3.II: Performances of five different species (Sp1 – Sp5) at different environmental values (Env) when competition is negligible due to the low population sizes. σ = Niche breadth. Adapted from Buchi & Vuilleumier, 2014.

3.6.4 Results across different dispersal scenarios
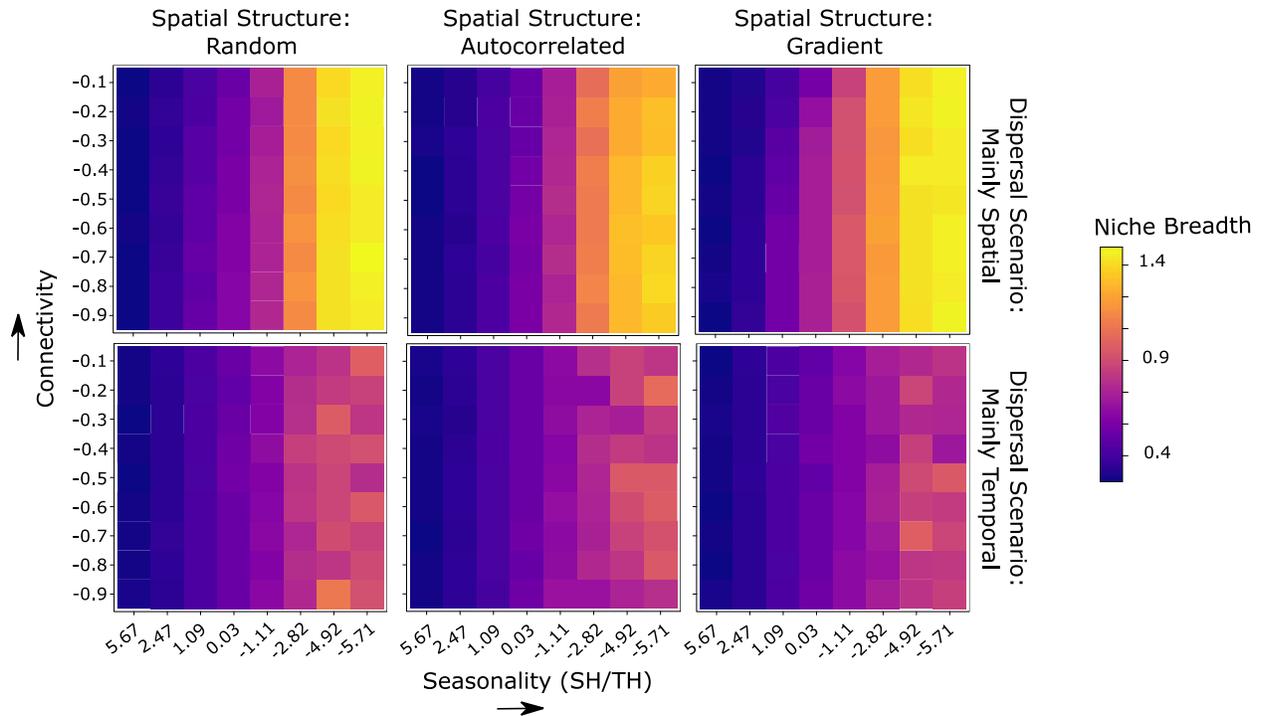
Figure SI. 3.III: Landscape attributes determine the dominant niche breadth in species pools. Aseasonal (SH/TH > 0) landscapes select for environmental specialists (i.e., narrow niche breadth). Seasonal (SH/TH < 0) landscapes favor the dominance of environmental generalists. Interestingly, when considering metacommunities mainly structured by temporal dispersal, we observe an increase in the persistence of species with a relatively narrow niche breadth in seasonal landscapes. Numerical relationships are depicted in Supp. Inf. tables SI 3.III and 3.IV
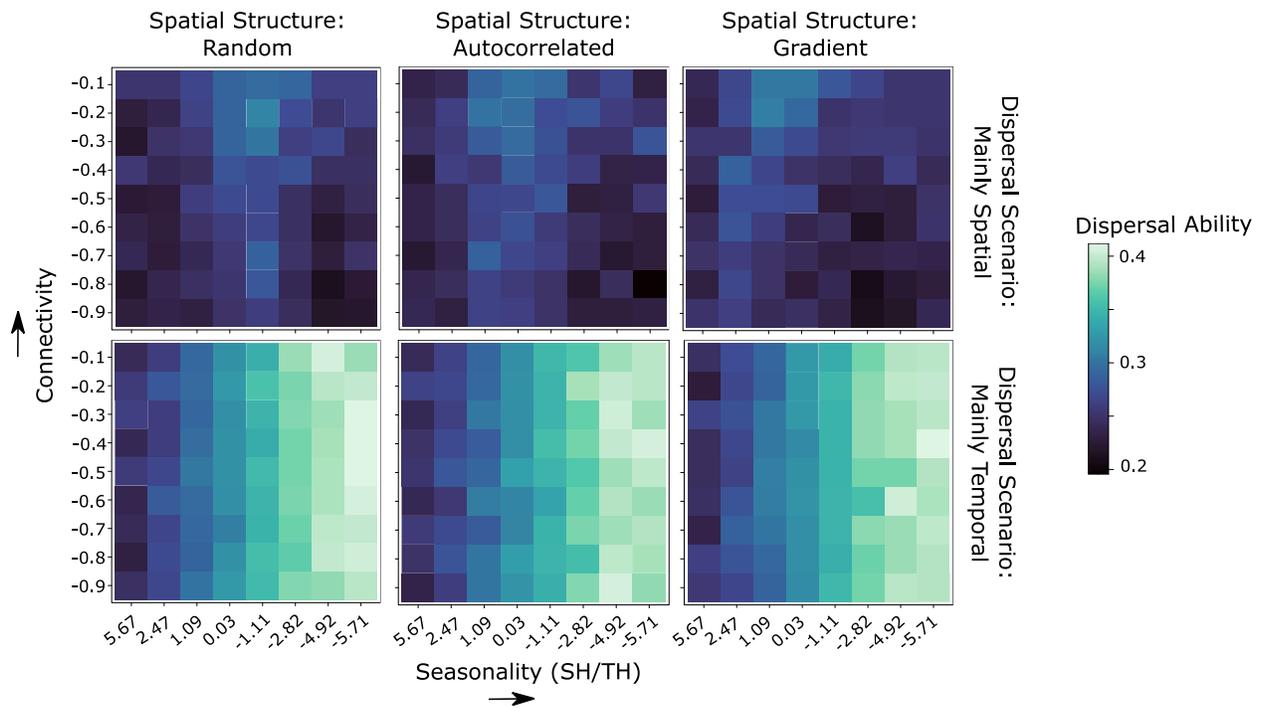
Figure SI. 3.IV: Landscape attributes determine the dominant dispersal ability in species pools. When spatial dispersal is more frequent than temporal dispersal, dispersal ability was maximized in highly connected landscapes where environmental heterogeneity in space and time were similar (i.e., SH/TH ≈ 0). In contrast, when species were constrained to disperse mainly in time (the "Mostly Temporal" scenario), dispersal ability was maximized at high levels of seasonality (i.e., SH/TH < 0). Numerical relationships are depicted in Supp. Inf. tables III and IV

Table SI 3.I: Results of Path Analyses fitted considering data on all dispersal scenarios pooled together. "NA" indicates non-applicable parameter estimation. "R.I. Joint All" is the amount of variation in the species data attributable to Environment ∩ Space (MEMs) ∩ Time (AEMs). In bold are the two most relevant pathways (larger standardized estimates) per moderators (Niche breadth and Dispersal ability) and endogenous variables (components of the variation partitioning approach)

Pathway	Std. Estimate	Std. Error	R-squared
Niche Breadth <-- SH/TH	**-0.8876**	0.0182	
Niche Breadth <-- Connectivity	-0.0309	0.0182	0.79
Niche Breadth <-- Spatial Structure	**-0.0311**	0.0182	
Dispersal Ability <-- SH/TH	**-0.5003**	0.0338	
Dispersal Ability <-- Connectivity	**0.1219**	0.0338	0.28
Dispersal Ability <-- Spatial Structure	-0.0228	0.0338	
R.I. Environment <-- SH/TH	0.0908	0.0652	
R.I. Environment <-- Connectivity	0.0839	0.0163	
R.I. Environment <-- Spatial Structure	-0.3020	0.0163	0.83
R.I. Environment <-- Niche Breadth	**-0.665**	0.0568	
R.I. Environment<-- Dispersal Ability	**-0.3094**	0.0305	
R.I. Space (MEMs) <-- SH/TH	-0.1251	0.1154	
R.I. Space (MEMs) <-- Connectivity	**-0.3444**	0.0288	
R.I. Space (MEMs) <-- Spatial Structure	0.1318	0.0288	0.48
R.I. Space (MEMs) <-- Niche Breadth	**0.414**	0.1005	
R.I. Space (MEMs) <-- Dispersal Ability	0.1689	0.0541	
R.I. Time (AEMs)<-- SH/TH	**-0.586**	0.054	
R.I. Time (AEMs)<-- Connectivity	0.1125	0.0135	
R.I. Time (AEMs) <-- Spatial Structure	-0.1109	0.0135	0.89
R.I. Time (AEMs) <--Niche Breadth	**0.2906**	0.047	
R.I. Time (AEMs) <-- Dispersal Ability	0.1644	0.0253	
R.I. Environment ∩ Space (MEMs) <-- SH/TH	-0.2405	0.0712	
R.I. Environment ∩ Space (MEMs) <-- Connectivity	-0.0124	0.0178	
R.I. Environment ∩ Space (MEMs) <-- Spatial Structure	0.2371	0.0178	0.8
R.I. Environment ∩ Space (MEMs) <-- Niche Breadth	**-0.9829**	0.0621	
R.I. Environment ∩ Space (MEMs) <-- Dispersal Ability	**-0.3681**	0.0334	
R.I. Environment ∩ Time (AEMs) <-- SH/TH	0.2847	0.0684	
R.I. Environment ∩ Time (AEMs) <-- Connectivity	0.0326	0.0171	0.82
R.I. Environment ∩ Time (AEMs) <-- Spatial Structure	0.0876	0.0171	

R.I. Environment ∩ Time (AEMs) <--Niche Breadth	**1.0117**	0.0596	
R.I. Environment ∩ Time (AEMs) <-- Dispersal Ability	**0.4752**	0.032	
R.I. Space (MEMs) ∩ Time (AEMs) <-- SH/TH	**1.4666**	0.1217	
R.I. Space (MEMs) ∩ Time (AEMs) <-- Connectivity	-0.0564	0.0304	
R.I. Space (MEMs) ∩ Time (AEMs) <-- Spatial Structure	0.2004	0.0304	0.42
R.I. Space (MEMs) ∩ Time (AEMs) <-- Niche Breadth	**0.8471**	0.106	
R.I. Space (MEMs) ∩ Time (AEMs) <-- Dispersal Ability	0.2706	0.057	
R.I. Joint All <-- SH/TH	**-0.3035**	0.1521	
R.I. Joint All <-- Connectivity	0.0169	0.038	
R.I. Joint All <-- Spatial Structure	-0.2171	0.038	0.1
R.I. Joint All <-- Niche Breadth	-0.2497	0.1325	
R.I. Joint All <-- Dispersal Ability	**-0.3409**	0.0713	
Correlated Errors			
Niche Breadth<--> Dispersal Ability	-0.7892	NA	NA
R.I. Space (MEMs) <--> R.I. Environment	-0.1255	NA	NA
R.I. Time (AEMs) <--> R.I. Environment	-0.4176	NA	NA
R.I. Environment ∩ Space (MEMs) <--> R.I. Environment	0.058	NA	NA
R.I. Environment ∩ Time (AEMs) <--> R.I. Environment	-0.3087	NA	NA
R.I. Space (MEMs) ∩ Time (AEMs) <--> R.I. Environment	-0.3943	NA	NA
R.I. Joint All <--> R.I. Environment	-0.2855	NA	NA
R.I. Time (AEMs) <--> R.I. Space (MEMs)	0.2067	NA	NA
R.I. Environment ∩ Space (MEMs) <--> R.I. Space	-0.387	NA	NA
R.I. Environment ∩ Time (AEMs) <--> R.I. Space	-0.4693	NA	NA
R.I. Space (MEMs) ∩ Time (AEMs) <-->RI.Space	-0.1753	NA	NA
R.I. Joint All <--> R.I. Space	-0.2244	NA	NA
R.I. Environment ∩ Time (AEMs) <--> R.I. Time (AEMs)	0.0451	NA	NA
R.I. Space (MEMs) ∩ Time (AEMs) <--> R.I. Time	-0.2935	NA	NA
R.I. Joint All <--> R.I. Time (AEMs)	0.5094	NA	NA
R.I. Environment ∩ Time (AEMs) <--> R.I. Environment ∩ Space (MEMs)	-0.6875	NA	NA
R.I. Space (MEMs) ∩ Time (AEMs) <--> R.I. Environment ∩ Space (MEMs)	-0.3975	NA	NA
R.I. Joint All <--> R.I. Environment ∩ Space (MEMs)	-0.3803	NA	NA
R.I. Space (MEMs) ∩ Time (AEMs) <--> R.I. Environment ∩ Time (AEMs)	0.5846	NA	NA
R.I. Joint All <--> R.I. Space (MEMs) ∩ Time (AEMs)	0.0545	NA	NA

Table SI 3.II: Results of Path Analyses fitted considering the Equal scenario. "NA" indicates non-applicable parameter estimation. "R.I. Joint All" is the amount of variation in the species data attributable to. Environment ∩ Space (MEMs) ∩ Time (AEMs). In bold are the two most relevant pathways (larger standardized estimates) per moderators (Niche breadth and Dispersal ability) and endogenous variables (components of the variation partitioning approach)

Pathway	Std. Estimate	Std. Error	R-squared
Niche Breadth <-- SH/TH	**-0.9551**	0.0213	
Niche Breadth <-- Connectivity	-0.044	0.0213	0.91
Niche Breadth <-- Spatial Strucure	**-0.0357**	0.0213	
Dispersal Ability <-- SH/TH	**-0.8014**	0.0225	
Dispersal Ability <-- Connectivity	**0.227**	0.0225	0.69
Dispersal Ability <-- Spatial Strucure	-0.0114	0.0225	
R.I. Environment <-- SH/TH	-0.1066	0.0906	
R.I. Environment <-- Connectivity	0.0078	0.0209	
R.I. Environment <-- Spatial Strucure	**-0.3412**	0.0196	0.91
R.I. Environment <-- Niche Breadth	**-1.0356**	0.0699	
R.I. Environment<-- Dispersal Ability	0.0664	0.0662	
R.I. Space (MEMs) <-- SH/TH	0.1354	0.2248	
R.I. Space (MEMs) <-- Connectivity	**-0.4432**	0.052	
R.I. Space (MEMs) <-- Spatial Strucure	0.174	0.0486	0.53
R.I. Space (MEMs) <-- Niche Breadth	**0.5053**	0.1736	
R.I. Space (MEMs) <-- Dispersal Ability	0.2624	0.1643	
R.I. Time (AEMs)<-- SH/TH	**-0.3312**	0.0885	
R.I. Time (AEMs)<-- Connectivity	0.1897	0.0205	
R.I. Time (AEMs) <-- Spatial Strucure	-0.1307	0.0191	0.92
R.I. Time (AEMs) <--Niche Breadth	**0.628**	0.0684	
R.I. Time (AEMs) <-- Dispersal Ability	-0.013	0.0647	
R.I. Environment ∩ Space (MEMs) <-- SH/TH	**-0.4309**	0.124	
R.I. Environment ∩ Space (MEMs) <-- Connectivity	-0.0784	0.0287	
R.I. Environment ∩ Space (MEMs) <-- Spatial Strucure	0.2196	0.0268	0.86
R.I. Environment ∩ Space (MEMs) <-- Niche Breadth	**-1.293**	0.0957	
R.I. Environment ∩ Space (MEMs) <-- Dispersal Ability	-0.0124	0.0906	
R.I. Environment ∩ Time (AEMs) <-- SH/TH	**0.4961**	0.1002	
R.I. Environment ∩ Time (AEMs) <-- Connectivity	0.1064	0.0232	0.9
R.I. Environment ∩ Time (AEMs) <-- Spatial Strucure	0.0957	0.0217	

R.I. Environment ∩ Time (AEMs) <--Niche Breadth	**1.2949**	0.0774	
R.I. Environment ∩ Time (AEMs) <-- Dispersal Ability	0.166	0.0732	
R.I. Space (MEMs) ∩ Time (AEMs) <-- SH/TH	**1.4316**	0.2024	
R.I. Space (MEMs) ∩ Time (AEMs) <-- Connectivity	0.0082	0.0468	
R.I. Space (MEMs) ∩ Time (AEMs) <-- Spatial Strucure	0.2915	0.0438	0.5
R.I. Space (MEMs) ∩ Time (AEMs) <-- Niche Breadth	**1.0485**	0.1562	
R.I. Space (MEMs) ∩ Time (AEMs) <-- Dispersal Ability	-0.1613	0.1479	
R.I. Joint All <-- SH/TH	**-0.6459**	0.2549	
R.I. Joint All <-- Connectivity	0.1774	0.0589	
R.I. Joint All <-- Spatial Strucure	-0.2539	0.0551	0.35
R.I. Joint All <-- Niche Breadth	-0.0782	0.1968	
R.I. Joint All <-- Dispersal Ability	**-0.9339**	0.1862	
Correlated Errors			
Niche Breadth<--> Dispersal Ability	-0.4485	NA	NA
R.I. Space (MEMs) <--> R.I. Environment	-0.2327	NA	NA
R.I. Time (AEMs) <--> R.I. Environment	-0.2005	NA	NA
R.I. Environment ∩ Space (MEMs) <--> R.I. Environment	-0.5315	NA	NA
R.I. Environment ∩ Time (AEMs) <--> R.I. Environment	0.2409	NA	NA
R.I. Space (MEMs) ∩ Time (AEMs) <--> R.I. Environment	-0.3842	NA	NA
R.I. Joint All <--> R.I. Environment	-0.4287	NA	NA
R.I. Time (AEMs) <--> R.I. Space (MEMs)	0.2485	NA	NA
R.I. Environment ∩ Space (MEMs) <--> R.I. Space	-0.559	NA	NA
R.I. Environment ∩ Time (AEMs) <--> R.I. Space	-0.4387	NA	NA
R.I. Space (MEMs) ∩ Time (AEMs) <-->RI.Space	-0.1381	NA	NA
R.I. Joint All <--> R.I. Space	0.0712	NA	NA
R.I. Environment ∩ Time (AEMs) <--> R.I. Time (AEMs)	-0.1717	NA	NA
R.I. Space (MEMs) ∩ Time (AEMs) <--> R.I. Time	-0.3601	NA	NA
R.I. Joint All <--> R.I. Time (AEMs)	0.3121	NA	NA
R.I. Environment ∩ Time (AEMs) <--> R.I. Environment ∩ Space (MEMs)	-0.62	NA	NA
R.I. Space (MEMs) ∩ Time (AEMs) <--> R.I. Environment ∩ Space (MEMs)	-0.5673	NA	NA
R.I. Joint All <--> R.I. Environment ∩ Space (MEMs)	-0.1733	NA	NA
R.I. Space (MEMs) ∩ Time (AEMs) <--> R.I. Environment ∩ Time (AEMs)	0.6593	NA	NA
R.I. Joint All <--> R.I. Space (MEMs) ∩ Time (AEMs)	-0.0548	NA	NA

Table SI 3.III: Results of Path Analyses fitted considering the "Mostly spatial" scenario. "NA" indicates non-applicable parameter estimation. "R.I. Joint All" is the amount of variation in the species data attributable to Environment ∩ Space (MEMs) ∩ Time (AEMs). In bold are the two most relevant pathways (larger standardized estimates) per moderators (Niche breadth and Dispersal ability) and endogenous variables (components of the variation partitioning approach)

Pathway	Std. Estimate	Std. Error	R-squared
Niche Breadth <-- SH/TH	**-0.9594**	0.0226	
Niche Breadth <-- Connectivity	**-0.0298**	0.0225	0.92
Niche Breadth <-- Spatial Structure	-0.0023	0.0226	
Dispersal Ability <-- SH/TH	**0.071**	0.0258	
Dispersal Ability <-- Connectivity	**0.5063**	0.0148	0.46
Dispersal Ability <-- Spatial Structure	-0.0197	0.018	
R.I. Environment <-- SH/TH	0.1061	0.0898	
R.I. Environment <-- Connectivity	-0.0051	0.024	
R.I. Environment <-- Spatial Structure	**-0.2663**	0.0198	0.91
R.I. Environment <-- Niche Breadth	**-0.7439**	0.0794	
R.I. Environment<-- Dispersal Ability	0.1896	0.0695	
R.I. Space (MEMs) <-- SH/TH	-0.1579	0.2158	
R.I. Space (MEMs) <-- Connectivity	**-0.4226**	0.0577	
R.I. Space (MEMs) <-- Spatial Structure	0.0204	0.0475	0.51
R.I. Space (MEMs) <-- Niche Breadth	**0.4189**	0.1908	
R.I. Space (MEMs) <-- Dispersal Ability	0.0139	0.167	
R.I. Time (AEMs)<-- SH/TH	**-0.2635**	0.0916	
R.I. Time (AEMs)<-- Connectivity	0.2482	0.0245	
R.I. Time (AEMs) <-- Spatial Structure	-0.1388	0.0202	0.89
R.I. Time (AEMs) <--Niche Breadth	**0.6416**	0.081	
R.I. Time (AEMs) <-- Dispersal Ability	-0.1196	0.0709	
R.I. Environment ∩ Space (MEMs) <-- SH/TH	**-0.5518**	0.1288	
R.I. Environment ∩ Space (MEMs) <-- Connectivity	-0.0466	0.0345	
R.I. Environment ∩ Space (MEMs) <-- Spatial Strucure	0.3076	0.0283	0.83
R.I. Environment ∩ Space (MEMs) <-- Niche Breadth	**-1.4026**	0.1139	
R.I. Environment ∩ Space (MEMs) <-- Dispersal Ability	-0.0703	0.0997	
R.I. Environment ∩ Time (AEMs) <-- SH/TH	**0.4617**	0.1201	
R.I. Environment ∩ Time (AEMs) <-- Connectivity	0.0096	0.0321	0.85
R.I. Environment ∩ Time (AEMs) <-- Spatial Structure	0.024	0.0264	

R.I. Environment ∩ Time (AEMs) <--Niche Breadth	**1.3963**	0.1062	
R.I. Environment ∩ Time (AEMs) <-- Dispersal Ability	0.1912	0.0929	
R.I. Space (MEMs) ∩ Time (AEMs) <-- SH/TH	**1.4685**	0.1641	
R.I. Space (MEMs) ∩ Time (AEMs) <-- Connectivity	-0.1843	0.0439	
R.I. Space (MEMs) ∩ Time (AEMs) <-- Spatial Strucure	0.3442	0.0361	0.31
R.I. Space (MEMs) ∩ Time (AEMs) <-- Niche Breadth	**1.1932**	0.1451	
R.I. Space (MEMs) ∩ Time (AEMs) <-- Dispersal Ability	0.1504	0.127	
R.I. Joint All <-- SH/TH	-0.0807	0.3	
R.I. Joint All <-- Connectivity	0.3189	0.0803	
R.I. Joint All <-- Spatial Strucure	-0.1707	0.066	0.34
R.I. Joint All <-- Niche Breadth	**-0.4103**	0.2653	
R.I. Joint All <-- Dispersal Ability	**-0.6547**	0.2322	
Correlated Errors			
Niche Breadth<--> Dispersal Ability	-0.6538	NA	NA
R.I. Space (MEMs) <--> R.I. Environment	-0.0362	NA	NA
R.I. Time (AEMs) <--> R.I. Environment	-0.2486	NA	NA
R.I. Environment ∩ Space (MEMs) <--> R.I. Environment	-0.2255	NA	NA
R.I. Environment ∩ Time (AEMs) <--> R.I. Environment	0.0639	NA	NA
R.I. Space (MEMs) ∩ Time (AEMs) <--> R.I. Environment	-0.3977	NA	NA
R.I. Joint All <--> R.I. Environment	-0.5913	NA	NA
R.I. Time (AEMs) <--> R.I. Space (MEMs)	0.3263	NA	NA
R.I. Environment ∩ Space (MEMs) <--> R.I. Space	-0.5183	NA	NA
R.I. Environment ∩ Time (AEMs) <--> R.I. Space	-0.3609	NA	NA
R.I. Space (MEMs) ∩ Time (AEMs) <-->RI.Space	-0.3008	NA	NA
R.I. Joint All <--> R.I. Space	-0.1541	NA	NA
R.I. Environment ∩ Time (AEMs) <--> R.I. Time (AEMs)	-0.0218	NA	NA
R.I. Space (MEMs) ∩ Time (AEMs) <--> R.I. Time	-0.2049	NA	NA
R.I. Joint All <--> R.I. Time (AEMs)	0.474	NA	NA
R.I. Environment ∩ Time (AEMs) <--> R.I. Environment ∩ Space (MEMs)	-0.6659	NA	NA
R.I. Space (MEMs) ∩ Time (AEMs) <--> R.I. Environment ∩ Space (MEMs)	-0.5368	NA	NA
R.I. Joint All <--> R.I. Environment ∩ Space (MEMs)	-0.3417	NA	NA
R.I. Space (MEMs) ∩ Time (AEMs) <--> R.I. Environment ∩ Time (AEMs)	0.6927	NA	NA
R.I. Joint All <--> R.I. Space (MEMs) ∩ Time (AEMs)	-0.0912	NA	NA

Table SI 3.IV: Results of Path Analyses fitted considering the "Mostly temporal" scenario. "NA" indicates non-applicable parameter estimation. "R.I. Joint All" is the amount of variation in the species data attributable to Environment ∩ Space (MEMs) ∩ Time (AEMs). In bold are the two most relevant pathways (larger standardized estimates) per moderators (Niche breadth and Dispersal ability) and endogenous variables (components of the variation partitioning approach)

Pathway	Std. Estimate	Std. Error	R-squared
Niche Breadth <-- SH/TH	**-0.9597**	0.0226	
Niche Breadth <-- Connectivity	-0.0204	0.0225	0.92
Niche Breadth <-- Spatial Structure	**-0.0944**	0.0226	
Dispersal Ability <-- SH/TH	**-0.9739**	0.0258	
Dispersal Ability <-- Connectivity	0.0092	0.0258	0.95
Dispersal Ability <-- Spatial Structure	**-0.0489**	0.0258	
R.I. Environment <-- SH/TH	-0.2863	0.0898	
R.I. Environment <-- Connectivity	0.019	0.024	
R.I. Environment <-- Spatial Structure	-0.3604	0.0198	0.87
R.I. Environment <-- Niche Breadth	**-0.5859**	0.0794	
R.I. Environment<-- Dispersal Ability	**-0.5833**	0.0695	
R.I. Space (MEMs) <-- SH/TH	**1.1738**	0.2158	
R.I. Space (MEMs) <-- Connectivity	-0.186	0.0577	
R.I. Space (MEMs) <-- Spatial Structure	0.2789	0.0475	0.61
R.I. Space (MEMs) <-- Niche Breadth	0.3517	0.1908	
R.I. Space (MEMs) <-- Dispersal Ability	**1.4976**	0.167	
R.I. Time (AEMs)<-- SH/TH	**-0.9475**	0.0916	
R.I. Time (AEMs)<-- Connectivity	0.0566	0.0245	
R.I. Time (AEMs) <-- Spatial Structure	-0.0933	0.0202	0.96
R.I. Time (AEMs) <--Niche Breadth	-0.1084	0.081	
R.I. Time (AEMs) <-- Dispersal Ability	**0.1361**	0.0709	
R.I. Environment ∩ Space (MEMs) <-- SH/TH	-0.3803	0.1288	
R.I. Environment ∩ Space (MEMs) <-- Connectivity	-0.0267	0.0345	
R.I. Environment ∩ Space (MEMs) <-- Spatial Structure	0.1555	0.0283	0.84
R.I. Environment ∩ Space (MEMs) <-- Niche Breadth	**-0.7419**	0.1139	
R.I. Environment ∩ Space (MEMs) <-- Dispersal Ability	**-0.5294**	0.0997	
R.I. Environment ∩ Time (AEMs) <-- SH/TH	**0.7365**	0.1201	
R.I. Environment ∩ Time (AEMs) <-- Connectivity	0.0409	0.0321	0.81
R.I. Environment ∩ Time (AEMs) <-- Spatial Structure	0.1929	0.0264	

R.I. Environment ∩ Time (AEMs) <--Niche Breadth	**0.9544**	0.1062	
R.I. Environment ∩ Time (AEMs) <-- Dispersal Ability	0.66	0.0929	
R.I. Space (MEMs) ∩ Time (AEMs) <-- SH/TH	**1.078**	0.1641	
R.I. Space (MEMs) ∩ Time (AEMs) <-- Connectivity	0.029	0.0439	
R.I. Space (MEMs) ∩ Time (AEMs) <-- Spatial Structure	0.1037	0.0361	0.68
R.I. Space (MEMs) ∩ Time (AEMs) <-- Niche Breadth	**0.8709**	0.1451	
R.I. Space (MEMs) ∩ Time (AEMs) <-- Dispersal Ability	-0.5634	0.127	
R.I. Joint All <-- SH/TH	**-0.8534**	0.3	
R.I. Joint All <-- Connectivity	-0.0279	0.0803	
R.I. Joint All <-- Spatial Structure	-0.2935	0.066	0.17
R.I. Joint All <-- Niche Breadth	-0.1682	0.2653	
R.I. Joint All <-- Dispersal Ability	**-0.3901**	0.2322	
Correlated Errors			
Niche Breadth<--> Dispersal Ability	0.3462	NA	NA
R.I. Space (MEMs) <--> R.I. Environment	0.1498	NA	NA
R.I. Time (AEMs) <--> R.I. Environment	-0.5352	NA	NA
R.I. Environment ∩ Space (MEMs) <--> R.I. Environment	-0.1065	NA	NA
R.I. Environment ∩ Time (AEMs) <--> R.I. Environment	-0.4334	NA	NA
R.I. Space (MEMs) ∩ Time (AEMs) <--> R.I. Environment	-0.1236	NA	NA
R.I. Joint All <--> R.I. Environment	-0.1193	NA	NA
R.I. Time (AEMs) <--> R.I. Space (MEMs)	0.4642	NA	NA
R.I. Environment ∩ Space (MEMs) <--> R.I. Space	-0.6688	NA	NA
R.I. Environment ∩ Time (AEMs) <--> R.I. Space	-0.7772	NA	NA
R.I. Space (MEMs) ∩ Time (AEMs) <-->RI.Space	-0.2832	NA	NA
R.I. Joint All <--> R.I. Space	-0.0047	NA	NA
R.I. Environment ∩ Time (AEMs) <--> R.I. Time (AEMs)	0.0276	NA	NA
R.I. Space (MEMs) ∩ Time (AEMs) <--> R.I. Time	-0.0298	NA	NA
R.I. Joint All <--> R.I. Time (AEMs)	0.1037	NA	NA
R.I. Environment ∩ Time (AEMs) <--> R.I. Environment ∩ Space (MEMs)	-0.6743	NA	NA
R.I. Space (MEMs) ∩ Time (AEMs) <--> R.I. Environment ∩ Space (MEMs)	-0.6418	NA	NA
R.I. Joint All <--> R.I. Environment ∩ Space (MEMs)	-0.336	NA	NA
R.I. Space (MEMs) ∩ Time (AEMs) <--> R.I. Environment ∩ Time (AEMs)	0.7693	NA	NA
R.I. Joint All <--> R.I. Space (MEMs) ∩ Time (AEMs)	0.1545	NA	NA

3.6.5 Empirical Support: Data information and analyses

Moth metacommunity in the MTC (data from Maicher et al. 2019)

Data publicly available taken from https://doi.org/10.5061/dryad.mgqnk98vr

Data cleaning:

Our analyses considered only moths that were collected using light traps (i.e., we filtered the data to remove butterflies and fruit-feeding moths). As such, out of the 1,099 species of Lepidoptera collected, our final analyses only considered 561 species (see MTC_Metacommunity.txt file). Information about the sampling design can be found in Maicher et al. (2019).

Species were sampled across 7 elevations (30, 300, 650, 1100, 1450, 1850, 2200m.a.s.l) in three different moments of their growing season between the years of 2014-2017. According to Table 1 in Maicher et al. (2019), the "Wet to Dry" season comprises the months of October to December; the Dry season goes from January to February; the Dry to Wet season goes from March to May. We used this information to extract the climate variables from (Karger et al. 2017) for each sample site between 2014-2017. The climate data used in further analysis (Climate_moth_MTC.txt) represent the monthly values of mean, maximum, minimum temperature, and precipitation per elevation averaged across 2014-2017 .

Moth metacommunity in the AEF (Miller and Jones 2005)

Data publicly available taken from
https://doi.org/10.6073/pasta/0cebe58bcc514e2bbf890ee7b2ea21c1

Data cleaning:

Moths were sampled in the AEF from 1994 to 2004. However, only in 2004, the year we used in our analyses, were moths systematically sampled in the totality of the growing season (from May to October). Moths were sampled in 12 plots (each with one or two sub-plots) along an elevational gradient ranging from 400 to 1400m.a.s.l. We summed all moths collected across sub-plots in a given moment in time to define species abundance in a community. Monthly climate data was extracted from the average latitude and longitude of subplots in a plot (Climate_moth_AEF.txt). No moths were collected in the plots in high elevations in May, so they were removed. In total, 367 species were considered in the final analyses (AEJ_Metacommunity.txt)

Reanalyzing empirical data considering elevational ranges of similar length

After removing samples in the MTC, we observed that the SH/TH decreased from 2.12 to 1.09, indicating that the strength of SH in relation to TH decreased when we reduced the total elevational range. However, it is important to note that the SH/TH remained positive, suggesting that even with similar elevational gradients, SH continued to outweigh TH in the tropical

mountain. This pattern aligns with Janzen's seasonality hypothesis and reinforces its applicability in the mountainous landscapes of our study system.

The differences in the metacommunity-weighted mean climate tolerance calculated for both species pools remained qualitatively the same. That is, the climatic tolerance of species in the AEF species pool is higher than the estimate for the MTC pool (AEF= 1.41, MTC =0.27).

As for the variation partitioning, we observed that the proportion of total explained variation attributable to "Purely Climate" decreased in the MTC, while the contribution of Space and Shared Climate and Space increased. However, it is worth noting that the relative importance of Purely Climate in the MTC was the second highest and remains qualitatively higher than what was observed in the AEF dataset.

Collectively, these results indicate that the predictions derived from our simulation models hold despite differences in the sampling design of both datasets
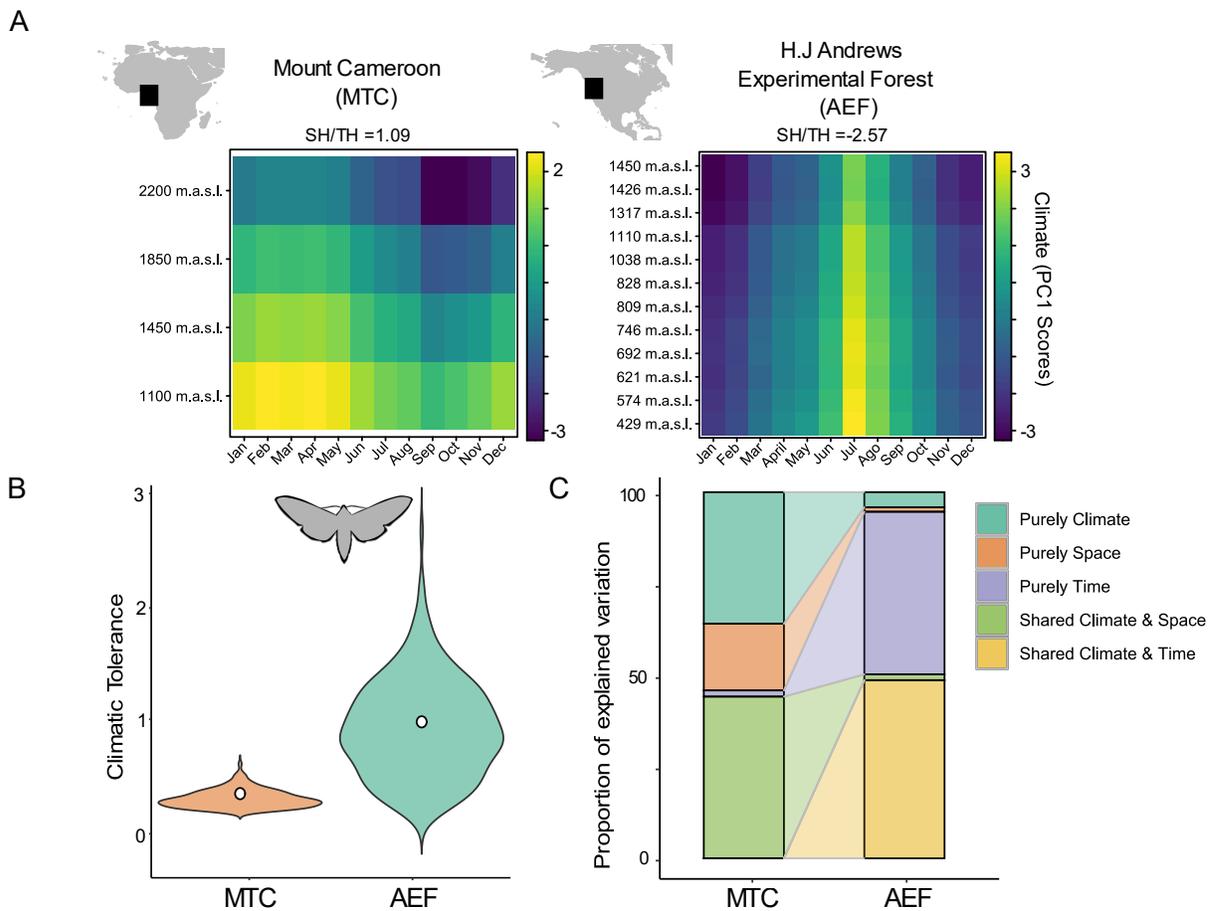
Figure SI. 3.V Results after reanalysing empirical data considering elevational ranges of similar size (approximately 1000m). Se more in main text and caption of Figure 5

Chapter 4 :Ecological selection of dispersal strategies in metacommunities: impact of landscape features and competitive dynamics

4.1 Abstract

Dispersal is simultaneously a cause and a consequence of metacommunity dynamics. While the influence of dispersal on metacommunities is subject of intense research, we still do not understand how species-species and species-environment relationships determine the success of different dispersal strategies in metacommunities. To address this, we employed simulation models considering species with distinct context-dependent dispersal strategies involved in the three stages of dispersal (departure, transience, and settlement). These species were allowed to reach coexistence at the metacommunity scale under various competitive hierarchies and different levels of spatial and temporal environmental variability. By assessing the dispersal strategies of species that persisted and dominated metacommunities, we could understand how metacommunity dynamics impose ecological selection on dispersal. Our simulation model reproduced empirical patterns in species dispersal across different scales, ranging from changes in the success of dispersal strategies caused by local intraspecific and interspecific competition, to observed shifts in dispersal strategies along broad-scale ecological gradients. Additionally, we derived new empirically testable predictions regarding how metacommunity dynamics select for different dispersal strategies. Collectively, our results foster a comprehensive understanding of the factors influencing the success and diversity of dispersal strategies in a large array of ecological contexts

4.2 Introduction

Dispersal is the ecological process where individuals depart from their natal patches, move across the landscape, and eventually establish themselves in breeding patches. This multi-stage process regulates the spatial and temporal dynamics of natural systems across all levels of ecological organization (Nathan et al. 2008; Kubisch et al. 2014; Bonte and Dahirel 2017). At the metacommunity level, dispersal governs species coexistence (Zhang et al. 2021) and influences community invasion success rates (Brown and Barney 2021). As a result, dispersal impacts diversity patterns within (alpha-diversity), between (beta-diversity), and across local communities (gamma-diversity).

Though much research has examined dispersal's influence on metacommunity dynamics and diversity (see Schlägel et al. 2020 and references within), our understanding of its role as a consequence of these dynamics is limited. For instance, density-dependent biotic interactions can regulate the decision of organisms to disperse or remain in their natal patches (De Meester et al. 2015; Fronhofer et al. 2015, 2018). Similarly, resource availability (Fronhofer et al. 2018), spatiotemporal heterogeneity (McPeek and Holt 1992; Büchi and Vuilleumier 2012), and landscapes' physical connectivity (Henriques-Silva et al. 2015) can impose ecological and evolutionary constraints on dispersal. The complex interplay between metacommunity dynamics

and dispersal remains underexplored, yet discerning the forces governing species dispersal is crucial to understand the potential impacts of global change on biodiversity (Urban et al. 2016).

To understand how and why metacommunity dynamics should influence dispersal patterns in metacommunities, we need first to acknowledge the multi-stage nature and context-dependence of dispersal strategies, which are intentionally simplified in the foundational framework of metacommunity theory (Büchi and Vuilleumier 2012; Thompson et al. 2020). Dispersal arises from balancing decisions involving the timing to leave natal patches (i.e., emigration propensity), travelling distances (i.e., traversal), and the selection of a suitable new patch to settlement (i.e., habitat selection). This sequence of decisions determine the three stages of dispersal events, namely departure, transience, and settlement (*sensu* Clobert et al. 2009). Dispersal decisions are context-dependent (plastic), meaning that organisms adjust them based on information about the surrounding biotic (e.g., predation, kin competition, and intra- and interspecific-competition) and abiotic (e.g., resource availability, spatiotemporal environmental variation) conditions (Bowler and Benton 2005). For instance, organisms are more propense to leave their natal patches when local performance (i.e., fitness) is reduced via strong competition (intraspecific or interspecific), predation, resource scarcity, or unsuitable abiotic conditions (Fronhofer et al. 2015; Campana et al. 2022). These context-dependent decisions underly changes in species' dispersal strategies to maximize regional fitness and/or minimize local mortality across different ecological contexts. Lastly, context-dependent variation in dispersal strategies are species-specific (De Meester et al. 2015; Fronhofer et al. 2018; Campana et al. 2022). Thus, even species that are evolutionary closely related may still exhibit contrasting changes in dispersal strategies when subjected to varying abiotic and biotic conditions (De Meester et al. 2015; Campana et al. 2022).

Given that metacommunity dynamics result directly from species-species and species-environment interactions, they should favor species from the regional pool whose dispersal strategies maximizes their persistence and dominance in the landscape (Büchi and Vuilleumier 2012). For instance, species exhibiting a dispersal strategy characterized by a rapid increase in emigration propensity when local performance is decreased should have an advantage in persisting and dominating metacommunities in landscapes that undergo temporally variable habitat conditions (McPeek and Holt 1992). Species that adopt a "risk-spreading" strategy, in which individuals choose to colonize suboptimal patches that have the potential to become optimal in the (relatively) short term, are expected to be favoured by temporal environmental variability. In metacommunities where competition dynamics hinder local coexistence (i.e., heterospecific competition is stronger than intraspecific competition, see Chesson 2000), species capable of reaching suitable habitat patches ahead of competitors have the potential to establish regional dominance through residency effects (*sensu* Kemp and Wiklund 2004).

Testing these (and potentially other) predictions about the ecological selective pressures of metacommunity dynamics on context-dependent dispersal strategies remains challenging for multiple reasons. Broad-scale observational data on multi-species dispersal strategies are scarce and experimental studies are commonly constrained by the number of species and environmental

predictors that can be manipulated (but see De Meester et al. 2015; Campana et al. 2022; Cote et al. 2022). Moreover, these studies are often conducted along a narrow range of ecological conditions, restricting their ability to capture the large range of conditions that could potentially drive variation in the success of distinct (context-dependent) dispersal strategies.

Here we sought out to expand our understanding about how landscape features and competition (ecologically) select for distinct context-dependent dispersal strategies in metacommunities (Figure 4.1). We employed process-based metacommunity models to assess how the interactions between landscape features and competition (i.e., metacommunity dynamics) select for dispersal strategies involving emigration propensity, habitat selection, and traversal (i.e., travelling distance). We allowed species with distinct context-dependent dispersal strategies to reach coexistence at the metacommunity scale under different types of competition types and under varying levels of spatiotemporal environmental variability (Figure 4.1). By assessing the dispersal strategies of the species that persisted and dominated metacommunities, we were able to derive informed predictions regarding how metacommunity dynamics impose ecological selection on dispersal. Moreover, we demonstrated how the integration of species-specific context-dependent dispersal strategies into the basis of metacommunity theory can help us to understand the interdependence between community assembly and dispersal.

4.3 Methods

For the sake of brevity, we only briefly describe how we simulated landscapes and within-patch metacommunity dynamics here. An extended description is found in Supp. Information.

4.3.1 Simulated landscapes

We generated 25 landscape types following a cross-factorial design combining 5 levels of spatial autocorrelation (top-left panel in Figure 4.1) 5 levels of seasonality (top-right panel in Figure 4.1) in environmental conditions. We chose these two landscape features because they have been shown to impose costs and risks to species movement dictating ecological and evolutionary constraints on dispersal (McPeek and Holt 1992; Büchi and Vuilleumier 2012). Each landscape type was composed of 50 habitat-patches with randomly generated x and y spatial coordinates (see Supp. Information I for more details).

4.3.2 Parametrizing competitive dynamics in metacommunities

We seeded landscapes with species pools of different sizes (richness; 2 levels: 50 and 300 species) and assumed distinct types of competition among species (bottom-left and bottom-right panels in Figure 4.1). Species pool sizes affect metacommunity competition strength, as larger pools intensify competition for limited patches. We considered different types of competition by manipulating the per-capita effects of species on themselves (α_{intra}) and their per-capita effect on other species (α_{inter}) (Thompson et al. 2020; Wisnoski and Shoemaker 2022). Stabilizing competition ($\alpha_{intra} > \alpha_{inter}$) promotes local stable coexistence, enabling rare species to growth positively when populations of dominant competitors are in equilibrium. Under equalizing

competitive dynamics ($\alpha_{intra} = \alpha_{inter}$), individuals are neutral with respect to their per-capita
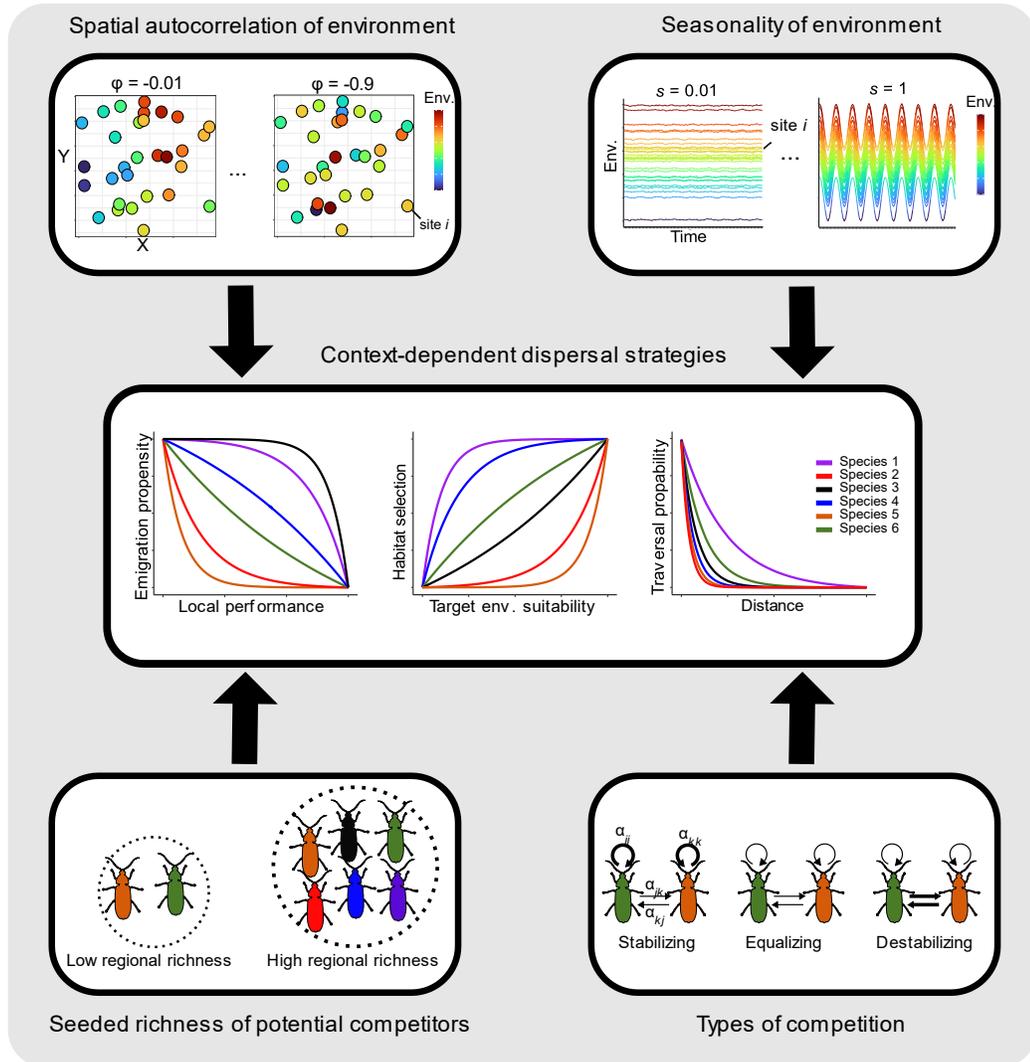
Figure 4.1: Simulation framework designed to understand how metacommunity dynamics impose ecological selection on context dependent dispersal strategies. Species were generated assuming random combinations of distinct dispersal strategies (i.e., represented by the plotted curves) for emigration propensity, habitat selection, and traversal probability (central panel). Species were allowed to colonise and reach coexistence in metacommunities subjected to different competitive dynamics (given by the factors represented in the lower panels) that took place in landscapes with different features (given by the factors represented in the upper panels). Parameters φ and s determine the landscape's spatial structure in environmental conditions and seasonality, respectively (See more in Supp. Information I). The size of regional pools gives the richness of potential competitors at the beginning of each simulation iteration. Variation in competition type was simulated by manipulating the per-capita effects of a species on itself (intraspecific competition α_{kk}, α_{jj}) and on other species (interspecific competition α_{kj}, α_{jk}). Here we considered species pools under stabilizing ($\alpha_{kk} = \alpha_{jj} > \alpha_{kj} = \alpha_{jk}$), equalizing ($\alpha_{kk} = \alpha_{jj} = \alpha_{kj} = \alpha_{jk}$), and destabilizing ($\alpha_{kk} = \alpha_{jj} < \alpha_{kj} = \alpha_{jk}$) competition.

effects on each other. This implies that coexistence is only possible for species having different

niche requirements and/or display contrasting dispersal strategies. Under destabilizing competition ($\alpha_{intra} < \alpha_{inter}$), local coexistence is unlikely because dominant species will lead rare species to local extinction even if those are better adapted to patch conditions.

4.3.3 Metacommunity dynamics

Our model simulates metacommunity dynamics that are spatially explicit, discrete in time, and governed by within-patch selection (density-dependent competition at the intraspecific and interspecific levels and density-independent species-environment sorting), dispersal, and ecological drift (as in Thompson et al. 2020). Within-patch selection was modelled as a Beverton-Holt growth model (Beverton and Holt 1957) with generalized Lotka-Volterra competition assuming distinct competitive structures (i.e., stabilizing, equalizing, or destabilizing). An extended description of the model is found in Supp. Information (see eq. SI 4.1 and SI 4.2).

Individuals able to persist in any given local community after within-patch selection and drift at time *t* could then disperse. Context-dependence in dispersal strategies was introduced by making species emigration propensity (EP), traversal probability (TP), and habitat selection (HS) to change as a function of local performance (given by joint influence of competition and niche-habitat matching), geographic distance, and environmental suitability, respectively. The shape of these relationships was made species-specific by randomly assigning different values of parameters *ep*, *tp*, and *hs* to each species in the regional pool.

Emigration propensity (EP) defines the probability that an individual leaves its natal patch given its current local performance. Based on previous experimental studies (e.g., (Fronhofer et al. 2018)), we assumed that species were more propense to emigrate when local habitat conditions and/or biotic interactions decreased their local performance ($P_{i,j,t}$, given by *eq SI-2*). As such, the emigration propensity of species *i* in site *j* at time *t* ($EP_{i,j,t}$) decreased with local performance as follows:

$$EP_{i,j,t} = u_i * exp\left(P_{i,j,t} * ep_i\right) \qquad (4.1)$$

where ep_i is species-specific (equally spaced values within the interval [-0.1,0) U (0,0.1] that were then randomly assigned to each species), and determined the rate of change in $EP_{i,j,t}$ as a function of $P_{i,j,t}$ (scaled to range between 0 and 100). u determines the concavity of the relationship and was defined as:

$$u = \begin{cases} -1 & if\ ep_i > 0 \\ 1 & if\ ep_i < 0 \end{cases}$$

If $ep_i < 0$ (and consequently $u = 1$), emigration propensity steeply decreased even at low levels of local performance (i.e., as observed in species 2, 5, and 6 in Figure 4.1, central panel). This strategy can be advantageous when local performance is temporally stable, allowing for competitive advantage due to residency effects. In contrast, species for which $ep_i > 0$ ($u = -1$) were prone to emigrate from patches even if those provided high local performance (as observed

in species 1, 3, and 4 Figure 4.1, central panel). This strategy can be advantageous when local performance changes abruptly due to environmental fluctuations. $EP_{i,j,t}$ (which is scaled to range between 0 to 1) set the probability of success in binomial trials determining the number of emigrants of species i departing from site j at time t ($E_{i,j,t}$) after within-patch dynamics (given by the first term of *eq. SI-4.1*).

We relied on a random sampling process to determine the total number of immigrants of species i that will arrive at patch j at time t ($I_{i,j,t}$) coming from other patches, e.g., patch k ($E_{i,k,t}$). That is, let D = {$x, y, ..., j$} be the potential destination patch for emigrants that departed from k. Let P = {$p_{i,kx,t}, p_{i,ky,t}, ..., p_{i,kj,t}$} be the set of probabilities for species i to immigrate to each patch in D when departing from k at time t (scaled to sum to unit). Note that $p_{i,kk,t}$ was set to 0 so that individuals that departed from k could not return to the same patch. Considering these two vectors, a random sampling process with unequal probabilities and replacement was repeated $E_{i,k,t}$ times. The number of times patch j was sampled determined the number of individuals out of the total that departed from k ($E_{i,k,t}$) that immigrated to patch j at time t ($I_{i,kj,t}$). It follows that $I_{i,j,t}$, i.e, the total number of individuals of species i that immigrated to j in time t, was then given by the sum of the total number of immigrants coming from all patches in the landscape.

We assumed that immigration probabilities increased with habitat suitability in the extant patch and decreased with the geographic distance between natal and extant patches. Thus, $p_{i,kj,t}$ was given by:

$$p_{i,kj,t} = HS_{i,j,t} * TP_{i,kj} \quad (4.2)$$

$HS_{i,j,t}$ represents species i probability to move to patch j at time t based on the match between the environment in j ($Env_{j,t}$) and their environmental requirements (i.e., *environmental suitability*, computed in the second term of *eq. SI. 4.2*). We assumed that species could assess the environmental suitability of extant patches at time t before deciding where to immigrate (i.e., informed dispersal). As such $HS_{i,j,t}$ increases with habitat suitability as follows:

$$HS_{i,j,t} = w * exp(environmental\ suitability * hs_i) \quad (4.3)$$

where hs_i is species-specific (equally spaced values within the interval [-0.1,0) U (0,0.1] that were then randomly assigned to each species) determining the rate of change in $HS_{i,j,t}$ with *environmental suitability* (scaled to range between 0 and 100). w set the concavity and direction of this relationship and was defined as:

$$w = \begin{cases} -1 & if\ hs_i < 0 \\ 1 & if\ hs_i > 0 \end{cases}$$

As such, species with $hs_i < 0$ (and consequently a $w = -1$) displayed a risk-spreading strategy as they also tended to immigrate to patches with suboptimal habitat conditions (e.g., species 1, 4, and 6 in Figure 4.1, central panel). This strategy can be advantageous in landscapes

with temporally variable habitat conditions. Conversely, species having $hs_i > 0$ (and consequently a $w = 1$) tended to immigrate only to patches with optimal or close to optimal environmental conditions (e.g., species 2, 3, and 5 in Figure 4.1, central panel). This strategy can be advantageous when environmental conditions are temporally constant across all patches.

$TP_{i,\text{kj}}$ is the probability of species i to move from patch k to j according to their geographic distance. $TP_{i,\text{kj}}$ decayed with the (Euclidean) distance between k and j ($dist_{kj}$)as follows:

$$TP_{i,kj} = exp(tp_i * dist_{jk}) \quad (4.4)$$

where tp_i is species-specific (equally spaced values within the interval [-0.99, -0.01], that were then randomly assigned to each species) and determined the rate at which $TP_{i,kj}$ decayed with $dist_{kj}$. The smaller the tp_i, the more limited the number of nearby patches a species could reach in a single dispersal event (e.g., species 2, 4, and 5, Figure 4.1 central panel). High traversal capacity (e.g., species 1, 3, and 6) should be favored in landscapes where local habitat conditions are temporally variable, weakly autocorrelated in space, and when intraspecific competition is strong.

Lastly, we repeated the simulation framework described above considering two distinct assumptions about species ecological equivalence (see Supp. Information for details). In one, species were set to exhibit distinct (non-neutral) performances along the environmental gradient. Then, species niche optima μ_i took values equally spaced along the interval [0,5] to scale with environmental variation. Note that niche tolerance (σ) was set to be equal for all species and narrow enough to make them respond to environmental variation (see Supp. Information). In the second assumption, we generated species to have equal (neutral) species' responses (changes in performance) to environmental variation. This was operationalized by assigning the same environmental optima to all species (μ_i = average value of environmental conditions observed in the generated landscape). By contrasting simulation outputs between these two assumptions, we could investigate how species ecological equivalence can modulate (either buffer or amplify) the ecological selection of metacommunity dynamics on successful dispersal strategies.

4.3.4 Simulation iterations

Metacommunity dynamics were set to run for 1200-time steps (100 seasonal cycles). Each combination of scenarios was replicated 20 times, totalizing 6000 simulation iterations (20 replicates × 5 seasonality levels × 5 spatial structure levels × 2 species pool sizes × 3 competitive structures × 2 types of responses to environmental variation). At the first-time step (*t1*), we seeded each patch with species abundances randomly drawn from a Poisson distribution ($\lambda = 0.5$). Then, during the first ten seasonal cycles of each iteration (i.e., the first 120 time-steps), we seeded each patch with species abundances randomly drawn from a Poisson distribution with $\lambda = 0.1$. This seeding procedure ensured that species had equal chances to be initially present in all patches and that patches with similar environmental conditions could harbour different communities over time due to priority effects (Thompson et al. 2020). Metacommunity dynamics (with no seeding) ran

for the remaining 1080 time-steps (i.e., 90 seasonal cycles) and we considered the metacommunity in the 100th seasonal cycle (last 12 time-steps) for our analyses (see below). This ensured that model summaries were based on stable rather than transient metacommunities.

We estimated the most successful dispersal strategy as the metacommunity-weighted mean values of *ep*, *hs,* and *tp,* where the weights were given by a species' regional abundance × occupancy (i.e., the relative number of patches occupied in the landscape). We also estimated the metacommunity-weighted standard deviation of *ep*, *hs,* and *tp* to quantify the "diversity" of dispersal strategies that resulted from metacommunity dynamics.

4.3.5 Understanding how landscape features and competition dynamics select for dispersal strategies in metacommunities

We used random forest to assess how the metacommunity-weighted mean and standard deviation of *ep*, *hs*, and *tp* (continuous response variables) changed as a function of variation in seasonality (ordinal predictor, 5 levels), spatial autocorrelation (ordinal predictor, 5 levels), the seeded richness of competitors (ordinal predictor, 2 levels),   different types of competition (categorical predictor, 3 levels), and different assumptions about species niche differentiation (categorical, 2 levels). Random forests identify general patterns in cross-factorial simulation data as they automatically model the effects of multilevel interactions among predictors on the response. We used the Boruta algorithm (Kursa and Rudnicki 2010) to reduce model dimensionality and identify the most relevant predictors explaining variation in the metacommunity-weighted mean and standard deviation of *ep*, *hs*, and *tp*. Partial dependence plots revealed the direction of predictor-response relationships, controlling for other model predictors. We used bootstrapping to estimate 95% confidence intervals of predicted responses depicted in the partial dependence plots (Ishwaran and Lu 2019).

4.4 Results and Discussion

Landscape features and competition dynamics had complex interactive effects on the success and diversity of dispersal strategies (i.e., metacommunity-weighted mean and standard deviation of *ep, hs, and tp*, respectively*)*. Overall, the emigration propensity of the dominant species in the metacommunity decreased relatively fast with initial increases in local performance (i.e., $ep < 0$). Additionally, dominant species were highly selective for optimal conditions in destination patches (i.e., $hs > 0$) and could traverse large distances in a single dispersal event (i.e., ts > -0.5) (Figure 4.2). Note though that optimal context-dependent dispersal strategies and their diversity changed consistently across levels of seasonality, environment' spatial structure, competitors' seeded richness, and competition types (Figure 4.3 and Figure 4.4). For the sake of tractability and synthesis, we focused on reporting and discussing the main effects of landscape features and competition types on the success and diversity of dispersal strategies. However, we also highlight and discuss some high-order interactions among predictors that increased our understanding of model outcomes (Figures SI. 4.I-IV).

4.4.1 The effects of landscape features on the success and diversity of dispersal strategies

In temporarily homogeneous (aseasonal) but spatially heterogeneous landscapes, species with lower average emigration rates (lower values of *ep*) were more successful in persisting and dominating metacommunities (Figure 4.3 and Figure SI. 4.I). This because high emigration from temporally stable and high- performing patches can negatively affect species' persistence, as reduced local abundances increases risks of local extinction from demographic stochasticity and competitive exclusion (Siqueira et al. 2020). In contrast, when habitat conditions were seasonal, metacommunity dynamics favoured nomadic behaviours characterized by high emigration rates despite high levels of local performance. These "nomad" species were particularly favoured when they were also able to colonize several extant patches in single dispersal events (i.e., high *tp* values, Figures 4.3 and SI. 4.I). Essentially, species with high emigration rates and traversal capacity can rapidly shift their spatial distribution to cope with abrupt changes in local performance due to spatial and temporal variability in habitat conditions (McPeek and Holt 1992; Sheard et al. 2020).

Empirical studies investigating latitudinal clines in species dispersal observed a similar influence of seasonality in selecting for higher emigration rates and traversal capacities. For instance, the relationship between seasonality and dispersal observed in our model serves as theoretical evidence that the conditions for the emergence of well-known latitudinal patterns on species range sizes and dispersal capacity (e.g., Sheard et al. 2020; Alzate and Onstein 2022) can emerge by only considering metacommunity dynamics at fine spatiotemporal scales (i.e., no need to consider trait evolution and speciation, but see conclusions).

Note that when we assumed equivalent species' environmental responses to environmental variation (i.e., neutrality), the relationship between seasonality and metacommunity-weighted mean *ep* and *tp* remained positive (Figures 4.3 and SI 4.I). However, under this assumption, seasonality had a negative effect on the metacommunity-weighted mean *hs*. This implies that when species are neutral in relation to their habitat requirements, dispersing to temporally unsuitable patches, despite the risks, allows their persistence and dominance in seasonal landscapes. This is because colonizing briefly suboptimal patches enable them to broaden their spatial occupancy while avoiding intense competition.

In our model, the effect of spatial autocorrelation fostered insights into the influence of spatial uncertainty in habitat conditions on species' dispersal strategies (Figures 4.3 and SI 4.I). Species capable of minimizing dispersal risks by effectively tracking habitat conditions and reaching a larger number of patches (i.e., higher *hs* and *ts* values) were more successful in landscapes with weak spatial autocorrelation. In contrast, when habitat conditions were strongly autocorrelated in space, less selective species exhibiting weaker traversal capacity (i.e., lower *hs* and *ts* values) were more likely to persist in the metacommunity. Spatial autocorrelation modulates the respective success of dispersal strategies by determining the costs and risks of dispersal events.
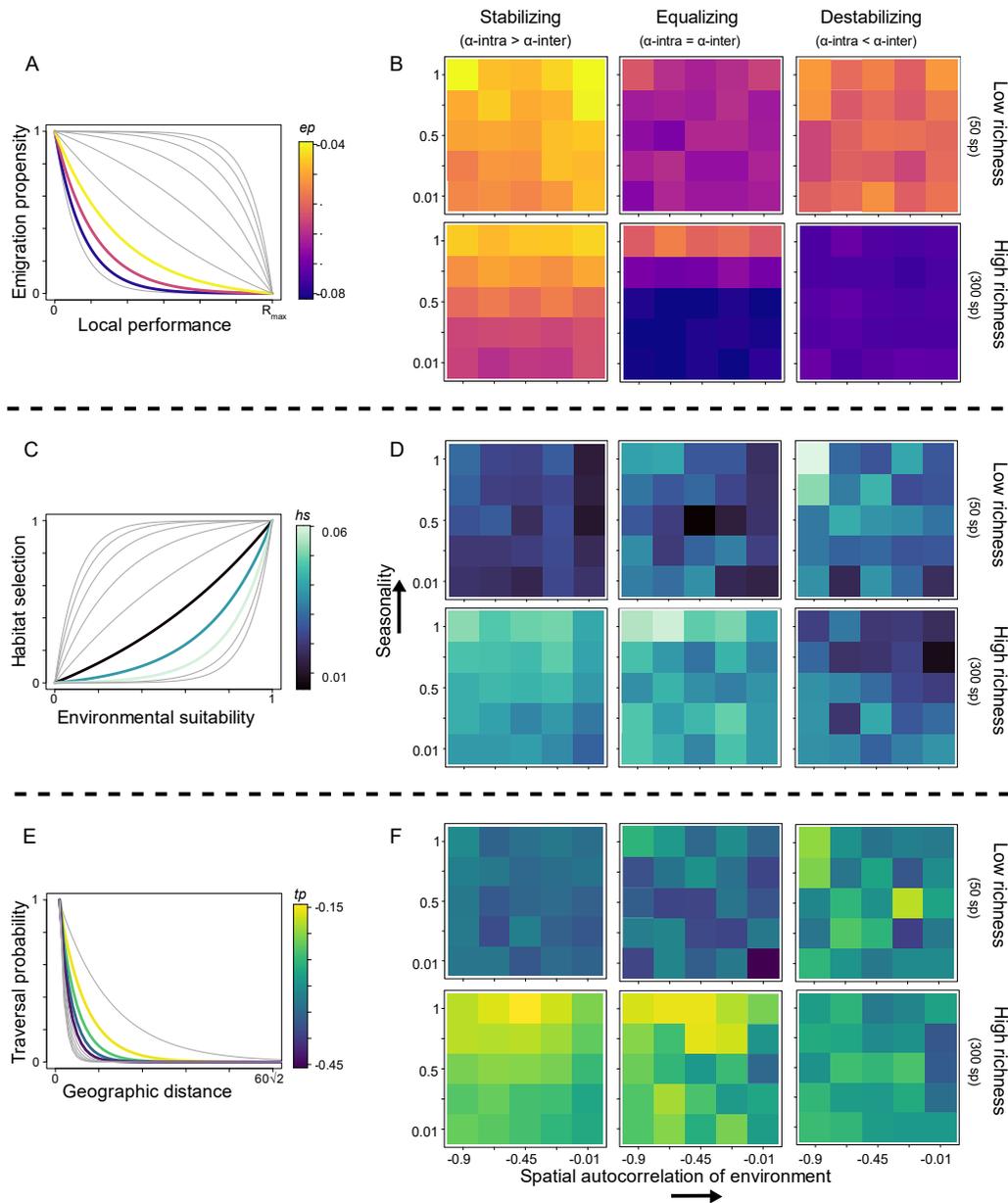
Figure 4.2 Dominant dispersal strategies for emigration propensity, habitat selection, and traversal probability observed across simulation scenarios. The curves in plots A, C, and E illustrate a small subset of the full range of context-depedent dispersal strategies seeded into metacommunities at each simulation iteration. The shape of these curves depends on the species-specific parameters ep, hs, and ts. The colour scales indicate the range of dispersal strategies that have dominated metacommunities at the end of each simulation iteration (estimated as the metacommunity-weighted mean values for hs, ep, ts). For illustrative purposes, some of these strategies are represented by the coloured curves in plots A, C, and E. The colour scales also serve as a reference for the heatmaps (B, D, and F) illustrating the changes in dominant context-dependent strategies across levels of seasonality, spatial autocorrelation, competitors' richness, and competitive types. Each entry (square) in the heatmap represents the average value of the 20 metacommunity-weighted means obtained for a given simulation scenario. The results reported here were obtained under the assumption of niche differentiation (i.e., simulated species differed in habitat requirements). See Figs 4.3, SI-4.I, and S-4.III for results obtained under the assumption of neutrality.

In landscapes with high spatial autocorrelation, the benefits of strong habitat selectivity and high

traversal capacity diminish, as there are fewer unsuitable patches accessible from suitable ones. Our results align with previous theoretical models, which observed that strong spatial autocorrelation favoured species with reduced traversal capacity, while weak autocorrelation favored species with strong traversal capacity (Büchi and Vuilleumier 2012). They also support empirical findings demonstrating that species adopt passive dispersal, a strategy characterized by weak habitat selection, when landscape structure is characterized by large clusters of patches with suitable habitat conditions (Bonte et al. 2006).

It is important to note that when the niche differentiation assumption was relaxed, spatial autocorrelation had negligible effects on selecting for optimal traversal capacity strategy (Figures 4.3 and SI 4.I). In our model, the neutrality assumption increases the number of species competing for a limited number of patches having equally suitable habitat conditions. As discussed below, a rise in species competing for similar habitat conditions favoured those with greater traversal capacity that enable them to escape local competition. Consequently, the selection for high-capacity traversers in highly competitive metacommunities offsets the selection of species with low traversal capacity in spatially structured landscapes (Figure SI. 4.I).

The diversity of dispersal strategies (weighted standard deviation of *ep*, *hs,* and *tp,* Figure 4.4*)* also changed as a function of landscape seasonality and spatial autocorrelation (Figures 4.4 and SI 4.II). Seasonality decreased the diversity of strategies related to habitat selection and traversal (i.e., the metacommunity-weighted standard deviation of *hs* and *tc,* respectively)*.* Only highly selective species with long-distance traversal capacities could persist when local conditions fluctuated substantially over time. Alternatively, the diversity of emigration propensity strategies (i.e., metacommunity-weighted standard deviation of *ep*) increased with seasonality. Thus, although seasonality tended to favour species with relatively higher emigration rates, species with lower emigration rates could persist provided they accumulated enough individuals to buffer mortality in temporally unsuitable conditions (i.e., storage effects Chesson 2000).

Moreover, spatial autocorrelation in habitat conditions notably diversified traversal capacity and habitat selection strategies (Figures 4.4 and SI. 4.II). This is because large clusters of suitable habitat conditions decrease the risks of dispersing to unsuitable habitats which, in turn, facilitates the persistence of species with suboptimal strategies for traversal and habitat selection in the metacommunity.

4.4.2 The effects of competition dynamics on the success and diversity of dispersal strategies

Research in metapopulation and movement ecology extensively explores how density-dependent processes (Fronhofer et al. 2018; Baines et al. 2020), including intraspecific (Bitume et al. 2014) and inter-specific competition (Fronhofer et al. 2015) influence dispersal strategies. Consistent with these, our results indicate that increases in competition associated with larger species pools sizes favoured species with (i) reduced emigration propensity (lower *ep* values); (ii) pronounced selectivity towards habitat patch condition (i.e., higher *hs* values), and; (iii) strong

traversal capacity (i.e., higher *hs* values) (Figure 4.3). These results held true across assumptions about niche differentiation and all types of competitive dynamics (Figure SI 4.III)
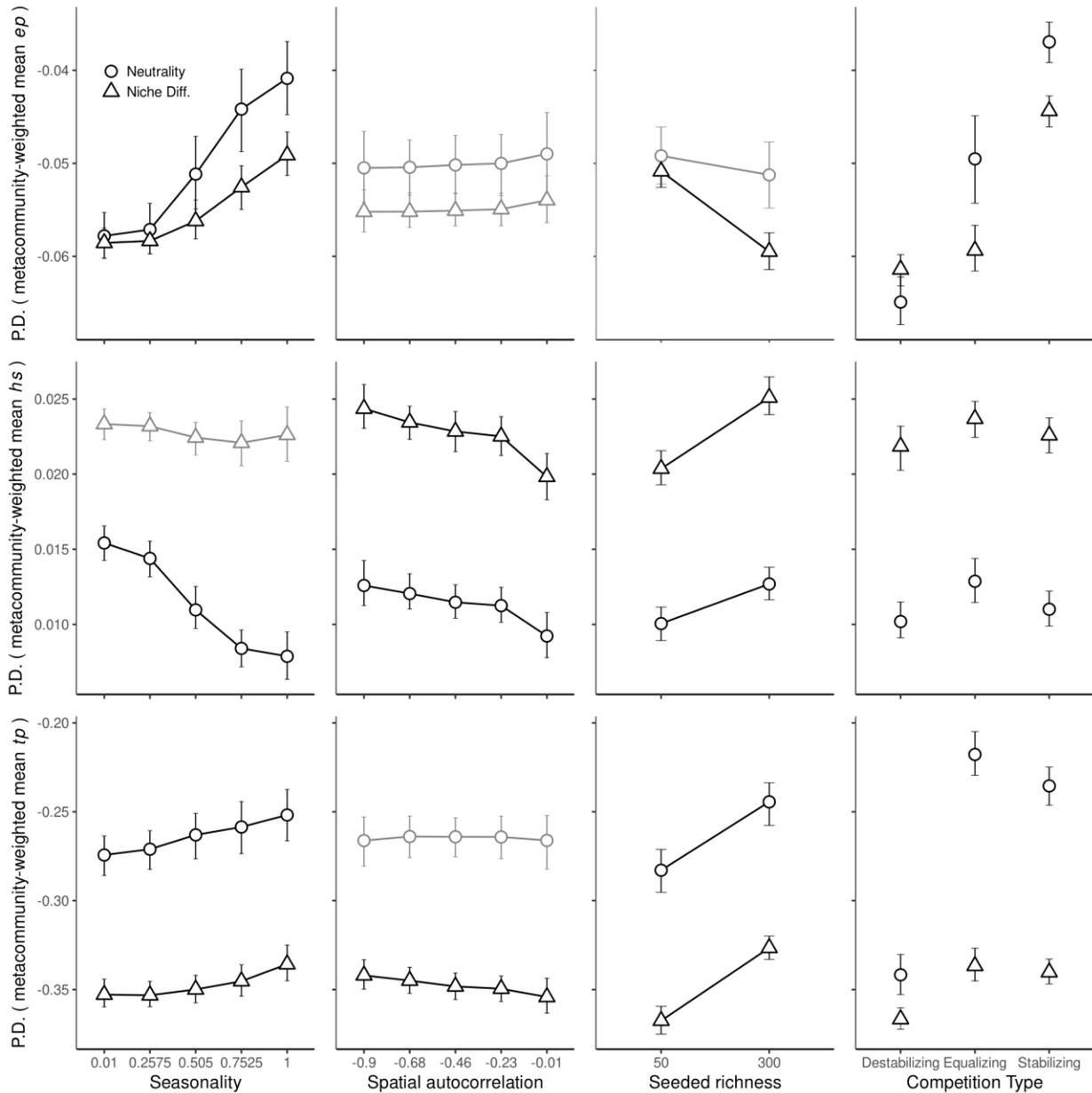
Figure 4.3 Partial dependence (PD) plots showing the predicted levels of dominant dispersal strategies in metacommunities (i.e., metacommunity-weighted mean values for ep, hs, ts) across levels of seasonality (first column), spatial autocorrelation in environmental conditions (second column), size of seeded regional pools (third column), and types of competition. Relationships were estimated for each level of the niche differentiation assumption (Neutrality = simulated species shared habitat requirements; Niche Diff. = simulated species differed in habitat requirements). Relationships of with variables that were not kept in the final random forests after feature selection were reported for illustrative purposes only (gray). R2 of random forest model fitted per row: metacommunity-weighted mean ep = 0.90, hs = 0.81, tp =0.95.

Species with reduced emigration propensity were favoured at higher species richness levels because they ensured competitive dominance by maintaining large local populations (Figures 4.3 and SI 4.III). The effectiveness of this strategy was maximized under destabilizing competition but minimized under stabilizing competition. Indeed, when intraspecific competition is stronger
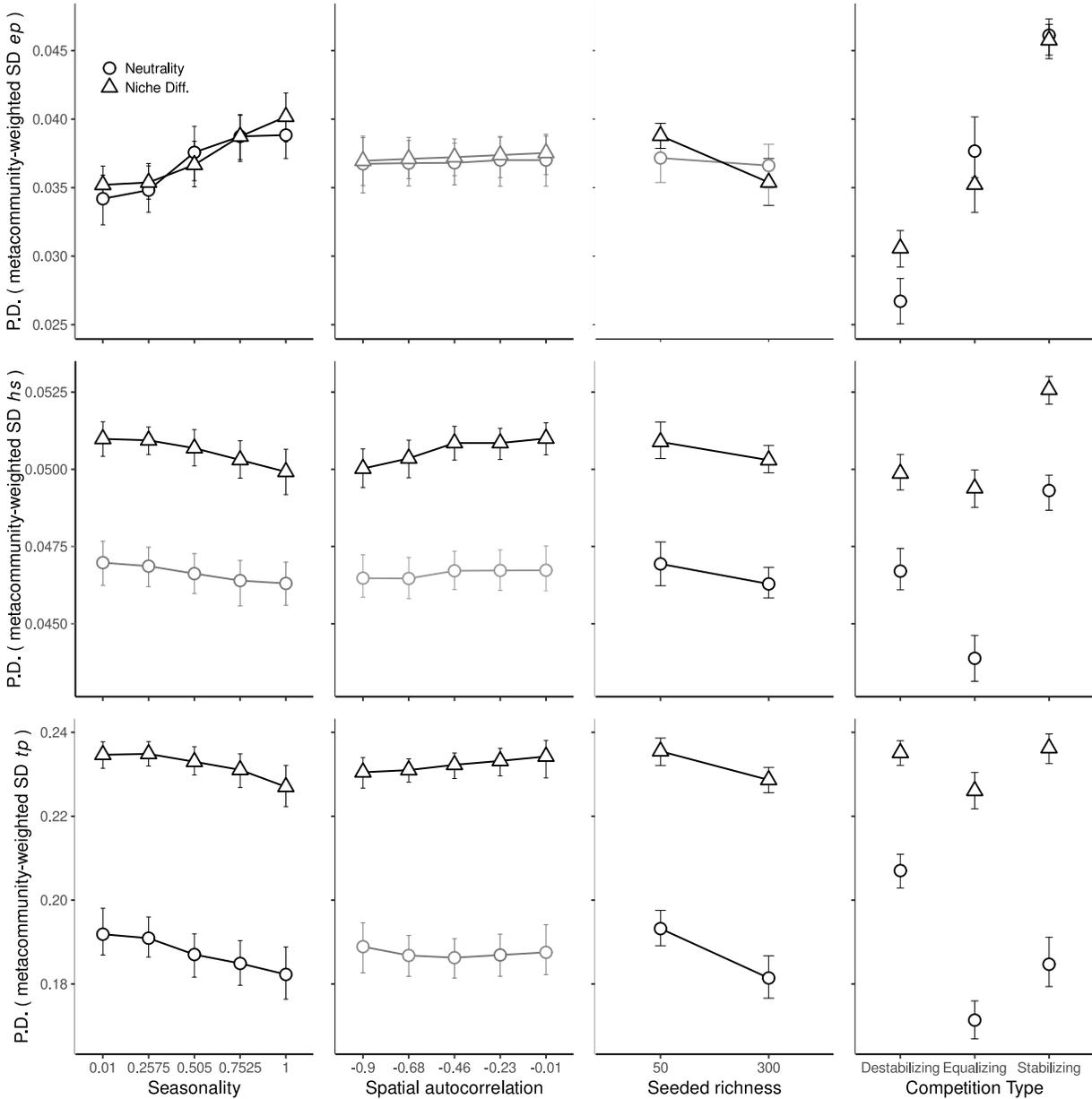
Figure 4.4: Partial dependence (PD) plots showing the predicted diversity of dispersal strategies (i.e., metacommunity-weighted standard-deviation for ep, hs, ts) across levels of seasonality (first column), spatial autocorrelation of environmental conditions (second column), size of seeded regional pools (third column), and types of competition. Relationships were estimated for each niche differentiation assumption (Neutrality = species shared habitat requirements; Niche Diff. = species' habitat requirements differed). Relationships of ep, hs, and ts with variables that were not kept in the final random forests after feature selection were reported in gray for illustrative purposes only. R2 of random forest model fitted per row: metacommunity-weighted standard deviation ep = 0.89, hs = 0.75, tp =0.9.

than interspecific competition, species with higher emigration propensity were able to mitigate the negative effects of intraspecific competition while keeping smaller viable populations in a larger number of suitable-habitat patches.

The success of dispersal strategies characterized by high traversal capacity increased with the initial number of competitors seeded in the metacommunity. This strategy had even greater success when intraspecific competition matched or surpassed interspecific competition (i.e., under equalizing and stabilizing competition, respectively, Figure SI 4.III). Such trends resonate with empirical and theoretical studies that examine how population density, an indicative of intraspecific competition, impacts dispersal. Typically, these studies demonstrate that high densities increase emigration rates, particularly to patches farther away from their natal patch (see Matthysen 2005 and references within, and Bitume et al. 2014). Our models shed deeper light on this dynamic, demonstrating that the relative strength of interspecific to intraspecific competition can either intensify or mitigate the latter's impact on emigration rates and traversal.

We also observed that increasing the number of competitors favoured species that were more efficient in tracking and colonising suitable habitat conditions (higher hs values, Figure 4.3). This efficiency was even greater under neutral competition dynamics (equalizing competition, Figure. SI 4.III). Under equalizing competition, only species capable of tracking the limited number of patches where adequate niche-environment matching outweighs the negative effects of competitive interactions can coexist in the landscape. Taken together, species that were equal competitors but highly selective towards different habitats could coexist regionally through species-environment sorting dynamics.

Overall, the diversity of context-dependent dispersal strategies in the metacommunity decreased with the number of initial competitors seeded in the landscape (Figures 4.4 and SI 4.IV). Thus, when a larger number of species competed for a few suitable patches, only a narrow range of dispersal strategies could ensure their regional persistence. Notably, the diversity of traversal capacity and habitat selection strategies ($tp,$ and hs) did not follow this trend under destabilizing competition. In this case, only species with differences movement patterns in the landscape were able to coexist at the metacommunity scale by dominating distinct clusters of suitable patches (Zhang et al. 2021). In contrast, the diversity of dispersal strategies tended to increase when coexistence was facilitated through stabilizing competition.

4.5 Conclusions, assumptions, and future directions

In this study, we used models to demonstrate that dispersal not only shapes the structure of metacommunities but also emerges from metacommunity dynamics. Previous theoretical and empirical studies that shared similar goals focused on investigating the ecological drivers of fixed behaviours involved in one or two stages of dispersal (Büchi and Vuilleumier 2012). Our study stands out as the first to use metacommunity theory to generate predictions regarding the selective effects of landscape features and competition dynamics on species-specific context-dependent dispersal behaviours involved in all three dispersal stages (i.e., departure, transience, and

settlement). Our models effectively recreated well-known variations in dispersal patterns across spatial scales, including changes caused by different forms of intraspecific and interspecific competition at local spatial scales and also shifts in dispersal patterns along broad-scale ecological gradients. Thus, our study improves the understanding of the factors influencing the success and diversity of dispersal strategies in a large array of ecological contexts.

However, our simulation models did not encompass the full complexity of species dispersal and its intricate relationships with metacommunity dynamics. For instance, we did not consider cost-related trade-offs that can cause covariance between dispersal, morphological, and behavioural traits. For instance, colonization-competition and ecological specialization-dispersal trade-offs can emerge as eco-evolutionary consequences of community assembly in landscapes with varying levels of environmental stability and habitat heterogeneity (Chesson 2000; Jocque et al. 2010). Therefore, we should expect these dispersal trade-offs to also be selected by the metacommunity dynamics.

Moreover, we focused solely on how competition at both intra and interspecific levels affects dispersal patterns in metacommunities. Yet, empirical experimental studies suggest that other biotic interactions can select for optimum context-dependent dispersal strategies. For example, predation risk can drive emigration (Fronhofer et al. 2018),while parasitism can have dual effects on host dispersal: it can stimulate movement if the host perceives threat and relocates, or it can inhibit movement if the host stays and becomes infected (Baines et al. 2020).Incorporating these and other biotic interactions into our framework is a logical progression for future research.

Lastly, in our model, although we investigated how biotic and abiotic factors selected for a wide range of predefined traits (*ep, hs* and *tc*) that determine species-specific dispersal strategies, we did not consider trait evolution at the species level. Trait evolution is an important component of metacommunity theory (Urban et al. 2008; Goodnight 2011) and the literature includes examples where the role of dispersal for maintaining biodiversity in changing environments is counteracted by the evolution of traits determining species' habitat requirements (Thompson and Fronhofer 2019). Moreover, we also did not consider the full range of possible non-linear changes in dispersal strategies commonly observed in nature. For instance, we assumed that emigration propensity decreased monotonically and continuously with local performance. However, Allee effects or other density-dependent behaviours could lead to a wide range of non-linear relationships between population density, and hence local performance, and dispersal, e.g., u-shaped or threshold functions (Fronhofer et al. 2015; Poethke et al. 2016). To address these limitations, future studies could build upon our modelling framework to explore how trait evolution and a wider range of dispersal strategies are modulated by landscape features and competition dynamics.

4.6 Supplementary Information

4.6.1 Simulated landscapes: Extended description

We generated 25 landscape types following a cross-factorial design combining 5 levels of seasonality (top-right panel in Figure 4.1) and 5 levels of spatial autocorrelation (top-left panel in Figure 4.1) in environmental conditions. We chose these two landscape features because they have been shown to impose costs and risks to species movement that ultimately dictate ecological and evolutionary constraints on dispersal (McPeek and Holt 1992; Büchi and Vuilleumier 2012).

We started by assigning 50 patches into the landscape by drawing their x and y spatial coordinates from a uniform distribution ranging from 0 to 60. Environmental conditions at the patch level were defined by a single continuous variable (Env), spatially autocorrelated across patches and bound within the interval [0,5]. Spatial habitat autocorrelation in the environment was generated by random draws from a multivariate normal distribution (mu=0) with a covariance matrix set as the desired spatial autocorrelation level. Covariance matrices were created such that environmental similarity decayed exponentially with geographic distance according to φ. By manipulating the levels of φ, we generated landscapes where environmental conditions ranged from weakly autocorrelated (e.g., $\varphi = -0.9$) to strongly autocorrelated (e.g., $\varphi = -0.01$).

To simulate seasonal environmental variation, we set local environmental conditions to follow a sinusoid function with 100 periods, each composed of 12-time steps (i.e., 100 years) plus a random error $N(0,0.1)$ that served to mimic the effects of temporal environmental stochasticity. The amplitude of the sinusoidal environmental variation was modulated by a multiplicative factor s (constant across all patches). By manipulating the values of s, we created landscapes ranging from highly aseasonal ($s = 0.1$) to highly seasonal ($s = 1$).

We then seeded metacommunity simulations (see *Metacommunity dynamics* below) with the final matrices containing environmental values of each patch over time (**Env**, 50 patches x 1200 times), and their corresponding spatial coordinates.

4.6.2 Metacommunity dynamics: Extended description of within-patch mechanisms

The dynamics of ecological communities were spatially explicit, discrete in time, and governed by habitat selection, demographic stochasticity, competition at intra and interspecific levels, and dispersal (Shoemaker and Melbourne 2016; Thompson et al. 2020).

Considering that $N_{i,j,t}$ is the abundance of species i in site j at time t, population dynamics is governed by:

$$N_{i,j,t} = Poisson(N_{i,j,t-1} * P_{i,j,t}) - E_{i,j,t} + I_{i,j,t} \qquad \text{(SI 4.1)}$$

The first term of eq. 1 is a modified version of the Beverton-Holt competition model (Beverton and Holt 1957) that sets population growth as a function of habitat selection, competitive dynamics, and demographic stochasticity. $E_{i,j,t}$ and $I_{i,j,t}$ are the total number of individuals of i that emigrate from and immigrate to site j at time t, respectively.

$P_{i,j,t}$, is the local performance (i.e., growth rate) of species i when conditioned to competition and habitat selection and is modelled as follows:

$$P_{i,j,t} = R.max * exp\left(\frac{-(Env_{j,t}-\mu_i)^2}{2\sigma^2}\right) * \frac{1}{(1+\alpha_{intra}N_{i,j,t}+\alpha_{inter}\sum_{k\neq i}^{S}N_{k,j,t})} \qquad \text{(SI 4.2)}$$

$R.max$ represented the species' maximum intrinsic growth rate and was assumed to be equal to 3 for all species. Making $R.max$ equal across species ensured that they could reach the same maximum growth rate when optimum habitat conditions and competition at the intra and interspecific levels are negligible. $P_{i,j,t}$ ranges from 0 to $R.max$

The second term of eq. SI 4.2 sets *environmental suitability, i.e.,* the match between local environmental conditions (Env_{jt}) and species' environmental requirements. *environmental suitability* values ranges between 0 to 1, representing a complete mismatch and a complete match between species niche optima and local habitat conditions, respectively . We simulated species exhibiting distinct (non-neutral) performances along the same environmental gradient. To that end, species niche optima μ_i took values equally spaced along the interval [0,5] to scale with environmental variation. Note that niche tolerance (σ) was fixed across all species and narrow enough to make species respond to environmental variation. Preliminary simulations showed that fixing $\sigma = 1$ makes species sensitive to environmental variation without making them overly prone to local extinctions when conditions are suboptimal.

In parallel, to investigate how the assumption of niche differentiation can modulate (either buffer or potentialize) the influence of metacommunity dynamics on successful dispersal strategies, we also ran simulations with equal (neutral) species' responses to environmental variation. This was operationalized by assigning the same environmental optima to all species (μ_i = average value of **Env**).

The third term of eq. SI-4.2 models the effects of density-dependent competition on population dynamics. Stabilizing competition was set as α_{intra}= 0.0066 > α_{inter} = 0.0033; equalizing competition as $\alpha_{intra} = \alpha_{inter} = 0.005$; destabilizing competition as $\alpha_{intra} = 0.0033$ < $\alpha_{inter} = 0.0066$. These values imply the assumption that when locally dominant species make up 50% of the total community abundance, they face the same level of per-capita competition across the three types of competition. Prior tests (not shown) demonstrated that relaxing this assumption by considering different combinations of values for α_{intra} and α_{inter} did not influence the observed patterns of local coexistence and exclusion. However, they did alter the abundances at which locally coexisting species reached stable equilibria.

To incorporate the influence of demographic stochasticity on local birth and survival, we draw the final local species abundances from a Poisson distribution whose mean was determined by the estimated population size after habitat suitability and density-dependent competition (Shoemaker and Melbourne 2016; Shoemaker et al. 2020*b*).

Individuals able to persist in any given local community after within-patch selection and drift at time t could then disperse. See main text for a detailed description of how dispersal (departure, movement, and settlement) was computed.

4.6.3 Extended results and 3-way interactions
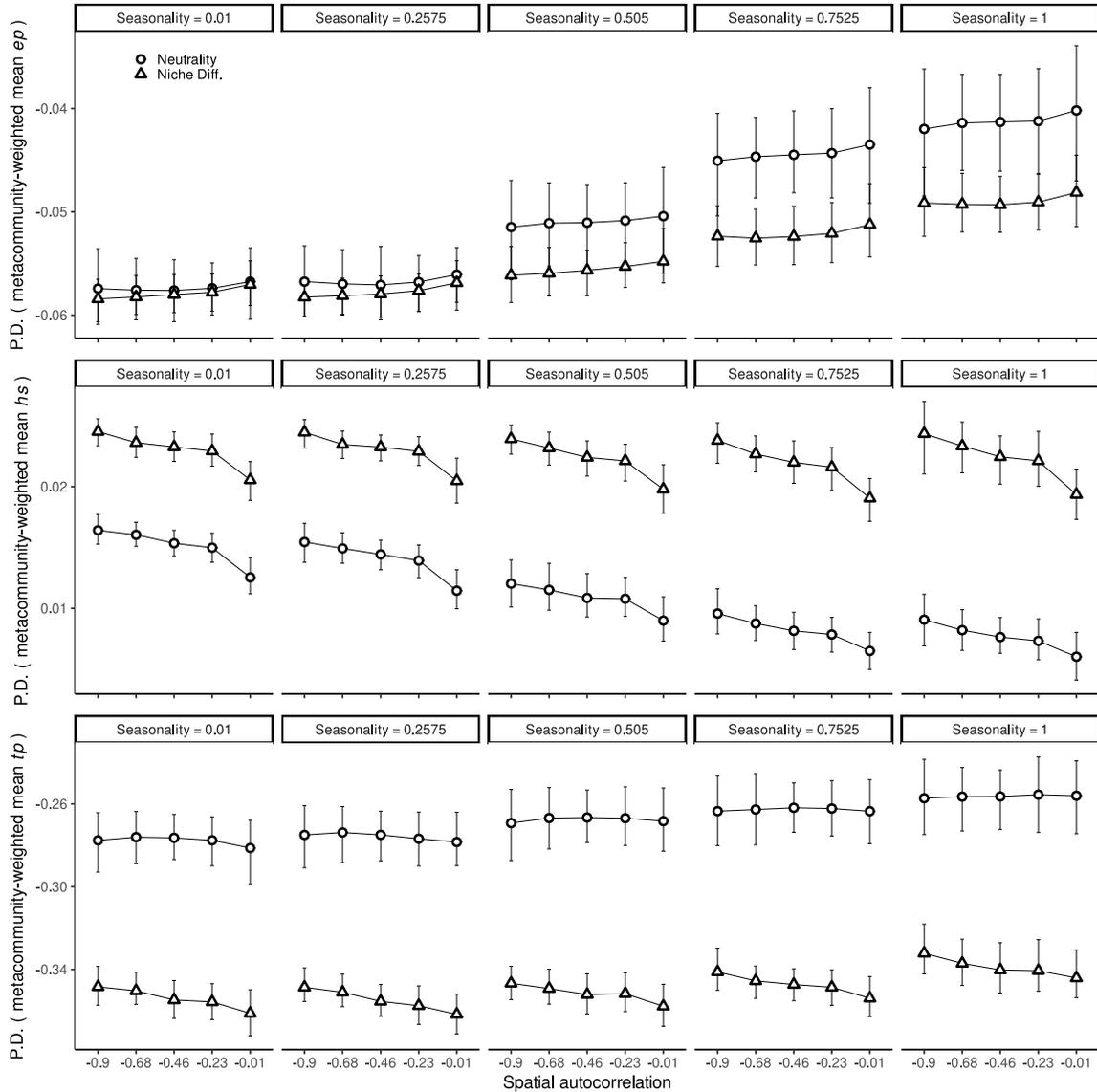
Figure SI. 4.I: Partial dependence (PD) plots showing interactive effects of seasonality (panels), spatial autocorrelation (x-axis), and the niche differentiation assumption (symbols) on the dominant dispersal strategies in metacommunities (i.e., metacommunity-weighted mean values for *ep, hs, ts).*
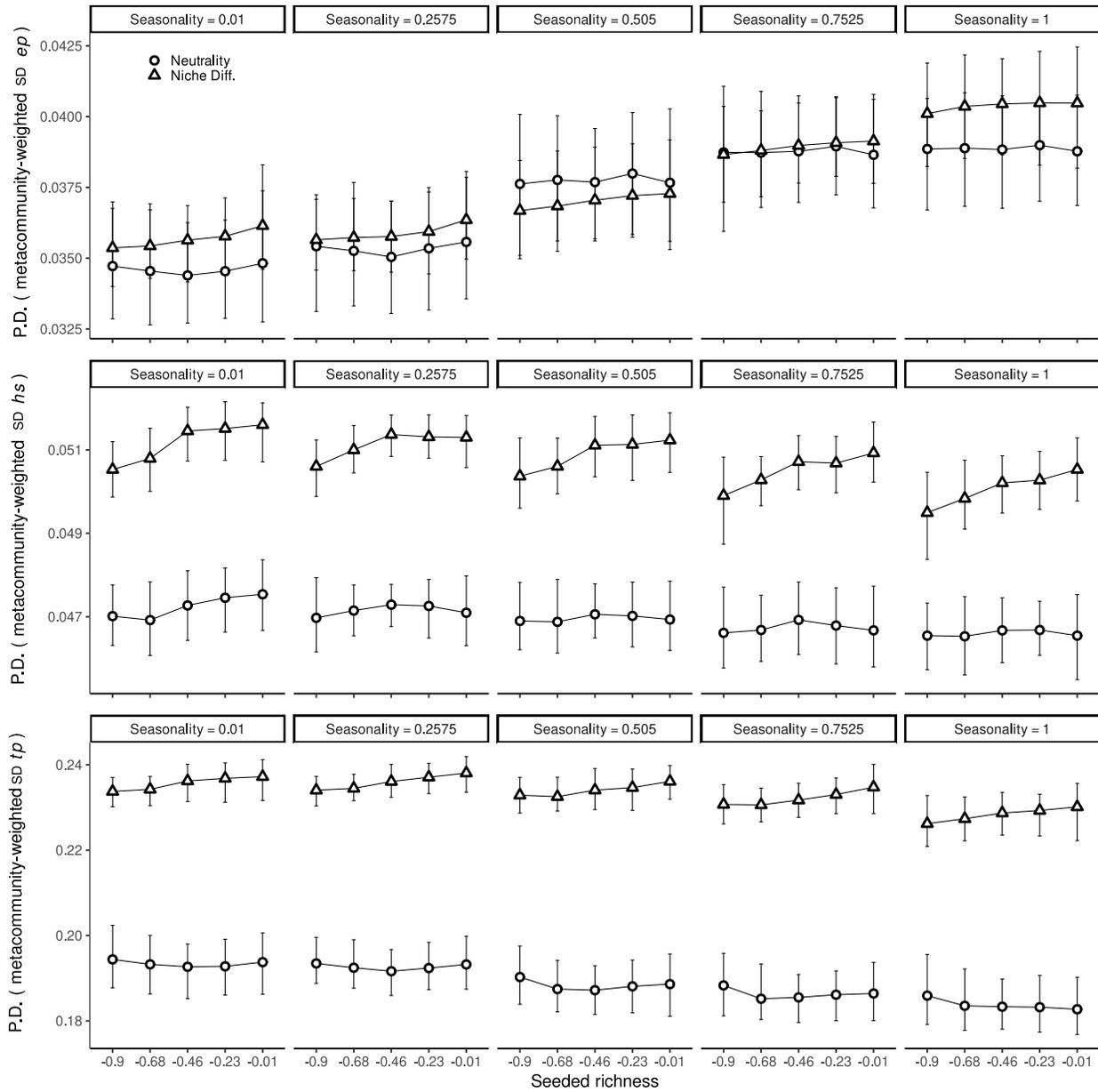
Figure SI. 4.II: Partial dependence (PD) plots showing interactive effects of seasonality (panels), spatial autocorrelation (x-axis), and the niche differentiation assumption (symbols) on the diversity of dispersal strategies in metacommunities (i.e., metacommunity-weighted standard deviation values for ep, hs, ts).
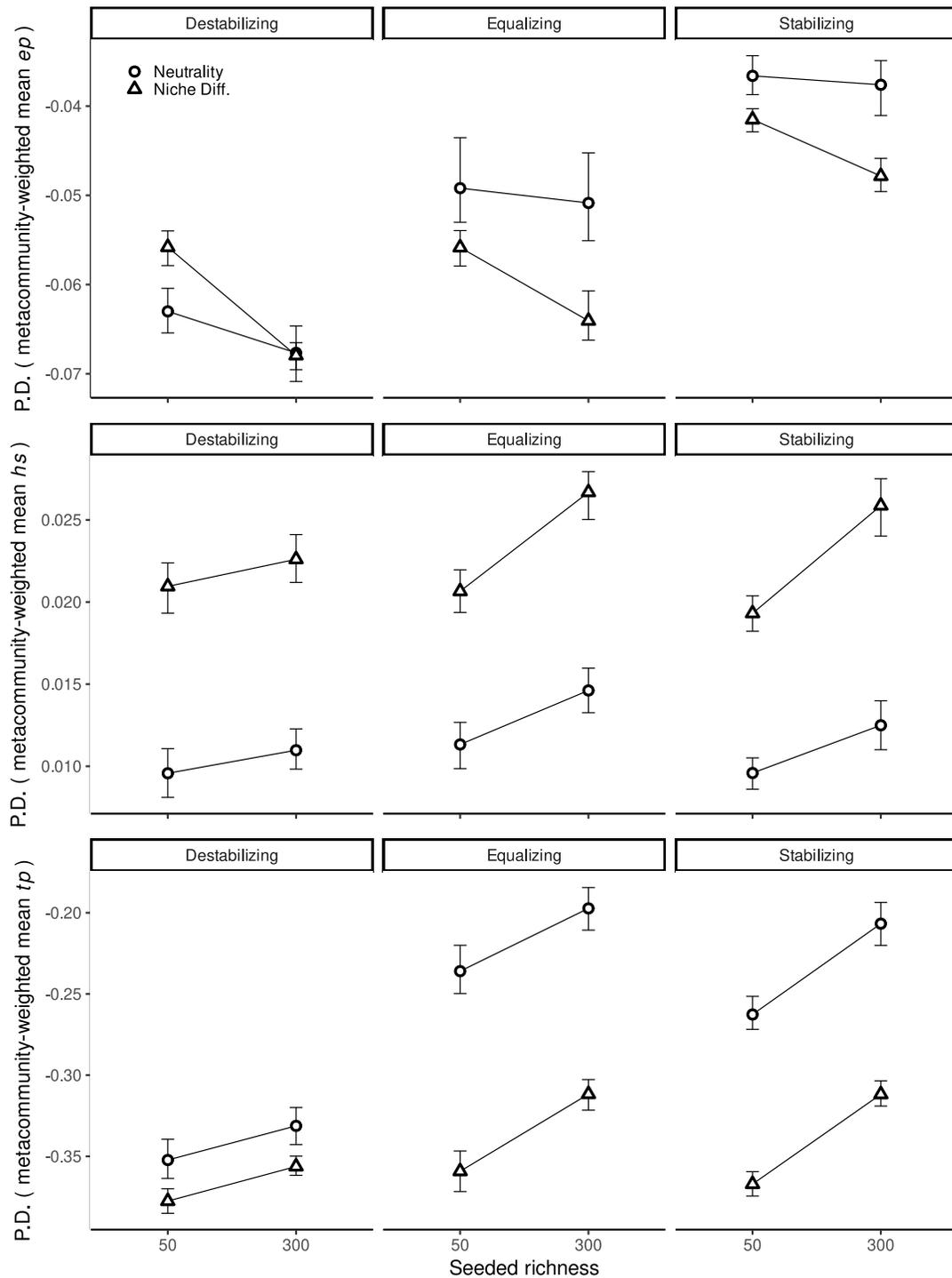
Figure SI. 4.III: Partial dependence (PD) plots showing interactive effects of competition type (panels), seeded richness of competitors (x-axis), and the niche differentiation assumption (symbols) on the dominant dispersal strategies in metacommunities (i.e., metacommunity-weighted mean values for ep, hs, ts).
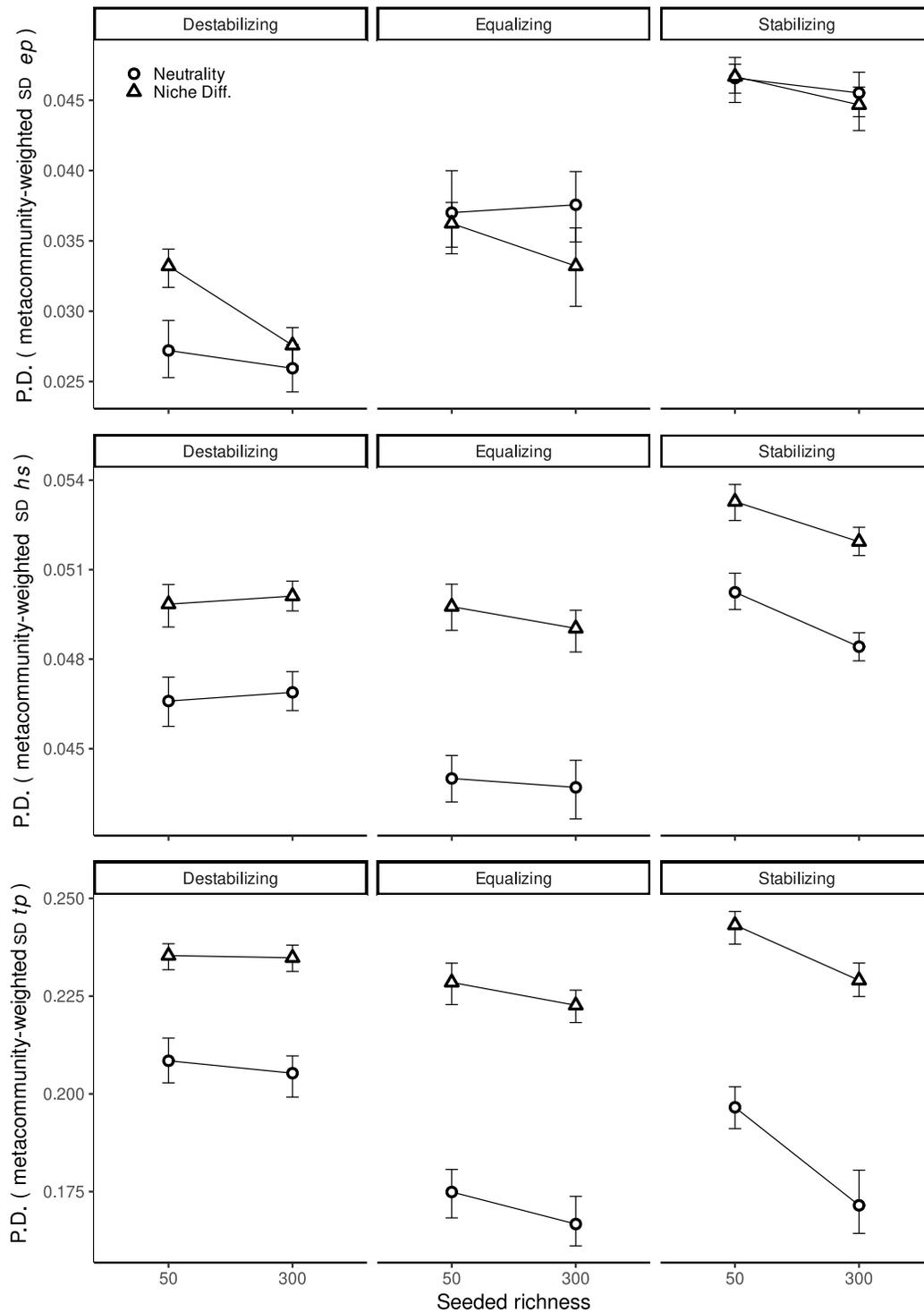
Figure SI. 4.IV: Partial dependence (PD) plots showing interactive effects of competition type (panels), seeded richness of competitors (x-axis), and the niche differentiation assumption (symbols) on diversity of dispersal strategies in metacommunities (i.e., metacommunity-weighted standard deviation values for ep, hs, ts).

Chapter 5 : Uncovering the trajectories of metacommunities: insights gained from the keystone community concept

5.1 Abstract

Understanding how metacommunities change over time due to natural and/or anthropogenic disturbances is a key goal in ecology. Here, we developed an analytical framework that is robust in identifying communities whose extirpation triggers stronger (i.e., keystone communities) or weaker (i.e., idle communities) cascading effects on extinction and colonization patterns that ultimately drive temporal changes in metacommunities (i.e., changes in compositional patterns among the remaining communities). Through mechanistic simulation models, we demonstrated that our framework correctly yields "keystoneness" estimates that rank local communities based on their significance in maintaining the structure of metacommunities. We also demonstrate how community keystoneness relates to habitat patch characteristics and how this relationship varies among different metacommunity archetypes. We applied the framework to examine a moth metacommunity situated in a protected mountainous region, which, due to its proximity to urban areas, is subject to both direct and indirect (artificial skyglow) effects of light pollution. As expected, we observed that light pollution was positively associated with local diversity but negatively correlated with the keystoneness moth communities. We conclude our study with an in-depth discussion on the application of our proposed analytical framework in assessing the conservation significance of local ecological communities.

5.2 Introduction

The internal structure of metacommunities (i.e., networks of communities linked through the dispersal of potentially interacting species) is an emergent property of the compositional (dis)similarities among their local communities (Leibold et al. 2021). Temporal variation in local community composition due to colonization and local extinction events renders the internal structure of metacommunities temporally dynamic. Losses and gains of species across communities in a large metacommunity are natural phenomena and result from the cumulative effects of dispersal, environmental variation, demographic stochasticity, and biotic interactions (Fukami 2015). Consequently, the internal structure of metacommunities is intrinsicality related to its spatial dynamics. For instance, a metacommunity can become more homogenous (i.e., low dissimilarity among its communities or low spatial beta-diversity) over time when there is an increase in species occupancy through colonization and/or an increase in extinctions of locally rare species (Olden & Poff 2003). Metacommunities can become more heterogenous over time when dominant species go extinct across multiple communities and/or new (alien) species colonize a few local communities, resulting in increased spatial beta-diversity (Tatsumi et al., 2021). A growing body of evidence reveals that the internal structure of metacommunities has been

changing at alarming rates owing to human-driven transformations in the abiotic and biotic milieu of landscapes (e.g., Pilotto et al., 2020). Consequently, a pressing task in theoretical and applied ecology is to understand the means by which we can assess, manage, and predict the magnitude and the trajectory (i.e., either towards homogenization or differentiation) of temporal changes in the internal spatial structure of metacommunities after natural and anthropogenic disturbance events (Chase et al. 2020).

To effectively manage the temporal trajectories of metacommunities, we must first recognize that their local communities play unequal roles in shaping their internal structure. This principle has been formalized by the "Keystone community concept" (hereafter KCC), which postulates that communities can be sorted along a continuum based on their significance in maintaining essential properties of a metacommunity (Mouquet et al. 2013). At one end of this continuum lie keystone communities. Their extirpation triggers pronounced temporal shifts in local extinction and colonization patterns that ultimately drive changes in the remaining communities' compositional (dis)similarity (i.e., structure) over time. At the opposite end, there are idle communities. Their absence neither disrupts the functioning of metacommunities nor alters the compositional (dis)similarity of the remaining communities over time.

Ranking local communities along the continuum between keystone and idle - termed "keystoneness" - holds significance for conservation and management efforts as it can provide deep insights into the fundamental processes that shape ecological communities and govern their interactions. This knowledge can be invaluable while prioritizing resources and efforts for successful conservation and management. For instance, in the context of conservation biology, information about community keystoneness can assist managers in reducing the effects of anthropogenic disturbances on metacommunities (Economo 2011, O'Sullivan et al. 2023). In invasion ecology, assessing community keystoneness can guide decisions on which communities should be prioritized for interventions to curb the spread of invasive species across the metacommunity (Brown and Barney 2021). In the context of sustainable natural resource management, assessments of keystoneness can assist managers in determining which communities are (or are not) suitable for harvesting while preserving the internal dynamics of metacommunities.

Furthermore, the knowledge about a community's keystoneness becomes particularly relevant when local species diversity does not correlate with habitat quality. This is the case when anthropogenic stressors lead to an increase in local diversity by driving a substantial influx of individuals into unsuitable habitats where they are unable to sustain viable populations over time (see theoretical example in Delibes et al. 2001, and empirical example below). In such cases, local diversity should not be considered a proxy for the conservation value of local communities. Instead, effective biodiversity protection depends on information about the role of local communities in regulating and maintaining metacommunity dynamics.

Here, we leveraged the well-established mathematical properties of graphs to build an analytical framework that quantifies a local community's importance in maintaining

metacommunity structure over time (i.e., hereafter, their keystoneness degree). Under specific conditions and assumptions (discussed in detail below), our framework can be implemented on observational data on species' spatiotemporal distributions to yield keystoneness estimates that are statistically independent of local diversity. This independence is relevant because it ensures the framework's utility in guiding management and conservation when local diversity does not accurately correlate with habitat quality, e.g., when anthropogenic stressors attract species to demographic sinks with higher death rates than birth rates (i.e., source-sink dynamics, Delibes et al. 2001).

We used mechanistic simulation models across various scenarios to provide theoretical validation that the proposed framework is robust in identifying communities whose extirpation causes the most (keystone) and the least (idle) temporal shifts their metacommunities. In essence, the framework assigns higher keystoneness values to communities that, when removed, triggers cascading effects in extinction and colonization dynamics, resulting in significant changes in the internal structure of metacommunities over time. Our framework's validation also demonstrated how our metric of community keystoneness relates to habitat patch characteristics and how this relationship varies among different metacommunity archetypes (i.e., species-sorting, mass-effects, neutral metacommunities). As such, our simulation framework shows how information about community keystoneness can help us to better understand the trajectories of different types of metacommunities after local disturbance events.

To showcase the unique insights derived from our analytical framework on the effects of anthropogenic disturbances on metacommunities, we applied it to investigate the effects of light pollution on the structure of a well-studied moth metacommunity (more information about the moth data in Choi & Na, 2020, Figures SI 5.I-II). Situated within a protected mountainous landscape surrounded by urban settlements, this metacommunity has been under direct and indirect (in the form of artificial skyglow) influences of light pollution (Figure SI 5.I). Light pollution poses a widely recognized threat to moth biodiversity, as even at incipient levels, it can increase local mortality by disrupting various aspects of their life history, including their ability to use the night sky to navigate the landscape (reviewed in Boyes et al., 2021). Given that most moths are positively phototactic (i.e., attracted to light), we predicted that light pollution would increase local diversity by attracting more individuals into artificially lit areas (Langevelde et al. 2011). This implies that anthropogenic disturbances affecting dispersal can render local diversity to be an unreliable indicator of habitat suitability. Additionally, we predicted that light pollution would reduce community keystoneness, suggesting that communities less impacted by light pollution play a crucial role in maintaining the internal structure of the metacommunity.

5.3 Methods

5.3.1 An analytical framework for community keystoneness

By definition, keystone communities are those whose removal/extirpation has a disproportionate impact on a given metacommunity property (here, its internal structure) (Mouquet

et al 2013). Consequently, analytical frameworks aiming to assess community keystoneness accurately must satisfy the following two criteria. Firstly (hereafter, C1), it needs to measure the impact of a single community's removal or extirpation on the metacommunity's internal structure. Secondly (hereafter, C2), it needs to determine whether this observed impact is disproportional (atypical) to what is expected given a specific community property (e.g., local diversity, biomass).

The framework described here capitalizes on the well-established mathematical properties of graphs (Figure 5.1) to meet C1. It draws inspiration from methods designed to evaluate the resilience of computer networks against failures or targeted attacks on specific computers (Liu et al. 2009) while incorporating additional steps to tailor its application to the context of the keystone community concept. While we do not provide a comprehensive introduction to graph theory (see Dale 2017 for an ecological perspective), Supp. Information offers a concise overview of key concepts, helping those unfamiliar with the subject to grasp our framework.

Metacommunities can be represented as graphs, where each node corresponds to a local community, and the weight of edges connecting pairs of communities reflects their compositional similarity. The underlying rationale is that a high degree of compositional similarity indicates the occurrence of significant past and current dispersal events connecting a pair of local communities within an ecological (or possibly evolutionary) timeframe (see more in Layeghifard et al., 2015).

The pairwise-compositional similarity matrix is used to calculate a weighted Laplacian matrix ($L^{weighted}$) following Liu et al. (2009):

$$L^{weighted} = \begin{cases} W_{(v_i)} & if \ i = j \ and \ degree \ (v_i) \neq 0 \\ -W_{(i,j)} & if \ i \ \neq j \ and \ v_i \ is \ adjacent \ to \ v_j \\ 0 & otherwise \end{cases}$$

where v_i and v_j are the ith and jth nodes (here, local communities). $W_{(v_i)}$ is the degree of the ith node and is given by the sum of the weights of edges (here, similarities) of all nodes connected to v_i. $W_{(i,j)}$ is the weight of the edge (similarity) between nodes i and j.

The second smallest eigenvalue of the resultant weighted Laplacian estimates how difficult it is to disconnect a graph (i.e., its algebraic connectivity). To estimate the importance of each community to the overall structure of the metacommunity (i.e., graph), one can remove local communities, one at a time, and recalculate the second smallest eigenvalue of the resulting sub-graphs (i.e., weighted Laplacian matrix recalculated without the focal community). By subtracting the second smallest eigenvalue based on the full metacommunity from the second smallest eigenvalue calculated when the focal community is removed, one estimates the impact of the removal of the focal local community on the overall structure (connectivity) of the metacommunity structure.

It is important to note that the estimated impacts of a community removal should not be readily interpreted as their level of keystoneness. This is because the local abundance and diversity of communities significantly influence their impact estimates. For instance, removing locally

diverse communities is more likely to have a higher impact on metacommunity structure because, by chance, they are more likely to share species with other communities. To accurately identify communities whose removal exerts a disproportionate (atypical) impact on the structure of a metacommunity, thus meeting C2, one must contrast a local community's observed impact against their expected impact considering its local species richness and abundance. This contrast highlights communities whose removal from the metacommunity has greater (i.e., keystone communities) or
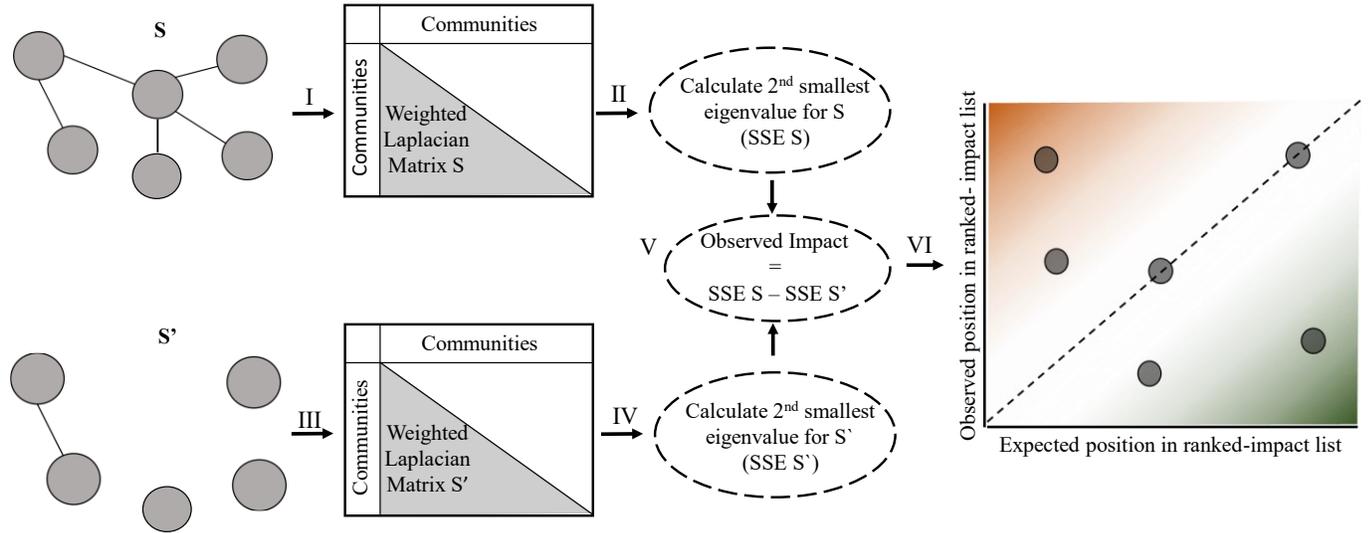
Figure 5.1: Schematic representation of the proposed analytical framework to estimate community keystoneness. We start by creating a graphical representation of a metacommunity (S) wherein nodes are communities and edges are weighted compositional similarities (Step I). We then estimate the second smallest eigenvalue (SSE) of the weighted Laplacian matrix representing **S** using eq. I (see Main text, and Supp. Material I for details) (Step II). The SSE informs how difficult it is to disconnect a graph (i.e., its algebraic connectivity). Then we remove a community of **S** to create **S'**, recalculate the SSE based on **S'** (Steps III and IV), and estimate the observed impact of community removal as the difference between SSE S and SSE S' (Step V). We then rank communities in ascending order of impact and use null models to assess their expected position in the ranked-impact list considering their size (abundance) and the abundance distribution of species in the regional pool. By contrasting the observed and expected position of communities in the ranked-impact list (see *Keystoneness* in main text), we can identify communities whose removal disproportionately impacts **S** (Step VI). "Keystone communities" are those whose observed position in the ranked-impact list is higher than expected based on their size (orange area in the scatterplot plot). "Idle communities" are those whose observed position in the ranked-impact list is lower than expected based on their size (green area in scatterplot plot).

lesser (i.e., idle communities) impacts than expected considering their size alone.

To assess the expected impact of the removal of a local community, we first rank local communities in ascending order of observed impact. This transformation is necessary because previous analyses (not shown) demonstrated that it enhances the statistical tractability of expected impact estimates. Specifically, it corrects for the skewness of the null distribution of expected impact values generated by the null models described below. Note that this rank transformation preserves the original interpretation of observed impacts, i.e., communities whose removal has the highest impact on metacommunity structure rank higher on the impact list.

Following, we used null models that kept community sizes (local abundance) and species regional abundances equal (fixed) to their original values (following Kraft et al. 2011). These null models generate communities whose position in the ranked-impact list is determined solely by their observed total abundances and species abundance distribution at the metacommunity scale. After repeating this null model procedure multiple times (here 1000 times), we estimated each local community's expected (average) position in the ranked-impact list. Then the keystoneness of community i is given by:

$$Keystoneness_i = \frac{(Obs_i - Expected_i)}{SD_i}$$

where Obs_i is the observed position of community i in the ranked impact list as calculated earlier, $Expected_i$ is its average position across all iterations of the null model, and SD_i is the standard deviation of the simulated positions. This keystoneness metric quantifies the deviation, in standard deviation units, of a community's position in the ranked-impact list from what would be expected based on their size alone. If $Keystoneness_i$ is greater than 0, it means that the position of community i in the ranked-impact list is higher than expected by its size. As such, community i can be considered a keystone community. If $Keystoneness_i \approx 0$, its position is equal to the expected by its size. Finally, if $Keystoneness_i$ is smaller than 0, its position in the ranked-impact list is lower than expected by its size. As such, community i can be considered an idle community.

In Supp. Information (Figure SI 5.III and 5.IV), we demonstrated that contrary to other indexes previously used as proxies of community keystoneness (e.g., the local contribution to beta-diversity- LCBD- developed by Legendre and De Cáceres 2013, and used in Ruhí et al. 2017), the framework described above can yield estimates that are either statistically independent or weakly correlated with community diversity (here, Hill-numbers of different orders).

5.3.2 Theoretical validation

We used process-based metacommunity simulations to: (i) provide theoretical validation that the proposed framework accurately ranks communities based on their importance in maintaining the internal structure of metacommunities; (ii) understand how the proposed estimate of community keystoneness relates to patch connectivity and habitat quality; (iii) and how these relationships vary in magnitude among metacommunities varying in their species ecological specialization and dispersal abilities. Due to space limitations, a condensed summary of the simulations is provided below, while a detailed description can be found in Supp. Information.

We started by randomly distributing 30 habitat patches in a geographic space defined by xy coordinates ranging from 0 to 60. Environmental conditions were spatially structured and fluctuated over time to emulate environmental stochasticity. In our simulations, individuals could only move between physically connected patches. The physical connectivity between any given pair of patches decayed exponentially with distance and was truncated to zero if below a fixed threshold (here 10^{-3}). Thus, each randomly generated landscape contained patches with varied

environmental conditions and differed in the number of connections to neighboring patches (i.e., patch degree).

Population dynamics of 100 species across the 30 patches were discrete in time and generated by a Beverton-Holt growth model with generalized Lotka-Volterra competition. Community assembly at time t was the outcome of three sequential steps: 1) within-patch dynamics (i.e., environmental selection, demographic stochasticity, and competition at both intra and interspecific levels); 2) emigration; and 3) immigration (see how each one of these steps was calculated in Supp. Information). Dispersal was spatially explicit, meaning that individuals could only immigrate to patches connected to their natal patch.

Simulations were carried out in two phases. In the first phase, we allowed metacommunities to reach stability (here, rates of regional extinction close to zero) by letting population dynamics in the landscape run for 1000 time steps (from t_1 to t_{1000}). Estimating community keystoneness in stable rather than transient metacommunities is relevant because it ensures that patterns in community keystoneness would remain consistent over time, provided that no external factors moved the entire metacommunity from stable equilibrium (see Discussion below). We halted population dynamics at t_{1000} and applied our framework to estimate community keystoneness based on local community-by-species matrices, where the entries were obtained by rounding the average abundance of each species at each local community between time intervals from t_{950} to t_{1000}.

We then used a random forest to model community keystoneness as a function of habitat quality (i.e., average patch suitability across all species in the metacommunity, see equation SI 5.2 in Supp Material) and physical connectivity (i.e., number of patches a focal patch is connected, hereafter, patch degree of a local community). Random forest modeling is suitable for these analyses because: (i) it does not assume any shape of the relationship between response and predictors; (ii) it is robust to collinearity and automatically models the influence of complex multi-level interactions among predictors on the response, and; (iii) they provide robust estimates of the importance of predictors in explaining changes in the response (Breiman 2001). Effect size estimates for each predictor (based on Cafri & Bailey (2016)) allowed us to understand the strength and direction of the relationships between predictors and response after accounting for (i.e., averaging out) the influence of other predictors. Since all predictors were standardized (mean = 0, sd =1) before model fitting, estimated effect sizes are directly comparable. We expected a positive correlation between community keystoneness and habitat quality and/or patch degree (connectivity).

In the second phase, we conducted removal experiments to demonstrate that, when used on stable metacommunities, the proposed framework can identify communities whose extirpation triggers strong cascading effects on colonization and extinction dynamics that ultimately alter the structure of the metacommunity over time. These removal experiments were carried out as follows:(1) we permanently removed a focal community i from the metacommunity by eradicating

its patch and connections to other patches; (2) we estimated the overall spatial beta-diversity among the remaining communities in the metacommunity (hereafter, Spatial $\beta_{t1001-i}$); (3) we resumed simulated population dynamics across the remaining patches for an additional 300 time-steps; (4) we estimated the spatial beta-diversity of the remaining communities in the metacommunity considering the average abundance of species per patch from t_{1285} to t_{1300} (Spatial $\beta_{final-i}$); (5) following Tatsumi et al. (2021), we measured how the removal of community i influenced the trajectory of the remaining communities in the metacommunity as Spatial $\beta_{final-i}$ - Spatial $\beta_{t1001-i}$ (Δ Spatial β). When Δ Spatial β > 0, the removal of patch i caused an increase in metacommunity differentiation over time; Δ Spatial $\beta \approx 0$, the removal of patch i caused no changes in metacommunity structure; Δ Spatial β < 0, the removal of patch i caused an increase in metacommunity homogenization over time. Spatial $\beta_{t1001-i}$ and Spatial $\beta_{final-i}$ were estimated using the abundance-based and incidence-based Whittaker multiplicative indexes. By doing so, we were able to estimate how much of Δ Spatial β was driven by either species extinctions or colonisations following the extirpation of the focal community (Tatsumi et al. 2021, 2022).

We repeated steps 1 to 5 across all local communities in the metacommunity. By plotting the values Δ Spatial β obtained with the removal of each community against their level of keystoneness, we could assess whether local communities assigned high or low values of keystoneness corresponded to those whose removal caused major or minimal temporal changes in the internal structure of metacommunities, respectively.

We employed this simulation setup across combinations of three levels of species` niche breadth (narrow, intermediate, broad, see Supp. Information for details on the numeric specification of each level), and dispersal rates (low, intermediate, high, see Supp. Information for details on the numeric specification of each level), yielding a total of 9 simulation scenarios replicated 50 times (i.e., 450 simulation iterations in total). Considering these different combinations of species traits allowed us to test the proposed framework across the parameter space that encompasses various well-known metacommunity "archetypes", namely, species-sorting, mass-effect, and neutral metacommunities (Leibold et al. 2002; see also Thompson 2020).

We also conducted further analyses to determine the robustness of our framework. This was done by comparing the outcomes of removal experiments where single local communities were removed against those where groups of local communities, characterized by keystoneness deciles were extirpated (i.e., from Q1 to Q10 where Q1 = the three communities with the lowest levels of keystoneness were removed at once; Q10 = the three communities with the highest levels of keystoneness were removed at once). Although the results of the removal experiment were qualitatively similar between the two approaches, the effects of the removal experiment on metacommunity structure were stronger when groups of communities rather than single communities were removed. We reported the results of the "decile-removal" approach in the main manuscript and the "single-community-removal" in Supp. Information. We then discuss the biological meaning of both approaches (see Discussion).

5.3.3 Empirical Assessment: The influence of light pollution on the diversity and keystoneness of moth communities

We used published data on the spatiotemporal distributions of moths in Mount Hallasan National Park, located on the volcanic island of Jeju-do, South Korea (Choi and Na 2020). In total, 13,249 individuals across 587 species of macro-moths were sampled using light traps along an elevational gradient (11 sampled elevations) during the growing and reproductive season (May to October) over six consecutive years (2013-2018). We considered all moths surveyed at a given elevation during the growing season as the operational definition of local community in further analyses. By incorporating this temporal dimension, we aligned the KCC with recent advancements in metacommunity ecology, which demonstrate the importance of accounting for temporal variation in compositional data to infer the mechanisms influencing empirical metacommunities more accurately (Khattar et al. 2021; Record et al. 2021; Guzman et al. 2022). Moreover, by integrating data on the natural inter-annual variation in community composition into our estimates of community keystoneness, we could evaluate the temporal consistency of community keystoneness patterns in a real metacommunity. This assessment is relevant because significant yearly fluctuations in community keystoneness would suggest a non-equilibrium (transience) state for this moth metacommunity. In such a scenario, relying on highly variable keystoneness estimates for management and conservation strategies would be ill-advised (see more in discussion). Therefore, our analyses considered 66 operational communities (i.e., species sampled in 11 sample locations × 6 years).

A critical assumption of the proposed analytical framework is that all local communities that belong to a given metacommunity have been equally and sufficiently sampled. Before analyses, we relied on the method proposed by Anderson and Santana-Garcon (2015) to evaluate whether the number of operational local communities surveyed was adequate to accurately characterize the extent of spatiotemporal compositional (dis)similarities among communities. The negative relationship between mean error in estimates of community composition and the number of communities surveyed reached an asymptote around 42 operational local communities (see Figure SI 5.II). The results indicated that the survey's sampling effort was adequate to capture the spatiotemporal variation in community composition within the moth metacommunity.

We used the Jaccard-Chao similarity index (i.e., 1- Jaccard-Chao dissimilarity) to construct a metacommunity graph. Simulations (see more in Supp. Information, Figure SI 5.III) indicated that this index yields robust estimates of community keystoneness that are statistically independent of community local diversity. We calculated the local diversity of each community using the Hill-Shannon diversity index. Hill-Shannon diversity expresses the effective number of species in the community, i.e., the number of equally abundant species that should be observed in the community so that it has the same Shannon's index as the one calculated (Roswell et al. 2021).

We used the geographic coordinates of each site (local community) to obtain estimates of their average monthly climatic variables and light pollution. The climatic variables considered were monthly temperature (mean, max, minimum monthly values) and log-transformed

precipitation data at 1 km × 1 km resolution (CHELSA dataset, Karger et al. 2017). We performed a principal component analysis on climate variables (all standardized to mean = 0, sd = 1) and used the multivariate scores (i.e., at the site level) on the first two PC axes as proxies of the climatic conditions through time. The proportion of total variance accounted for by both PC axes was 93.7%.

Data on artificial light at night (ALAN) was extracted from satellite imagery from the United States National Oceanic and Atmospheric Administration Earth Observation group (https://ngdc.noaa.gov/eog/download.html) at a resolution of 0.5 km^2 and harmonized by Li et al. (2020). These data represent the amount of light emitted or reflected by a surface (measured in radiance, units nW/cm2 × sr) detected using the visible infrared imaging radiometer day-night band (VIIRS DNB) satellite. We estimated the average radiance of all pixels included in buffers of 1 km and 5 km around the sample locations to assess the direct (local) and indirect (regional, via sky glow) effects of light pollution in moth metacommunities, respectively. Note that the VIIRS DNB satellite system measures light pollution in a 500 to 900 nm spectral range. This spectral range overlaps with the spectral sensitivity of the photoreceptors of many groups of moths (Van Der Kooi et al. 2021), representing an appropriate estimate of the large-scale effects of light pollution on the sensory environment of this taxonomic group.

We used site coordinates to generate Moran's eigenvector maps (MEMs; Dray et al. 2006) and selected the eigenvectors that captured positive spatial autocorrelation to represent the effects of spatial processes on metacommunity dynamics. We generated asymmetric eigenvector maps (AEMs; Blanchet et al. 2008) based on the year the community was sampled to account for temporal autocorrelation.

As detailed earlier, we used random forest modeling to investigate how community keystoneness and local diversity (response variables) are influenced by light pollution, climate, and the spatial and temporal variables. We used the Boruta algorithm described in Kursa & Rudnicki (2010) to reduce dimensionality and ensure that only relevant predictors (features) were considered in the final model. This algorithm selects predictors whose importance to model prediction (measured as % of increase in mean-squared error) is higher than a threshold established by data permutation. Additionally, we estimated the effect size of the selected predictors (see more in Cafri & Bailey 2016). We used bootstrapping to estimate the variation in estimates of MSE and effect sizes across all variables (following Ishwaran & Lu, 2019).

All simulations and statistical analyses were conducted using R (v.4.2.0) (R Core Team 2023). We used the *vegan* (Oksanen et al. 2020) package to calculate compositional similarities. AEMs and MEMs were calculated using the *adespatial* package (Dray et al. 2022). Random forest models were run using the *randomForest* package (Liaw and Wiener 2002), while the Boruta algorithm and estimates of effect sizes were run using the *Boruta* (Kursa and Rudnicki 2010) and *rfUtilities* (Evans and Murphy 2018) packages, respectively.

5.4 Results

5.4.1 Theoretical validation

In all simulation scenarios, keystoneness increased with patch connectivity (i.e., effect sizes > 0), and these effects were more substantial with increased dispersal rates (Figure 5.2). Keystoneness also increased with habitat quality, but only when species were sensitive to environmental variation (narrow and intermediate levels of niche breadth) and were weaker dispersers. Outside this parameter range, keystoneness was not affected by habitat quality (i.e., effect sizes not significantly different from 0).
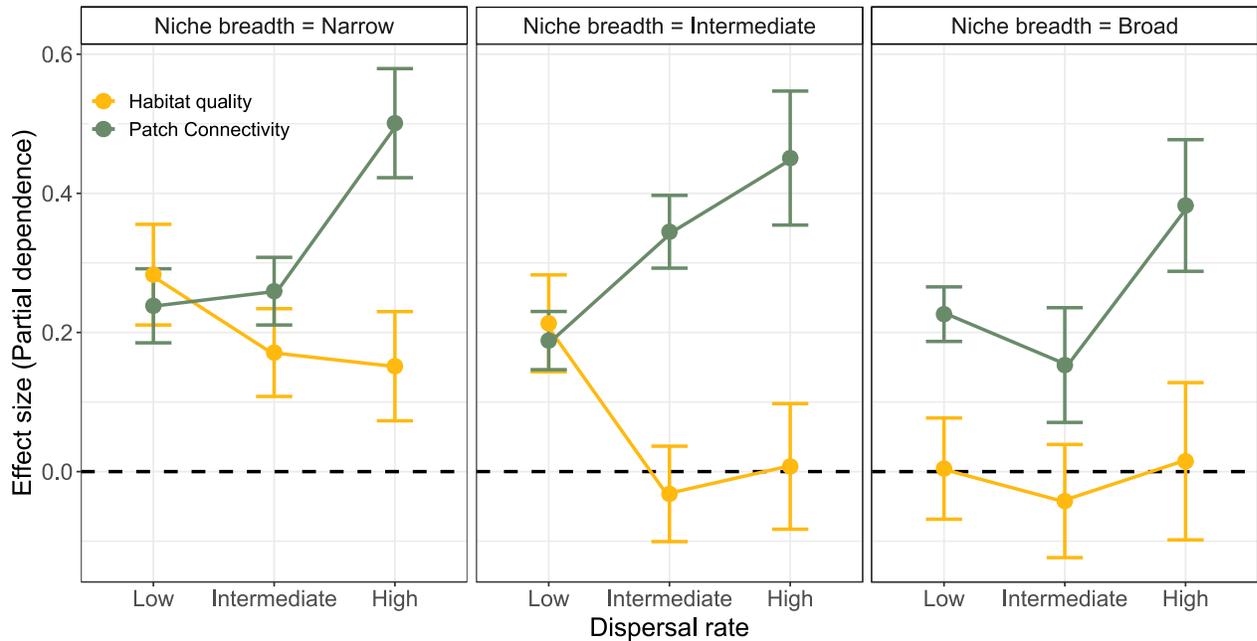
Figure 5.2: Results of random forest modeling considering community keystones as a function of habitat quality and patch connectivity. Community keystoneness increased with habitat quality and patch connectivity (effect sizes > 0). Note that the magnitude of these relationships varied as a function of species' niche breadths and dispersal abilities. Points represent average effect sizes across 50 simulation replicates per combination of levels of niche breadth and dispersal rate in the simulated regional pool. Whiskers represent 95% confidence intervals. Dashed line represents effect size = 0.

Across most combinations of species niche breadths and dispersal rates, we observed that our estimates of community keystoneness correctly sorted communities based on their influence on the internal structure of metacommunity (Figure 5.3, Figures SI 5.V and SI 5.VI). That is, when communities with low levels of keystoneness (i.e., Q1) were removed, little to no changes in metacommunity were observed over time (Δ Spatial $\beta \approx 0$). In contrast, when communities with high levels of keystoneness were removed (Q10), there was an increase in metacommunity differentiation over time (Δ Spatial $\beta > 0$). When Δ Spatial β was estimated based on species abundances (Figure 5.3), we observed that such an increase in metacommunity differentiation was driven by reductions in the abundance of common species across communities combined with an increase in the abundance of locally rare species. When Δ Spatial β was estimated based on

incidence data (Figure SI 5.V), we observed an increase in metacommunity differentiation led by the extinctions of species that were dominant in the metacommunity prior to the extirpation of local communities. These results were qualitatively similar, although weaker when the removal experiments were based on the extirpation of individual rather than multiple (groups) communities (Figure SI 5.VI).
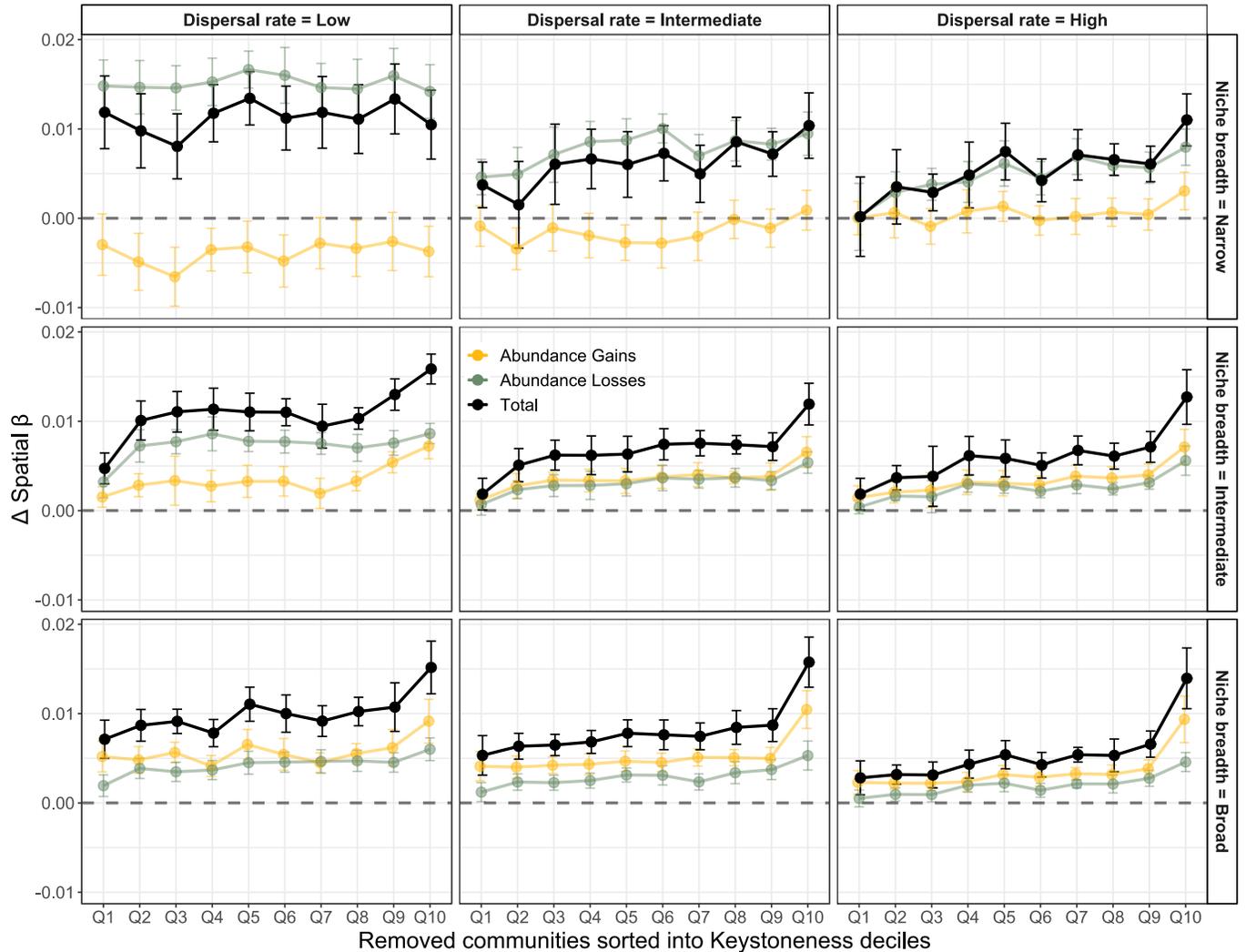
Figure 5.3: Results of simulated removal experiments. Impact of community removal on temporal changes in the metacommuity structure (total Δ Spatial β) caused by losses and gains in species local abundances. Communities were sorted into deciles of keystoneness (i.e., Q1 = the 3 communities with the lowest levels of keystoneness, Q10 = the 3 communities with the highest keystoneness). Losses and gains on local species abundances over time sum up to total Δ Spatial β. Points represent average effect sizes across 50 simulation replicates per combination of levels of niche breadths and dispersal rates in the simulated regional pool. Whiskers represent 95% confidence intervals. Dashed line represent Δ Spatial β = 0.

5.4.2 Empirical Assessments of community keystoneness

Mapping the degree of community keystoneness across space and time in the focal moth metacommunity revealed that high-elevation communities consistently exhibited a greater degree

of keystoneness compared to low-elevation communities (Figure 5.4). While climate and spatial variables primarily drove keystoneness and local diversity (Figure 5.5), it is noteworthy that light pollution (at both scales) also played a role, as indicated by its inclusion in the final random forest models for local diversity and keystoneness. More specifically, after controlling for the influence of climate, spatial MEMs, and temporal variables AEMs, light pollution was positively associated with local diversity (effect size > 0) but negatively associated with community keystoneness (effect size < 0). Together, these results suggest that anthropogenic stressors can significantly impact the diversity and keystoneness of this moth metacommunity.
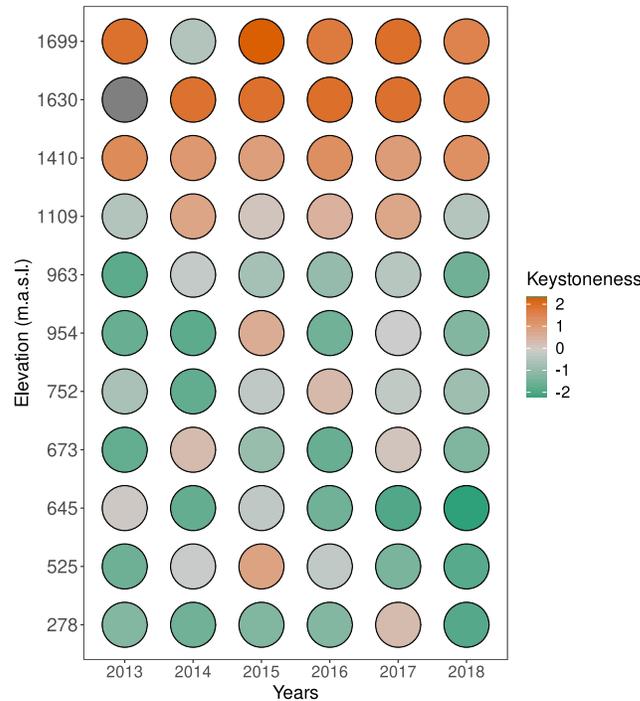
Figure 5.4: Mapping the keystoneness of moths communities in space (across elevations) and time. High-elevation communities consistently exhibited a greater degree of keystoneness compared to low-elevation communities.

5.5 Discussion

In this study, we proposed a novel analytical framework that identifies keystone communities, i.e., those whose extirpation promotes strong cascading effects on extinction and colonization patterns that ultimately drive temporal changes in compositional patterns among the remaining local communities (i.e., temporal changes in the internal structure of metacommunities). Through simulation models that replicated in silico metacommunity dynamics and removal experiments, we provide theoretical evidence that understanding the keystoneness of ecological communities aids in predicting metacommunities' trajectories—either towards homogenization or differentiation—when confronted with local disturbances. This simulation framework also helped us understand how our proposed community keystoneness metric relates to habitat patches' characteristics and how this relationship varies among different metacommunity archetypes. We

119

then used this framework to investigate the influence of light pollution on the diversity and keystoneness of moth communities. This demonstrated how the proposed framework deepens our understanding of the effects of anthropogenic stressors on the internal structure of metacommunities.

5.5.1 Keystoneness and the trajectories of metacommunities: insights gained from simulation models

Our analyses revealed a positive correlation between community keystones and both habitat quality and connectivity of the patch inhabited by the local community. Notably, the magnitude (strength) of these relationships varied across levels of species ecological specialization and dispersal rates (Figure 5.2). In the extreme case of species-sorting metacommunities, characterized by narrow niche breadths and low dispersal rates, habitat quality was the most important driver of community keystoneness. However, by increasing species dispersal rates and/or the degree of ecological specialization (i.e., moving towards mass-effect or neutral

Figure 5.5 Results of random forest modeling considering local diversity (Panel A, $R^2 = 0.58$) and keystoneness (Panel B, $R^2 = 0.61$) as a function of climate, spatial MEMs, temporal AEMs, and light pollution (ALAN, in yellow) at different scales, spatial and temporal variables of the empirically studied moth metacommunity. Only predictors retained via model selection are reported here. The first column shows variables ranked in order of importance to model fit (estimated as % of the increase in mean squared errors when permuted). The second column shows their standardized effect sizes. Dashed line represents effect size = 0, and whiskers indicate a confidence interval at 95%. The third column reports partial dependence plots. They serve as a graphical depiction of how the average predicted values of the response (local diversity and keystoneness) change with variation in the level of ALAN at smaller (1km buffer) and larger (5km buffer) scales. The shaded area encompasses +- 1 Standard error of the mean.

metacommunities), patch connectivity became the most relevant driver of community keystoneness.

The observed interplay between species niche breadth and dispersal ability, in relation to patch attributes and community keystoneness, suggest biogeographic variations on the potential success biodiversity conservation strategies centered on landscape connectivity or habitat quality. For instance, in temperate regions, where species typically display broader ecological tolerances and superior dispersal abilities (Carscadden et al. 2020; Sheard et al. 2020), managing the internal structure of metacommunities should require special attention to landscape connectivity. Conversely, in tropical regions, where ecological specialists with constrained dispersal capacities dominate species pools (Jocque et al. 2010; Carscadden et al. 2020), the internal structure of metacommunities should be better managed by protecting the quality and suitability of chosen habitats. Should these theoretical predictions be validated, it could pave the way for integrating the KCC, a concept derived from metacommunity theory, into the realm of conservation biogeography (*sensu* Whittaker et al. 2005).

Our simulations provided theoretical validation that the proposed analytical framework accurately categorizes communities along a continuum of importance in maintaining the internal structure of metacommunities (Figures 5.3, SI 5.V- VI). That is, it assigns high values of keystoneness to communities that, when removed, promote cascading changes in extinction and colonization dynamics that ultimately increase biodiversity differentiation over time (i.e., Δ Spatial $\beta > 0$). In contrast, it assigns low values of keystoneness to communities that, when removed, cause little to no changes in the internal structure of the remaining communities in the metacommunity (i.e., Δ Spatial $\beta \approx 0$). By testing the framework across various combinations of species traits, we demonstrated that it correctly identifies community keystoneness within a parameter space that reflects different types of metacommunity dynamics (archetypes).

Note that the shape, but not the direction, of the relationship between Δ Spatial β and the keystoneness of extirpated communities varied depending on whether Δ Spatial β was estimated based on incidence (Figure SI 5.V) or abundance (Figure 5.3) data. For instance, in species sorting metacommunities, we did not observe a positive relationship between keystoneness and Δ Spatial β when estimated using abundance data. However, a strong positive relationship emerged when considering incidence data (Figure SI 5.V). Estimates of Δ Spatial β based on abundance and incidence data assign different weights to the losses and gains of species that were locally rare but had a high occupancy at the metacommunity scale. As such, when species distribution across the metacommunity is strongly constrained by habitat selection and dispersal limitation, only a few species can keep small populations across multiple local communities through the dispersal of individuals from source communities. The removal of such source (keystone) communities causes cascading losses of these small-sized populations of widespread species across the entire metacommunity, ultimately increasing metacommunity differentiation (Δ Spatial $\beta > 0$). Such losses of small populations supported by dispersal will have a negligible effect on estimates of Δ Spatial β based on abundance data but a significant effect on estimates based on incidence data.

These findings demonstrate that incorporating different types of data when estimating Δ Spatial β provides a more in-depth understanding of the effects of community removal on the internal structure of metacommunities (Tatsumi et al. 2022).

Similarly, in the extreme case of neutral metacommunities, the relationship between keystoneness and Δ Spatial β could only be observed when we increased the magnitude of disturbance events and extirpated multiple (groups) communities rather than single communities during removal experiments (Figure SI 5.VI). This suggests that within neutral metacommunities, the impact of single local communities on the broader metacommunity structure is less pronounced.

5.5.2 Assumptions and considerations for the use of the proposed framework

An important consideration underlying our simulation model is its assumption that metacommunities were sampled at stable equilibrium and that all communities have been equally and sufficiently sampled. In this ideal circumstance, our simulation models indicate that our framework can accurately estimate community keystoneness based on single spatial snapshot surveys. However, these assumptions can be challenging to validate in real-world empirical data. Thus, we propose guidelines to assist researchers in assessing the applicability of our framework.

First, we encourage potential users of our framework to meticulously evaluate the adequacy of their sampling efforts before estimating the keystoneness of focal communities (see example in Figure SI 5.II that used the approach proposed by Anderson & Santana-Garcon, 2015). Second, we advocate for considering temporally replicated data in keystoneness estimates. Extending the timeframe of sampling designs reduces biases in inferring metacommunity patterns and processes that arise from the false absence of species often noted in snapshot surveys (Record et al. 2021). Moreover, by considering temporally replicated data on estimates of community keystoneness, one can examine the temporal consistency of community keystoneness patterns. Consistent keystoneness patterns over time (as observed in the empirical example in Figure 5.4) would indicate metacommunity stability. Conversely, if keystoneness patterns fluctuate unpredictably over the sampling duration, it would indicate that the metacommunity is not in stable equilibrium. In this case, it would be prudent to refrain from utilizing estimates of community keystoneness derived from this framework to guide conservation and management plans. This is not inherently a limitation of our proposed framework as non-stable metacommunities are intrinsically challenging to conserve and manage.

5.5.3 Insights gained from empirical analyses

In this study, we have presented compelling correlational evidence that light pollution can impact the dynamics of moth communities (Figures 5.4 and 5.5). We showed that light pollution was positively associated with local diversity. These results align with experimental and observational studies, providing further support that light pollution, both in its direct and indirect forms, disrupts the spatial orientation of moths by attracting and trapping them into artificially lit localities (see Boyes et al. 2021 and references within).

Even though light pollution seems to have a positive effect on local diversity, it has been shown to decrease the fitness of moth populations by exposing them to opportunistic predators and disrupting their life cycle (Firebaugh and Haynes 2019). Consequently, at least in the case of this moth metacommunity, we provided evidence that the commonly held assumption that local diversity is a proxy of habitat quality/suitability can be violated when anthropogenic stressors affect species distributions (Delibes et al. 2001). In contrast, we showed that, after controlling for the effects of other predictors, light pollution was negatively correlated with community keystoneness. This indicates that the protection of patches where the night sky is less affected by light pollution are critical to the maintenance of the internal structure of this moth metacommunity.

The relevance of our findings is twofold; first, they offer further support to other observational studies indicating that light pollution, particularly in its indirect form, transcends the physical boundaries of protected areas and can persistently impact the dynamics of insect populations and communities (Vaz et al. 2021; Khattar et al. 2022). These findings are especially concerning given the alarming historical and predicted increases in light pollution at global scales (Kyba et al., 2017). Second, they illustrate that using alpha diversity to inform inferences about the conservation value of communities can lead to allocating conservation efforts to areas of high mortality and low fitness. Given the alarming increase in the influence of anthropogenic stressors on the redistribution of biodiversity, we argue that the relationship between diversity and habitat quality is likely becoming weaker (less positive) across taxa at varying spatial scales (e.g., Viana and Chase 2022).

5.5.4 Keystoneness and the conservation values of communities

Most of the research on the KCC has been conducted on micro and mesocosms experimental studies (Resetarits et al. 2018; Yang et al. 2020). The few attempts carried out in natural metacommunities considered estimates of local richness and regional uniqueness (i.e., unique compositions) as proxies of community's conservation value (e.g., Ruhí et al. 2017; Hitchman et al. 2018; Sullivan et al. 2021; Dansereau et al. 2022). The rationale is that protecting locally rich or unique communities can counteract immediate species richness declines at regional scales, underscoring the critical role of these communities as keystone in maintaining overall biodiversity. While estimates of keystoneness based on community contribution to regional richness are valuable, it is important to recognize that management and conservation plans often encompass goals beyond preserving regional species numbers (Capmourteres and Anand 2016). This is specifically relevant when regional diversity is artificially inflated due to the introduction of invasive species (Venevskaia et al. 2013).

As such, despite the implicit differences in conservation value that the words "idle" and "keystone" evoke, we argue that estimates of community keystoneness should not be readily thought of as a proxy of the conservation value of any given local community. Both keystone and idle communities can have a high or low conservation value depending on whether their protection contributes to the success of conservation and management planning. For instance, keystone communities may have a decreased conservation value when their extirpation is necessary to

disrupt colonization dynamics that can ensure the regional dominance of unwanted invasive species (Brown and Barney 2021). Conversely, the conservation value of keystone communities should be high if the goal is to preserve the current structure of metacommunities (assuming metacommunities in equilibrium). These examples should clearly illustrate that community keystoneness and conservation value are distinct concepts that may or may not align depending on specific contexts. Therefore, it is critical to note that the analytical framework proposed and validated in this study should not be regarded as a standalone metric of conservation value. Rather, it should be seen as an analytical tool that can assist managers and conservation biologists in achieving proposed specific goals, whether it involves halting or promoting changes in the internal structure of metacommunities in space and over time.

5.6 Supplementary Information

5.6.1 Empirical data and supporting analyzes
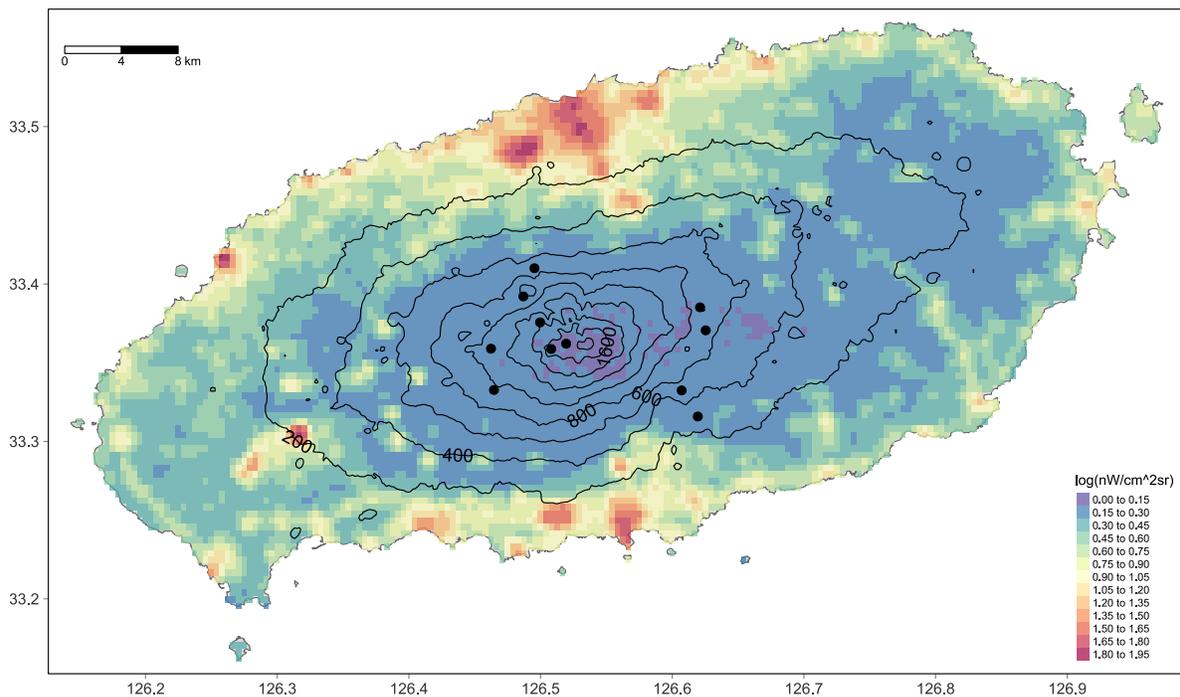
Figure SI. 5.I: Artificial light at night (ALAN) in the Jeju Island, South Korea. Color scale depicts levels of radiance (log-transformed to facilitate visualization) captured by the visible infrared imaging radiometer day-night band (VIIRS DNB) satellite. Black dots represent the sampling sites within the Mount Hallasan National Park.
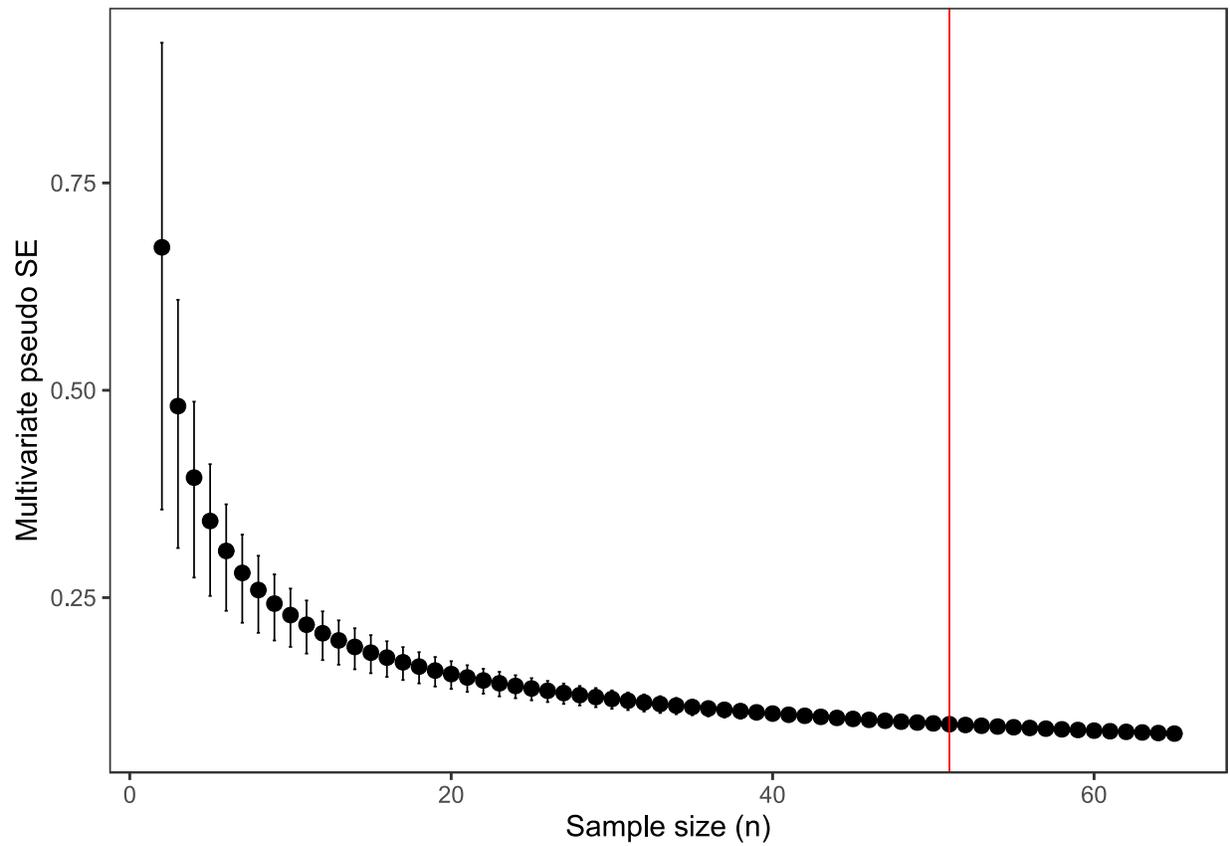
Figure SI. 5.II: Sampling sufficiency estimated through the decrease in Multivariate pseudo SE (Anderson and Santana-Garcon. 2015). The red vertical line represents the minimum sample size above which the decreases in Multivariate pseudo SE with additional samples are not significantly different.

5.6.2 A brief description of core concepts in graph theory

Considering a graph (S), we call its adjacency matrix a *n x n* symmetric matrix where *n* is the number of nodes and $i, j \in \{1, \ldots, n\}$ are their indices. The entries of the adjacency matrix represent the presence (or weights) of the edges between nodes. If node i is connected to node j (i.e., if $edge_{i,j} \neq 0$), we say that this two nodes are adjacent. We call the degree of node "i" the total number of (or sum of weights) of edges that connect i to other nodes. The degree matrix of S is a *n x n* and the main diagonal entries represent the degrees of each node.

The Laplacian matrix of S is given by the difference between its adjacency matrix and its degree matrix. An important characteristic of Laplacian matrices is that they are symmetric and their eigenvalues are all positive (i.e., it is a positive definite matrix). This is relevant because the spectral decomposition of a Laplacian matrix can capture many properties of the graph it represents. For instance, while its smallest eigenvalue is always zero, its second smallest eigenvalue (called algebraic connectivity) informs how difficult it is to disconnect a network. That is, considering two networks S_1 and S_2, if S_1 has fewer connections than S_2, then the algebraic connectivity of S_1 (i.e., the second smallest eigenvalue of its Laplacian matrix) is smaller than that of S_2.

5.6.3 Demonstrating the statistical independence between community Keystoneness and estimates of local diversity

Estimates of keystoneness that are statistically independent or weakly correlated with local diversity can yield standalone insights about local communities that can be instrumental in fostering further inferences about their conservation value and importance at the metacommunity scale. Here, we used simple simulations to establish the baseline correlation between our proposed estimates of community keystoneness and local diversity in the absence of any confounding deterministic ecological process (e.g., habitat selection, dispersal limitation, anthropogenic stressors) dictating community composition. If our proposed estimate of community keystoneness is not inherently influenced by local diversity, one should not observe strong correlations between these two estimates when community composition is simply the outcome of stochastic sampling from regional species pools.

To test that, we simulated 50 regional species pools whose species abundance distribution followed a log-normal distribution. Each regional pool was composed of 1000 individuals distributed across 50 species. These individuals were then randomly allocated across 30 sample sites (communities). For each simulated metacommunity, we estimated local community diversity using Hill-numbers of orders 0 to 2 (i.e., species richness, Hill-Shannon diversity, and Hill-Simpson diversity).

We used six different similarity indexes to estimate community keystoneness for each metacommunity: Jaccard-Chao, Bray Curtis-Chao, Hellinger, Chord, Canberra, and Kulczynski. These indexes have desired mathematical properties when measuring community compositional variation (see more in Legendre and De Cáceres 2013; Cao et al. 2021). Similarity matrices were

126

estimated as the difference between the potential maximum dissimilarity obtained for each index (i.e., 1 for the Jaccard-Chao, Bray Curtis-Chao, Canberra, and Kulczynski indexes; $\sqrt{2}$ for Hellinger and Chord) and the observed dissimilarities calculated across pairs of simulated communities. Then we used these compositional similarities to estimate community keystoneness following the steps described in the main text (see "An analytical framework for community keystoneness", and Figure 5.1). Testing the relationship between keystoneness and local diversity considering different dis(similarity) indexes can help us to identify those better suited to yield statistically independent estimates of community keystoneness.

We also aimed to contrast our proposed estimates of community keystoneness with other estimates that rely on community (dis)similarity to evaluate a community's influence on regional metacommunity patterns. Notably, the 'Local contribution to beta-diversity" index (hereafter LCBD) has been widely adopted (Legendre and De Cáceres 2013). When estimated through composition dissimilarity matrices, LCBD quantifies the compositional uniqueness of ecological communities within a metacommunity and has frequently been employed as an indicator of communities' conservation value (Ruhí et al. 2017*b*; Hill et al. 2021; Perez et al. 2023). The fundamental premise of these studies is that by protecting unique communities, one can mitigate immediate losses of species richness at regional scales.

To assess the baseline association between keystoneness, local diversity, and LCBD, we estimated their pairwise correlations across the 50 simulated metacommunities. Our simulation framework demonstrated that the proposed estimates of community keystoneness are only weakly positively correlated with local diversity, regardless of the compositional similarity index used to estimate it (average ρ across all similarity and local diversity indexes = 0.13, Figure SI 5.III). This result indicates that any strong association between these two variables observed in empirical data must be underpinned by confounding environmental, biotic, and spatial variables and not by any statistical association contrived by how they are estimated.

In contrast, we observed a moderate correlation between LCBD and local diversity (average ρ across all = -0.43, Figure SI 5.III), albeit it varied considerably among beta-diversity indexes. These findings suggest that previous studies that invoked different ecological mechanisms to explain observed negative relationships between LCBD and local diversity may have failed to account for their inherent statistical association (e.g., Hill et al. 2021; Dansereau et al. 2022; Perez et al. 2023).

Lastly, we detected a moderate negative correlation between LCBD and keystoneness (average ρ across all similarity and local diversity indexes = -0.38, Figure SI 5.IV). A relationship between these two metrics is expected because they are derived based on identical (dis)similarity indexes. Nevertheless, the weak correlation of our proposed estimate for keystoneness with various local diversity estimates, irrespective of the chosen (dis)similarity index, renders it more suitable for studies exploring communities' significance at the metacommunity scale. This is particularly

useful when local diversity serves as an unreliable indicator of habitat quality due to maladaptive habitat selection induced by anthropogenic stressors.
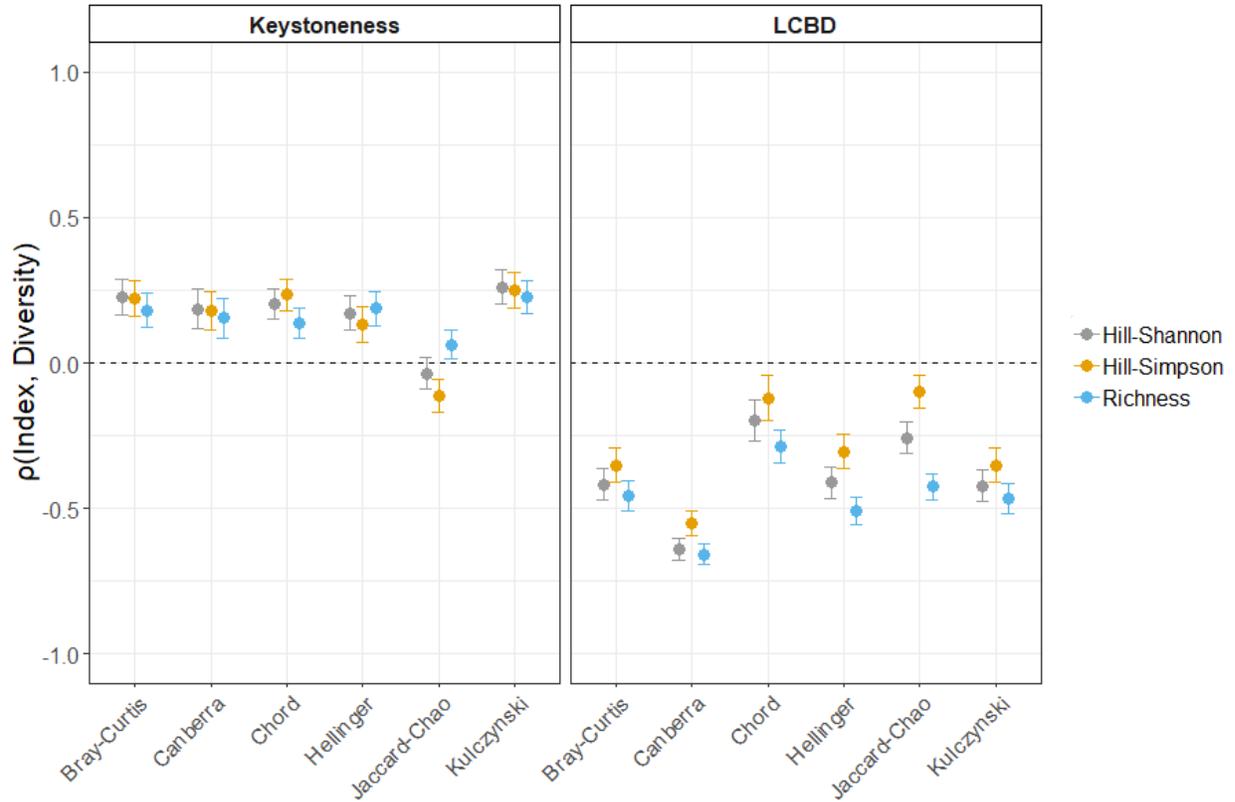
Figure SI. 5.III: Spearman correlation (ρ) between community keystoneness and local diversity (left panel) and LCBD and local diversity (right panel). Keystoneness and LCBD were estimated based on different (dis)similarity indexes (x-axis). Local diversity was estimated through Hill-numbers of different orders (q=0=Richness, q=1=Hill Shannon, q=2= Hill Simpson). Points represent the average correlation across 50 simulated metacommunities. Whiskers represent the confidence interval at 95%. Dashed line represents ρ correlation coefficient =0.
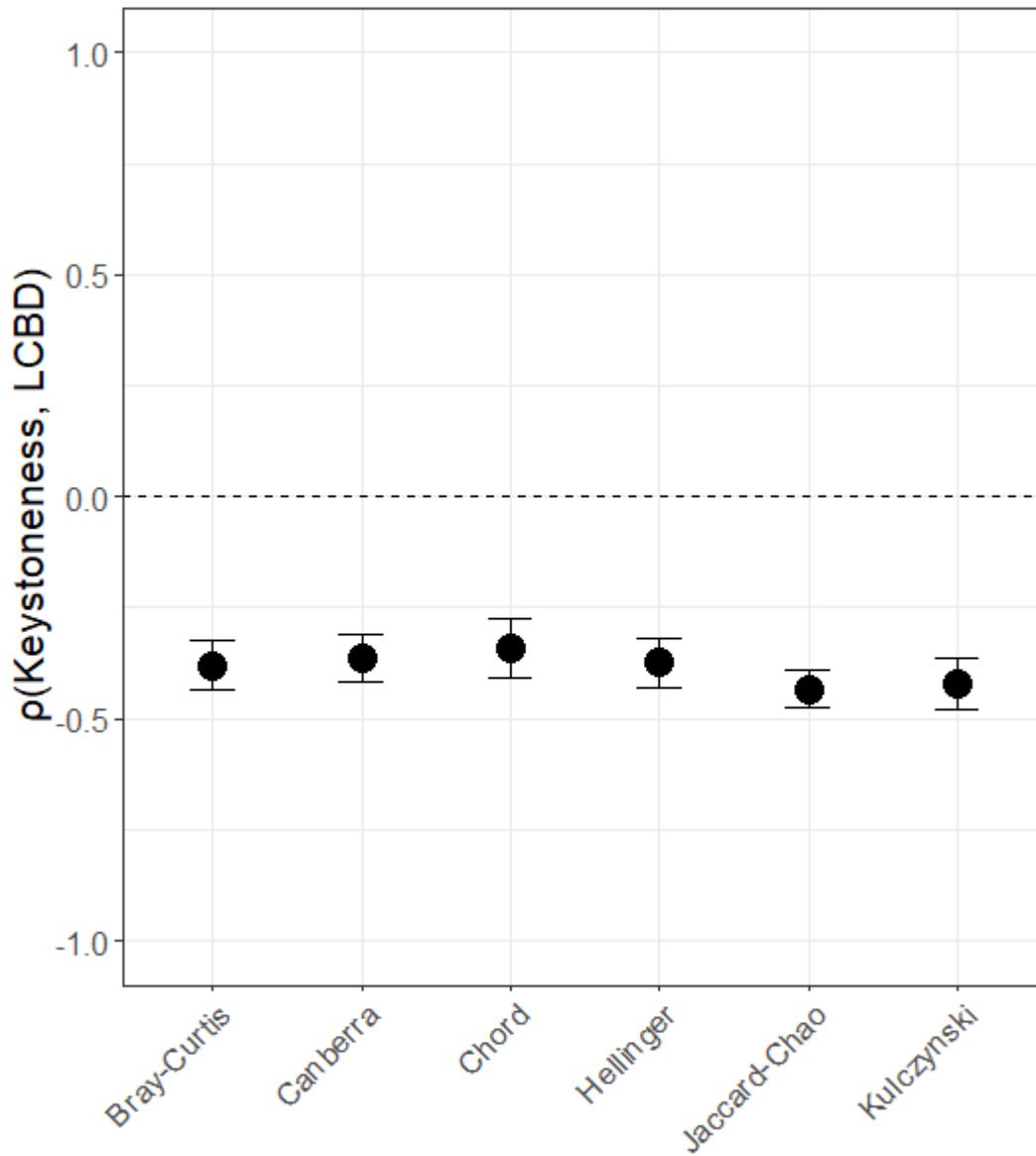
Figure SI. 5.IV: Spearman correlation (ρ) between community keystoneness and LCBD estimated from different (dis)similarity. Points represent the average correlation across 50 simulated metacommunities. Whiskers represent the confidence interval at 95%. Dashed line represents ρ correlation coefficient =0.

5.6.4 Extended description of mechanistic metacommunity simulations

We started by randomly distributing 30 habitat patches in a geographic space defined by x and y coordinates ranging from 0 to 60. The environmental conditions across patches were spatially autocorrelated and scaled to range in the interval [0,5]. Spatial habitat autocorrelation in the environment was given by random draws from a multivariate normal distribution (mu=0) with a covariance matrix that incorporates an exponential decay of environmental similarity with geographic distance and is modulated by the parameter φ (here, fixed at 0,1). In our simulations, individuals could only move between physically connected patches (see below). The degree into which any given two patches are physically connected decays exponentially with Euclidean distance (ΔS) according to the following kernel function:

$Connectivity_{i,j} = \exp(-0.5 * \Delta S_{ij})$ (eq. SI 5.1)

$Connectivity$ values below a fixed threshold (here 10^{-3}) where truncated to zero, so that individuals could not move between the focal pair of patches (as in Fournier et al. 2017). All values above this threshold were assumed to be equal to 1. As such, we could create a network of habitat patches that were environmentally heterogenous and differed in the number of edges (connections) shared with neighboring patches.

Population dynamics of 100 species across the 30 patches were discrete in time and spatially explicit. Community assembly at time t was the outcome of three sequential steps: 1) within-patch dynamics (i.e., environmental selection, demographic stochasticity, and competition at the intra and interspecific levels); 2) emigration; 3) immigration.

Considering that $N_{i,j,t}$ is the abundance of species i in site j in time t, population dynamics is governed by:

$N_{i,j,t} = Poisson(N_{i,j,t-1} * P_{i,j,t}) - E_{i,j,t} + I_{i,j,t}$ (SI 5.2)

The first term of SI 5.2 is a modified version of the commonly used Beverton-Holt model that models time discrete population growth as a function of within-patch selection and ecological drift (i.e., demographic stochasticity). $P_{i,j,t}$ is the local performance (i.e., growth rate) of species i when conditioned to competition and habitat selection in site j and time t, and is modeled as:

$P_{i,j,t} = R_{i,j,t} * \frac{1}{(1+\alpha_{intra}N_{i,j,t}+\alpha_{inter}\sum_{k\neq i}^{S} N_{k,j,t})}$ (SI 5.3)

where $R_{i,j,t}$ is the influence of local environmental conditions on species performance given by a Gaussian response:

$R_{i,j,t} = R.max * \exp\left(\frac{-(Env_{j,t}-\mu_i)^2}{2\sigma^2}\right)$ (SI 5.4)

where $R.max$ represents the maximum intrinsic growth rate and it was assumed to be equal across species and is fixed at 3. This assumption ensured that all species had the same maximum growth

when local conditions are optima and competition is negligible (see below). Env_{jt} represents local abiotic conditions, μ_i is species environmental optimum, and σ is species environmental tolerance (niche breadth). By manipulating the values of σ, we could model metacommunities wherein species were more (lower values of σ) or less (higher values of σ) sensitive to spatial and temporal environmental variation. Here we ran different simulation scenarios where all species had either a narrower ($\sigma = 1$), intermediate ($\sigma = 3$) or broader ($\sigma = 5$) niche breadths. Contrasting simulation outcomes across levels of niche breadth allowed us to understand how the degree of environmental specialization of species pools influenced the mechanisms underpinning community keystoneness (see Figures 5.2 and 5.3).

The term on the right of eq SI 5.3 models the effects of density-dependent competition at the intraspecific and interspecific levels on population dynamics. α_{intra} represents the per capita effects of species i on itself whereas α_{inter} is the per capita effect of all other species on the local performance of i. Here we assumed that a stabilizing structure of competition dynamics by assuming that $\alpha_{intra} > \alpha_{inter}$. Assuming a stabilizing competition is relevant because it facilitates stable local coexistence by allowing locally rare species to keep positive population growth when the populations of dominant competitors are in equilibrium at higher abundances (the so called invasion criterium for coexistence (HilleRisLambers et al. 2012; Shoemaker and Melbourne 2016). α_{intra} and α_{inter} were fixed at 1/200, and 1/400 across all species to allow species to coexist at relative high abundances provided that local habitat conditions were suitable. Previous tests (not shown) indicated that assuming different values for the inequality $\alpha_{intra} < \alpha_{inter}$ did not change coexistence patterns but changed the equilibrium abundances of locally coexisting species. We acknowledge that other types of competitive structure are observed in nature (equalizing: $\alpha_{intra} = \alpha_{inter}$; destabilizing $\alpha_{intra} < \alpha_{inter}$), but understanding their effects on estimates of community keystoneness goes beyond the scope of the present study.

We incorporated the influence of ecological drift on local birth and survival by drawing the final local abundance of species i from a Poisson distribution (eq. SI 5.2) whose mean is given by the deterministic influence of abiotic density-dependent and biotic density-independent selection on population dynamics (as in Thompson et al. 2020).

Individuals able to persist in the local community after within-patch selection and drift at time t could then disperse. The total number of emigrants of species i leaving site j in time t ($E_{i,j,t}$) is given by binomial trials whose size is equal to the outcomes of within-patch dynamics (first term of eq. SI 5.2) and whose probability of success were given by the parameter θ. θ represent a species dispersal rate (i.e., propensity to emigrate). By manipulating the values of θ we could simulate species more (higher levels of θ) less (lower levels of θ) propense to leave their natal patch. Here we ran different simulation scenarios where all species had either a lower ($\theta = .1$), Intermediate ($\theta = .2$) or higher ($\theta = .3$) propensity to emigrate. By doing so, we could understand how dispersal rates influenced the association between estimates of keystoneness and patch attributes (Figure 5.2). It also allowed us to understand how the ability of the proposed estimate

for community keystoneness to correctly identify keystones and burden communities changes across levels of dispersal (Figure 5.3).

The total number of immigrants of species i arriving at site j in time t ($I_{i,j,t}$) is determined by randomly sampling spatial emigrants from neighboring patches ($E_{i,k,t}$) with weights (probabilities) that are equal to their *connectivity* level with j (i.e., equal to either 0 or 1).

We replicated (50 times) this simulation setup across combinations of the three levels of species` niche breadth and dispersal rates, yielding a total of 9 simulation scenarios and 450 simulation iterations. In each simulation iteration, population dynamics of the 100 initial species in the regional pool were carried across all 30 patches over 1000-time steps plus 300 time steps after the removal of focal patches (or groups of patches) during removal experiments. We seeded each patch in the first time step (t_1) with species local (patch level) populations randomly drawn from a Poisson distribution with $\lambda = 0.5$. Then from t_2-t_{100}, populations randomly drawn from a Poisson distribution with $\lambda = 0.1$. This allowed the chance for establishment and population growth for all species, provided that local abiotic and biotic conditions were suitable. In addition, the random placement of species populations across patches allowed those with similar habitat conditions to develop communities with dissimilar compositions due to random dispersal and priority effects (Thompson et al. 2020).

5.6.5 Removal experiments: Extended results
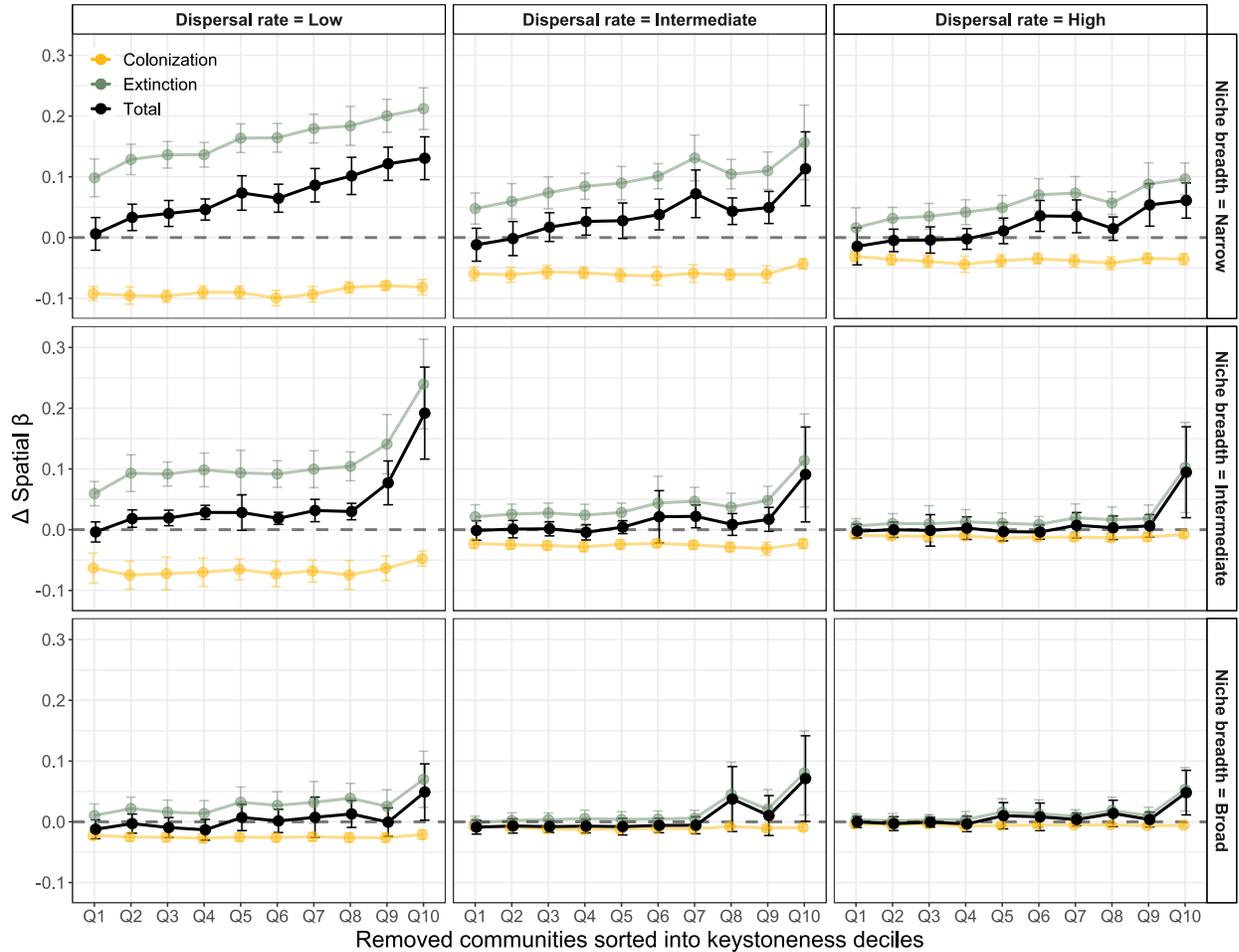
Figure SI. 5.V: Results of simulated removal experiments. Impact of community removal on temporal changes in the internal structure of metacommunities (total Δ Spatial β) caused by losses (through extinctions) and gains (through colonization) in species local occurrences. Communities are sorted into deciles of keystoneness (i.e., Q1 = 3 communities with the lowest levels of keystoneness, Q10 = 3 communities with the highest keystoneness). Losses and gains on local species abundances over time sum up to total Δ Spatial β. Points represent average effect sizes (n=50) across simulation replicates per combination of levels of niche breadths and dispersal rates in the simulated regional pool. Whiskers represent 95% confidence intervals. Dashed line marks Δ Spatial β = 0
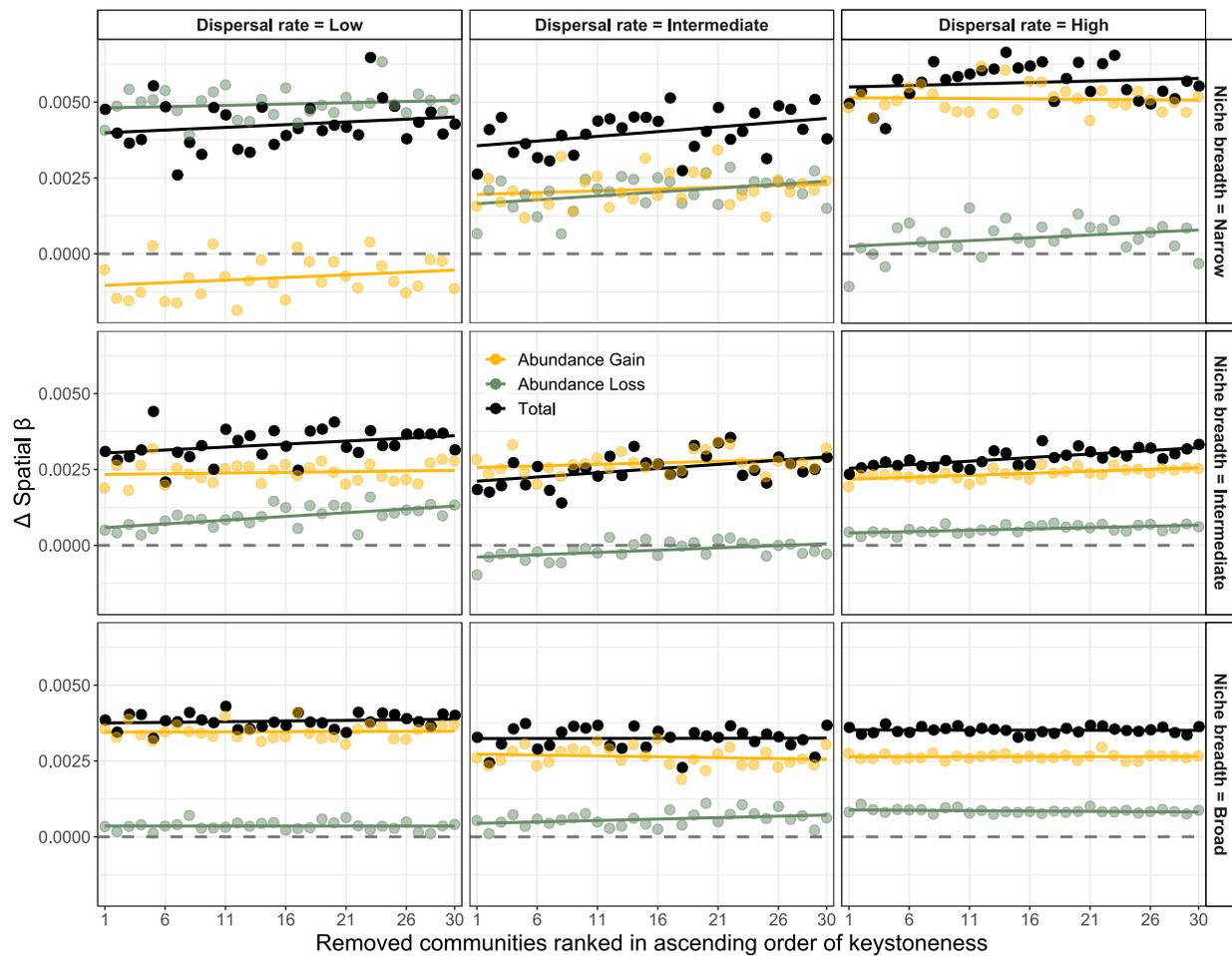
Figure SI. 5.VI: Results of simulated removal experiments. Impact of community removal on temporal changes in the internal structure of metacommunities (total Δ Spatial β) caused by losses and gains in species local abundances. Communities are ranked based on their level of keystoneness. Losses and gains on local species abundances over time sum up to total Δ Spatial β. Points represent average effect sizes (n=50) across simulation replicates per combination of levels of niche breadths and dispersal rates in the simulated regional pool. Dashed line marks Δ Spatial β = 0

Chapter 6 : Concluding remarks, assumptions, and future directions

"Theory and fact are equally strong and utterly interdependent; one has no meaning without the other. We need theory to organize and interpret facts, even to know what we can or might observe. And we need facts to validate theories and give them substance." (Gould, 1998, p. 155)

In this thesis, we set out to demonstrate that acknowledging the influence of species pools-mediated top-down and landscape-mediated bottom-up controls over community assembly helps us to better understand the ecological causes driving spatiotemporal context-dependence on community dynamics. Given that the scientific value of any conceptual model lies in its ability to generate robust predictions about the functioning and dynamics of natural systems (Houlahan et al. 2016), I employed process-based metacommunity simulation models to derive predictions regarding how the interplay between top-down and bottom-up controls governs: (i) Spatiotemporal shifts in community composition (Chapters 2 and 5); (ii) The prevailing life-history strategies in metacommunities (Chapters 3 and 4); (iii) The relative importance of assembly mechanisms across space and time (Chapter 2) and throughout large-scale ecological gradients (Chapter 3), and; (iv) The trajectories of metacommunities (towards differentiation or homogenization) when faced with natural or anthropogenic stressors (Chapter 5). Subsequently, by assessing the similarities between the predictions derived from our simulation models and patterns observed in empirical data, I could provide empirical validation that conceptual frameworks considering bottom-up and top-down controls on community assembly can enhance our comprehension of community dynamics.

I deliberately focused on a reduced number of landscape features and species pool attributes to investigate the interactive effects of bottom-up and top-down controls on community assembly. While this simplification can enhance understanding, we acknowledge that overly simplified frameworks can be stripped away from the ecological phenomena they are meant to represent and foster comprehension (Bergelson et al. 2021). Future research should build upon the frameworks presented in this thesis, exploring how aspects of landscapes and species pools other than the ones investigated here can control the assembly of ecological communities. For example, recent findings indicate that the spatial frequency of habitat conditions in a landscape affects the coexistence and prevalence of ecological specialists versus generalists (Fournier et al. 2020). Moreover, for the sake of simplicity, we considered landscape characteristics to be immutable over time (but see a slightly different approach in the simulated removal experiments described in Chapter 5). However, landscapes are constantly changing due to natural and anthropogenic disturbances. Such dynamic landscapes can also control the assembly process by inducing predictable shifts in metacommunity dynamics by altering species interaction rates, habitat patch sizes, and overall landscape connectivity (Fernandes et al. 2013; Li et al. 2023; Marco Palamara et al. 2023).

One of the primary goals of this thesis was to investigate the emergence of a feedback loop between bottom-up and top-down controls on community assembly. That is, we investigated how landscape attributes filter "Mainland Species pools" into "Metacommunity species pools" (Chapters 3 and Chapter 4) and how the composition of metacommunity pools determines the importance of different assembly mechanisms (Chapter 3). It was done by starting with a large diversity of fixed/non-mutable traits and allowing metacommunity dynamics (species-species and species-environment interactions) to select the most successful traits in the metacommunity. This modeling choice implies the assumption that ecological mechanisms operating at fine spatiotemporal scales can also scale up to drive patterns of trait distribution at regional scales.

However, our modeling design is not the only way to investigate the feedback loop between top-down and bottom-up controls on community assembly. For instance, one could also investigate how differences in the genesis of species pools that colonize a given landscape would lead to different metacommunity patterns (e.g., Thompson *et al.* 2020). This approach consists of generating species pools with distinct trait distributions (i.e., all species with the same trait values, but these trait values changed across simulation scenarios) and allowing them to colonize the same landscape to investigate the resulting metacommunity patterns. This framework is aligned with the assumption that idiosyncratic differences in the broad-scale evolutionary and historical mechanisms that generated species pools are the primary driver of variation in metacommunity patterns (Taylor *et al.* 1990; Lessard *et al.* 2012).

Similarly, one can investigate the feedback loops between both types of control on community assembly by investigating how landscape features influence trait evolution (either by a mutation-selection process or local adaptation) and how newly evolved trait states compete with already established ones to dominate metacommunities (Goodnight 2011; Leibold et al. 2019). This approach would be aligned with the assumption that landscape characteristics determine the "evolutionary stable communities" in metacommunities (*sensu* Edwards et al. 2018), i.e., communities that emerge in a given landscape that cannot be displaced by novel or initially rare alternative species.

In this thesis, we enhance the heuristic value of community assembly theory by proposing a reduced number of general (high-level) ecological conditions (here, landscapes and species pools) that modulate the relative importance of different fundamental assembly processes (i.e., selection, dispersal, and drift). However, what remains underexplored is how each fundamental assembly process (e.g., selection) influences, or is influenced by, the effects of the other processes on community assembly (e.g., drift and dispersal). For instance, a growing body of literature suggests that dispersal, one of the three fundamental processes in the community assembly theory, mediates the relative importance of drift and selection in community assembly (Ron et al. 2018). This is because dispersal can ultimately modulate the size of local populations and communities by mixing individuals of different communities in a large metacommunity. When dispersal results in reduced community sizes (e.g., emigration higher than immigration), the importance of ecological drift in community assembly increases because a few "negative" stochasticity events

(e.g., disease outbreaks or a particularly unsuccessful reproductive season) can have a large impact on populations' sizes, leading to significant shifts in community composition (Orrock and Watling 2010). In contrast, if dispersal increases the size of populations, the law of large numbers comes into play, providing a buffer against drastic composition changes due to random demographic events (Siqueira et al. 2020). In this thesis, we extend the understanding of the non-independence of dispersal and other high-level assembly processes by showing how the influence of landscape and species pool characteristics on ecological selection (i.e., the outcome and of species-species and species-environment interactions) determine the dispersal strategies of species in metacommunities (Chapter 4). However, there is a pressing need for further research to accommodate the mutual interference of high-level processes within the foundational framework of community assembly theory.

In conclusion, this thesis explores the interplay between species pool-mediated top-down and landscape-mediated bottom-up controls in shaping community assembly. While our conceptual frameworks and simulation models provided significant insights, they are mere simplifications of the functioning of natural systems. These simplified frameworks highlight the need for future research to delve deeper into the nuances not investigated here. Notably, the dynamic nature of landscapes and the intricate interdependence between assembly processes, such as dispersal, drift, and selection, warrant further exploration. Nevertheless, this thesis emphasizes the importance of integrative frameworks that assume the existence of bottom-up and top-down controls over community dynamics to investigate the dynamics of ecological communities and pave the way for further studies to refine our understanding about the context-dependent nature of community assembly. Beyond its theoretical significance, this knowledge is crucial for predicting how the impact of human activities on landscapes and species pools can alter the structure, dynamics, and natural regulation of ecological communities.

References

Adler, P. B., and J. M. Levine. 2007. Contrasting relationships between precipitation and species richness in space and time 221–232.

Alzate, A., and R. E. Onstein. 2022. Understanding the relationship between dispersal and range size. Ecology Letters 2303–2323.

Anderson, M. J., and J. Santana-Garcon. 2015. Measures of precision for dissimilarity-based multivariate analysis of ecological communities. Ecology Letters 18:66–73.

Baines, C. B., S. Diab, and S. J. McCauley. 2020. Parasitism Risk and Infection Alter Host Dispersal. The American Naturalist 196:119–131.

Bar-Massada, A., R. Kent, and Y. Carmel. 2014. Environmental heterogeneity affects the location of modelled communities along the niche–neutrality continuum. Proceedings of the Royal Society B: Biological Sciences 281:20133249.

Barabás, G., R. D. Andrea, and S. M. Stump. 2018. Chesson ' s coexistence theory. Ecological Monographs.

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

Belmaker, J. 2009. Species richness of resident and transient coral-dwelling fish responds differentially to regional diversity. Global Ecology and Biogeography 18:426–436.

Belyea, L. R., and J. Lancaster. 1999. Assembly rules within a contingent ecology. Oikos 86:402–416.

Bergelson, J., M. Kreitman, D. A. Petrov, A. Sanchez, and M. Tikhonov. 2021. Functional biology in its natural context: A search for emergent simplicity. eLife 10:1–12.

Beverton, R. J. H., and S. J. Holt. 1957. On the Dynamics of Exploited Fish Populations (Fish and F.). Springer Science and Business Media, London.

Bitume, E. V., D. Bonte, O. Ronce, I. Olivieri, and C. M. Nieberding. 2014. Dispersal distance is influenced by parental and grand-parental density. Proceedings of the Royal Society B: Biological Sciences 281.

Blanchet, F. G., P. Legendre, and D. Borcard. 2008. Modelling directional spatial processes in ecological data. Ecological Modelling 215:325–336.

Bonte, D., J. Vanden Borre, L. Lens, and Jean-Pierre Maelfait. 2006. Geographical variation in wolf spider dispersal behaviour is related to landscape structure. Animal Behaviour 72:655–662.

Borcard, D., P. Legendre, and P. Drapeau. 1992. Partialling out the Spatial Component of Ecological Variation. Ecology 73:1045–1055.

Box, G. E. P. 1976. Science and Statistics. Journal of the American Statistical Association

71:791.

Boyes, D. H., D. M. Evans, R. Fox, M. S. Parsons, and M. J. O. Pocock. 2021. Is light pollution driving moth population declines? A review of causal mechanisms across the life cycle. Insect Conservation and Diversity 14:167–187.

Braga, P. H. P., S. Kembel, and P. Peres-Neto. 2023. Historical and contemporary processes drive global phylogenetic structure across geographical scales: Insights from bat communities. Global Ecology and Biogeography 747–765.

Breiman, L. 2001. Random Forests. Machine Learning 45:5–32.

Brown, B. L., and J. N. Barney. 2021. Rethinking Biological Invasions as a Metacommunity Problem. Frontiers in Ecology and Evolution 8:1–23.

Brown, B. L., E. R. Sokol, J. Skelton, and B. Tornwall. 2017. Making sense of metacommunities: dispelling the mythology of a metacommunity typology. Oecologia 183:643–652.

Brown, J. H., J. F. Gillooly, A. P. Allen, V. M. Savage, and G. B. West. 2004. TOWARD A METABOLIC THEORY OF ECOLOGY. Ecology 85:1771–1789.

Büchi, L., P.-A. Christin, and A. H. Hirzel. 2009. The influence of environmental spatial structure on the life-history traits and diversity of species in a metacommunity. Ecological Modelling 220:2857–2864.

Büchi, L., and S. Vuilleumier. 2012. Dispersal strategies, few dominating or many coexisting: The effect of environmental spatial structure and multiple sources of mortality. PLoS ONE 7.

———. 2014. Coexistence of specialist and generalist species is shaped by dispersal and environmental factors. American Naturalist 183:612–624.

Buoro, M., and S. M. Carlson. 2014. Life-history syndromes: Integrating dispersal through space and time. Ecology Letters 17:756–767.

Cadena, C. D., K. H. Kozak, J. P. Gómez, J. L. Parra, C. M. Mccain, R. C. K. Bowie, A. C. Carnaval, et al. 2011. Latitude, elevational climatic zonation and speciation in New World vertebrates. Proceedings of the Royal Society B: Biological Sciences 279:194–201.

Cafri, G., and B. A. Bailey. 2016. Understanding Variable Effects from Black Box Prediction: Quantifying Effects in Tree Ensembles Using Partial Dependence. Journal of Data Science 14:67–96.

Cao, K., J. Svenning, C. Yan, J. Zhang, X. Mi, and K. Ma. 2021. Undersampling correction methods to control γ-dependence for comparing β-diversity between regions. Ecology 102.

Capmourteres, V., and M. Anand. 2016. " Conservation value ": a review of the concept and its quantification 7:1–19.

Carscadden, K. A., N. C. Emery, C. A. Arnillas, M. W. Cadotte, M. E. Afkhami, D. Gravel, S. W. Livingstone, et al. 2020. Niche Breadth: Causes and Consequences for Ecology, Evolution, and Conservation. The Quarterly Review of Biology 95:179–214.

Carstensen, D. W., J. Lessard, B. G. Holt, and M. K. Borregaard. 2013. Introducing the biogeographic species pool 1–9.

Castro, E. 2018. Changing a Brazilian Protected Areas Paradigm: Why Public Use is Not Just Optional. Journal of Park and Recreation Administration 36:129–140.

Catford, J. A., J. R. U. Wilson, P. Pyšek, P. E. Hulme, and R. P. Duncan. 2021. Addressing context dependence in ecology. Trends in Ecology & Evolution xx:1–13.

Cerezer, F. O., A. Machac, T. F. Rangel, and C. S. Dambros. 2022. Exceptions to the rule: Relative roles of time, diversification rates and regional energy in shaping the inverse latitudinal diversity gradient. Global Ecology and Biogeography 31:1794–1809.

Chase, J. M. 2014. Spatial scale resolves the niche versus neutral theory debate. Journal of Vegetation Science 25:319–322.

Chase, J. M., A. Jeliazkov, E. Ladouceur, and D. S. Viana. 2020. Biodiversity conservation through the lens of metacommunity ecology. Annals of the New York Academy of Sciences 1469:86–104.

Chase, J. M., and J. A. Myers. 2011. Disentangling the importance of ecological niches from stochastic processes across scales. Philosophical Transactions of the Royal Society B: Biological Sciences 366:2351–2363.

Chave, J. 2013. The problem of pattern and scale in ecology: What have we learned in 20 years? Ecology Letters 16:4–16.

Chesson, P. 2000. Mechanisms of Maintenance of Species Diversity. Annual Review of Ecology and Systematics 31:343–366.

Chesson, P. L. 1985. Coexistence of competitors in spatially and temporally varying environments. Theoretical Population Biology 28:263–287.

Choi, S. W., and S. H. Na. 2020. Quantitative data from six years (2013-2018) of light trap sampling of macromoths (Lepidoptera) in Mt. Hallasan National Park, South Korea. Biodiversity Data Journal 8.

Clobert, J., M. Baguette, T. G. Benton, and J. M. Bullock. 2012. Dispersal ecology and evolution. Oxford University Press.

Clobert, J., J. F. Le Galliard, J. Cote, S. Meylan, and M. Massot. 2009. Informed dispersal, heterogeneity in animal dispersal syndromes and the dynamics of spatially structured populations. Ecology Letters 12:197–209.

Collins, S. L., and S. M. Glenn. 1991. Importance of spatial and temporal dynamics in species regional abundance and distribution. Ecology 72:654–664.

Cornell, H. V., and S. P. Harrison. 2014. What Are Species Pools and When Are They Important? Annual Review of Ecology, Evolution, and Systematics 45:45–67.

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

Coyle, J. R., A. H. Hurlbert, and E. P. White. 2013. Opposing mechanisms drive richness patterns of core and transient bird species. American Naturalist 181:E83–E90.

da Silva, P. G., J. M. Lobo, M. C. Hensen, F. Z. Vaz-de-Mello, and M. I. M. Hernández. 2018. Turnover and nestedness in subtropical dung beetle assemblages along an elevational gradient. (K. Feeley, ed.)Diversity and Distributions 24:1277–1290.

Dale, M. R. T. 2017. Applying Graph Theory in Ecological Research. Cambridge University Press.

Dale, V. H., C. M. Crisafulli, and F. J. Swanson. 2005. 25 Years of ecological change at Mount St. Helens. Science 308:961–962.

Dansereau, G., P. Legendre, and T. Poisot. 2022. Evaluating ecological uniqueness over broad spatial extents using species distribution modelling. Oikos 2022:1–13.

De Meester, N., S. Derycke, A. Rigaux, and T. Moens. 2015. Active dispersal is differentially affected by inter- and intraspecific competition in closely related nematode species. Oikos 124:561–570.

Delibes, M., P. Gaona, and P. Ferreras. 2001. Effects of an Attractive Sink Leading into Maladaptive Habitat Selection. The American Naturalist 158:277–285.

Diamond, J. M. 1975. Assembly of species communities. Pages 342–444 *in*Ecology and evolution of communities. Havard University Press.

Dray, S., D. Bauman, F. G. Blanchet, D. Borcard, S. Clappe, G. Guenard, T. Jombart, et al. 2022. "Adespatial: Multivariate Multiscale Spatial Analysis (version 0.3-7)."

Dray, S., P. Legendre, and P. R. Peres-Neto. 2006. Spatial modelling: a comprehensive framework for principal coordinate analysis of neighbour matrices (PCNM). Ecological Modelling 196:483–493.

Economo, E. P. 2011. Biodiversity Conservation in Metacommunity Networks: Linking Pattern and Persistence. The American Naturalist 177:E167–E180.

Edwards, K. F., C. T. Kremer, E. T. Miller, M. M. Osmond, E. Litchman, and C. A. Klausmeier. 2018. Evolutionarily stable communities: a framework for understanding the role of trait evolution in the maintenance of diversity. Ecology Letters 21:1853–1868.

Engel, T., S. A. Blowes, D. J. McGlinn, F. May, N. J. Gotelli, B. J. McGill, and J. M. Chase. 2020. Resolving the species pool dependence of beta-diversity using coverage-based rarefaction. bioRxiv 1–21.

Evans, J. S., and M. A. Murphy. 2018. rfUtilities.

Fernandes, I. M., R. Henriques-Silva, J. Penha, J. Zuanon, and P. R. Peres-Neto. 2013. Spatiotemporal dynamics in a seasonal metacommunity structure is predictable: the case of floodplain-fish communities. Ecography no-no.

Ferrier, S., G. Manion, J. Elith, and K. Richardson. 2007. Using generalized dissimilarity modelling to analyse and predict patterns of beta diversity in regional biodiversity assessment. Diversity and Distributions 13:252–264.

Firebaugh, A., and K. J. Haynes. 2019. Light pollution may create demographic traps for nocturnal insects. Basic and Applied Ecology 34:118–125.

Fitzpatrick, M. C., N. J. Sanders, S. Normand, J.-C. Svenning, S. Ferrier, A. D. Gove, and R. R. Dunn. 2013. Environmental and historical imprints on beta diversity: insights from variation in rates of species turnover along gradients. Proceedings. Biological sciences 280:20131201.

Flinte, V., E. Hentz, B. M. Morgado, A. C. do M. Lima, G. Khattar, R. F. Monteiro, and M. V. de Macedo. 2015. Biology and phenology of three leaf beetle species (Chrysomelidae) in a montane forest in southeast Brazil. ZooKeys 2015:119–132.

Flinte, V., M. V. Macedo, and R. F. Monteiro. 2009. Chrysomelids And Their Host Plants Along An Altitudinal Gradient In An Atlantic Rain Forest In The State Of Rio De Janeiro, Brazil. Pages 31–56 *in* P. Jolivet, J. Santiago-Blay, and M. Schmitt, eds. Research on Chrysomelidae, Volume 2 (Vol. 2). BRILL.

Fournier, B., N. Mouquet, M. A. Leibold, and D. Gravel. 2016. An integrative framework of coexistence mechanisms in competitive metacommunities. Ecography 1–12.

———. 2017. An integrative framework of coexistence mechanisms in competitive metacommunities. Ecography 40:630–641.

Fournier, B., H. Vázquez-Rivera, S. Clappe, L. Donelle, P. H. P. Braga, and P. R. Peres-Neto. 2020. The spatial frequency of climatic conditions affects niche composition and functional diversity of species assemblages: the case of Angiosperms. (C. Violle, ed.)Ecology Letters 23:254–264.

Freestone, A. L., and B. D. Inouye. 2015. Nonrandom community assembly and high temporal turnover promote regional coexistence in tropics but not temperate zone. Ecology 96:264–273.

Fronhofer, E. A., J. Klecka, C. J. Melián, and F. Altermatt. 2015. Condition-dependent movement and dispersal in experimental metacommunities. (S. Navarrete, ed.)Ecology Letters 18:954–963.

Fronhofer, E. A., D. Legrand, F. Altermatt, A. Ansart, S. Blanchet, D. Bonte, A. Chaine, et al. 2018. Bottom-up and top-down control of dispersal across major organismal groups. Nature Ecology and Evolution 2:1859–1863.

Fukami, T. 2004. Community assembly along a species pool gradient: Implications for multiple-scale patterns of species diversity. Population Ecology 46:137–147.

———. 2015. Historical Contingency in Community Assembly: Integrating Niches, Species Pools, and Priority Effects. Annual Review of Ecology, Evolution, and Systematics 46:1–23.

Gálvez, Á., P. R. Peres-Neto, A. Castillo-Escrivà, F. Bonilla, A. Camacho, E. M. García-Roger, S. Iepure, et al. 2022. Inconsistent response of taxonomic groups to space and environment in mediterranean and tropical pond metacommunities. Ecology 104:1–16.

Ghalambor, C. K. 2006. Are mountain passes higher in the tropics? janzen's hypothesis revisited. Integrative and Comparative Biology 46:5–17.

Gómez-Rodríguez, C., and A. Baselga. 2018. Variation among European beetle taxa in patterns

of distance decay of similarity suggests a major role of dispersal processes. Ecography 41:1825–1834.

Goodnight, C. J. 2011. Evolution in metacommunities. Philosophical Transactions of the Royal Society B: Biological Sciences 366:1401–1409.

Graham, M. H., and P. K. Dayton. 2002. On the Evolution of Ecological Ideas : Paradigms and Scientific Progress. Ecological Society of America 83:1481–1489.

Grainger, T. N., J. M. Levine, and B. Gilbert. 2019. The Invasion Criterion: A Common Currency for Ecological Research. Trends in Ecology and Evolution 34:925–935.

Gravel, D., C. D. Canham, and C. Messier. 2006. Reconciling niche and neutrality : the continuum hypothesis. Ecology 399–409.

Guzman, L. M., R. M. Germain, C. Forbes, S. Straus, M. I. O'Connor, D. Gravel, D. S. Srivastava, et al. 2019. Towards a multi-trophic extension of metacommunity ecology. Ecology Letters 22:19–33.

Guzman, L. M., P. L. Thompson, D. S. Viana, B. Vanschoenwinkel, Z. Horváth, R. Ptacnik, A. Jeliazkov, et al. 2022. Accounting for temporal change in multiple biodiversity patterns improves the inference of metacommunity processes. Ecology 1–16.

Henriques-Silva, R., F. Boivin, V. Calcagno, M. C. Urban, and P. R. Peres-Neto. 2015. On the evolution of dispersal via heterogeneity in spatial connectivity. Proceedings of the Royal Society B: Biological Sciences 282:20142879.

Highland, S. A., J. C. Miller, and J. A. Jones. 2013. Determinants of moth diversity and community in a temperate mountain landscape: Vegetation, topography, and seasonality. Ecosphere 4.

Hill, M. J., J. C. White, J. Biggs, R. A. Briers, D. Gledhill, M. E. Ledger, I. Thornhill, et al. 2021. Local contributions to beta diversity in urban pond networks: Implications for biodiversity conservation and management. Diversity and Distributions 27:887–900.

Hillerislambers, J., P. B. Adler, W. S. Harpole, J. M. Levine, and M. M. Mayfield. 2012. Rethinking Community Assembly through the Lens of Coexistence Theory. Annual Review of Ecology, Evolution, and Systematics.

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

Hitchman, S. M., M. E. Mather, J. M. Smith, and J. S. Fencl. 2018. Identifying keystone habitats with a mosaic approach can improve biodiversity conservation in disturbed ecosystems. Global Change Biology 24:308–321.

Holyoak, M., T. Caspi, and L. W. Redosh. 2020. Integrating Disturbance, Seasonality, Multi-Year Temporal Dynamics, and Dormancy Into the Dynamics and Conservation of Metacommunities. Frontiers in Ecology and Evolution 8:1–17.

Houlahan, J. E., S. T. Mckinney, T. M. Anderson, and B. J. Mcgill. 2016. The priority of

prediction in ecological understanding. Oikos.

Hua, X. 2016. The impact of seasonality on niche breadth , distribution range and species richness : a theoretical exploration of Janzen ' s hypothesis. Proc. R. Soc. B.

Hughes, C. E. 2017. Are there many different routes to becoming a global biodiversity hotspot ? Proceedings of the National Academy of Sciences 114:4275–4277.

Hutchinson, G. . 1959. Homage to Santa Rosalia or Why are There So Many Kinds of Animals?

Ishwaran, H., and M. Lu. 2019. Standard errors and confidence intervals for variable importance in random forest regression, classification, and survival. Statistics in Medicine 38:558–582.

Jabot, F., F. Laroche, F. Massol, F. Arthaud, J. Crabot, M. Dubart, S. Blanchet, et al. 2020. Assessing metacommunity processes through signatures in spatiotemporal turnover of community composition. Ecology Letters.

Jankowski, J. E., A. L. Ciecka, N. Y. Meyer, and K. N. Rabenold. 2009. Beta diversity along environmental gradients: implications of habitat specialization in tropical montane landscapes. Journal of Animal Ecology 78:315–327.

Janzen, D. H. 1967. Why Mountain Passes are Higher in the Tropics. The American Naturalist 101:233–249.

Jocque, M., R. Field, L. Brendonck, and L. De Meester. 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.

Justus, J. 2021. The Philosophy of Ecology: An Introduction. Cambridge Introductions to Philosophy and Biology. Cambridge University Press.

Karger, D. N., O. Conrad, J. Böhner, T. Kawohl, H. Kreft, R. W. Soria-Auza, N. E. Zimmermann, et al. 2017. Climatologies at high resolution for the earth's land surface areas. Scientific Data 4:170122.

Karger, D. N., H. Tuomisto, V. B. Amoroso, D. Darnaedi, A. Hidayat, S. Abrahamczyk, J. Kluge, et al. 2015. The importance of species pool size for community composition. Ecography 1243–1253.

Keddy, P. A. 1992. Assembly and response rules: two goals for predictive community ecology. Journal of Vegetation Science 3:157–164.

Khattar, G., M. Macedo, R. Monteiro, and P. Peres-Neto. 2021. Determinism and stochasticity in the spatial–temporal continuum of ecological communities: the case of tropical mountains. Ecography 44:1391–1402.

Khattar, G., S. Vaz, P. Henrique, P. Braga, M. Macedo, and L. Felipe. 2022. Life history traits modulate the influence of environmental stressors on biodiversity : The case of fireflies , climate and artificial light at night 1–12.

Koffel, T., K. Umemura, E. Litchman, and C. A. Klausmeier. 2022. A general framework for species-abundance distributions: Linking traits and dispersal to explain commonness and rarity. Ecology Letters.

Kolasa, J., M. P. Hammond, and J. Yan. 2021. Metacommunity research can benefit from including context-dependency. bioRxiv 2021.09.26.461405.

Körner, C. 2000. Why are there global gradients in species richness? Mountains might hold the answer [1]. Trends in Ecology and Evolution 15:513–514.

Körner, C., and J. Paulsen. 2004. A world-wide study of high altitude treeline temperatures. Journal of Biogeography 713–732.

Kozak, K. H., and J. J. Wiens. 2010. Niche Conservatism Drives Elevational Diversity Patterns in Appalachian Salamanders. The American Naturalist 176:40–54.

Kraft, N. J. B. B., L. S. Comita, J. M. Chase, N. J. Sanders, N. G. Swenson, T. O. Crist, J. C. Stegen, et al. 2011. Disentangling the Drivers of β Diversity Along Latitudinal and Elevational Gradients. Science 333:1755–1758.

Kubisch, A., R. D. Holt, H.-J. Poethke, and E. A. Fronhofer. 2014. Where am I and why? Synthesizing range biology and the eco-evolutionary dynamics of dispersal. Oikos 123:5–22.

Kursa, M. B., and W. R. Rudnicki. 2010. Feature selection with the boruta package. Journal of Statistical Software 36:1–13.

Lamy, T., N. I. Wisnoski, R. Andrade, M. C. N. Castorani, A. Compagnoni, N. Lany, L. Marazzi, et al. 2021. The dual nature of metacommunity variability. Oikos 130:2078–2092.

Langevelde, F. Van, J. A. Ettema, M. Donners, M. F. Wallisdevries, and D. Groenendijk. 2011. Effect of spectral composition of artificial light on the attraction of moths. Biological Conservation 144:2274–2281.

Lawton, J. H. 1999. Are There General Laws in Ecology? Oikos 84:177–192.

Layeghifard, M., V. Makarenkov, and P. R. Peres-Neto. 2015. Spatial and species compositional networks for inferring connectivity patterns in ecological communities. Global Ecology and Biogeography 24:718–727.

Lees, D. C., and A. Zilli. 2019. Moths: A complete Guide to Biology and Behavior. Natural History Museum: London, UK.

Lefcheck, J. S. 2020. PiecewiseSEM (R package version 2.1.2).

Legendre, P. 2014. Interpreting the replacement and richness difference components of beta diversity. Global Ecology and Biogeography 23:1324–1334.

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

Legendre, P., M. De Cáceres, and D. Borcard. 2010. Community surveys through space and time: Testing the space-time interaction in the absence of replication. Ecology 91:262–272.

Legendre, P., and O. Gauthier. 2014. Statistical methods for temporal and space–time analysis of community composition data. Proceedings of the Royal Society B: Biological Sciences 281:20132728.

Leibold, M. A., and J. M. Chase. 2018. Metacommunity ecology, volume 59. Princeton

University Press.

Leibold, M. A., L. Govaert, N. Loeuille, L. De Meester, and M. C. Urban. 2022. Evolution and Community Assembly Across Spatial Scales. Annual Review of Ecology, Evolution, and Systematics 53:299–326.

Leibold, M. A., M. Holyoak, N. Mouquet, P. Amarasekare, J. M. Chase, M. F. Hoopes, R. D. Holt, et al. 2004. The metacommunity concept: a framework for multi-scale community ecology. Ecology Letters 7:601–613.

Leibold, M. A., F. J. Rudolph, F. G. Blanchet, L. De Meester, D. Gravel, F. Hartig, P. Peres-Neto, et al. 2021. The internal structure of metacommunities. Oikos 1–13.

Leibold, M. A., M. C. Urban, L. De Meester, C. A. Klausmeier, and J. Vanoverbeke. 2019. Regional neutrality evolves through local adaptive niche evolution. Proceedings of the National Academy of Sciences of the United States of America 116:2612–2617.

Lessard, J.-P. P., B. G. Weinstein, M. K. Borregaard, K. A. Marske, D. R. Martin, J. A. McGuire, J. L. Parra, et al. 2016. Process-Based Species Pools Reveal the Hidden Signature of Biotic Interactions Amid the Influence of Temperature Filtering. The American Naturalist 187:75–88.

Lessard, J. P., J. Belmaker, J. A. Myers, J. M. Chase, and C. Rahbek. 2012*a*. Inferring local ecological processes amid species pool influences. Trends in Ecology and Evolution 27:600–607.

Lessard, J. P., M. K. Borregaard, J. A. Fordyce, C. Rahbek, M. D. Weiser, R. R. Dunn, and N. J. Sanders. 2012*b*. Strong influence of regional species pools on continent-wide structuring of local communities. Proceedings of the Royal Society B: Biological Sciences 279:266–274.

Lester, S. E., B. I. Ruttenberg, S. D. Gaines, and B. P. Kinlan. 2007. The relationship between dispersal ability and geographic range size. Ecology Letters 10:745–758.

Levin, S. A. 1992. The problem of pattern and scale in ecology. Ecology 73:1943–1967.

Levins, R., and R. Lewontin. 1980. Dialectics and reductionism in ecology. Synthese 43:47–78.

Li, H., M. Holyoak, and Z. Xiao. 2023. Disentangling spatiotemporal dynamics in metacommunities through a species-patch network approach. Ecology Letters.

Li, X., Y. Zhou, M. Zhao, and X. Zhao. 2020. A harmonized global nighttime light dataset 1992–2018. Scientific Data 7:1–9.

Liaw, A., and M. Wiener. 2002. Classification and Regression by randomForest. R News 2:18–22.

Liu, W., H. Sirisena, and K. Pawlikowski. 2009. Weighted algebraic connectivity metric for non-uniform traffic in reliable network design. Pages 1–6 *in* 2009 International Conference on Ultra Modern Telecommunications & Workshops. IEEE.

MacArthur, R. H. 1972. Geographical ecology: patterns in the distribution of species. Princeton University Press.

MacArthur, R., and E. O. Wilson. 1967. The theory of island biogeography. Princinton Press. Princeton University Press.

Magurran, A. E., and P. A. Henderson. 2003. Explaining the excess of rare species in natural species abundance distributions. Nature 422:714–716.

Maicher, V., S. Sáfián, M. Murkwe, S. Delabye, Ł. Przybyłowicz, and P. Potocký. 2019. Seasonal shifts of biodiversity patterns and species' elevation ranges of butterflies and moths along a complete rainforest elevational gradient on Mount Cameroon. Dryad.

Manion, G., M. Lisk, S. Ferrier, D. Nieto-Lugilde, K. Mokany, and M. C. Fitzpatrick. 2018. gdm: Generalized Dissimilarity Modeling. R package ver. 1.3.11.

Marco Palamara, G., A. Rozenfeld, C. N. de Santana, J. Klecka, R. Riera, V. M. Eguíluz, and C. J. Melián. 2023. Biodiversity dynamics in landscapes with fluctuating connectivity. Ecography 1–12.

Marquet, P. A., A. P. Allen, J. H. Brown, J. A. Dunne, B. J. Enquist, J. F. Gillooly, P. A. Gowaty, et al. 2014. On theory in ecology. BioScience 64:701–710.

Matthysen, E. 2005. Density-dependent dispersal in birds and mammals. Ecography 28:403–416.

McCain, C. M. 2009. Vertebrate range sizes indicate that mountains may be 'higher' in the tropics. Ecology Letters 12:550–560.

Mcintosh, R. P. 1987. Pluralism in Ecology. Annual Review of Ecology and Systematics 18:321–341.

McPeek, M. A., and R. D. Holt. 1992. The evolution of dispersal in spatially and temporally varying environments. American Naturalist 140:1010–1027.

Miller, J., and J. A. Jones. 2005. Spatial and temporal distribution and abundance of moths in the Andrews Experimental Forest, 1994 to 2008. H. J. Andrews Experimental Forest. Forest Science Data Bank, Corvallis. http://andlter.forestry.oregonstate.edu/data/abstract.aspx?dbcode=SA015.

Mittelbach, G. G., and D. W. Schemske. 2015. Ecological and evolutionary perspectives on community assembly. Trends in Ecology & Evolution 30:241–247.

Moe, S. J., R. S. Stelzer, M. R. Forman, W. S. Harploe, T. Daufresne, and T. Yoshida. 2005. Recent advances in ecological stoichiometry: Insights for population and community ecology. Oikos 109:29–39.

Moritz, C., C. N. Meynard, V. Devictor, K. Guizien, C. Labrune, J. Guarini, and N. Mouquet. 2013. Disentangling the role of connectivity , environmental fi ltering , and spatial structure on metacommunity dynamics. Oikos 1401–1410.

Mouquet, N., D. Gravel, F. Massol, and V. Calcagno. 2013. Extending the concept of keystone species to communities and ecosystems. Ecology Letters 16:1–8.

Mouquet, N., and M. Loreau. 2003. Community Patterns in Source-Sink Metacommunities. The American naturalist.

Myers, J. A., J. M. Chase, I. Jiménez, P. M. Jørgensen, A. Araujo-Murakami, N. Paniagua-

Zambrana, and R. Seidel. 2013. Beta-diversity in temperate and tropical forests reflects dissimilar mechanisms of community assembly. Ecology Letters 16:151–157.

Myers, N., R. Milttermeier, C. Milttermeier, G. A. B. Fonseca, and J. Kent. 2000. Biodiversity hotspots for conservation priorities. Nature 115:853–862.

Nathan, R., Getz, W. M., Revilla, E., Holyoak, M., Kadmon, R., Saltz, D., & Smouse, P. E. (2008). A movement ecology paradigm for unifying organismal movement research. Proceedings of the National Academy of Sciences, 105(49), 19052-19059.

Nishizawa, K., N. Shinohara, M. W. Cadotte, and A. S. Mori. 2022. The latitudinal gradient in plant community assembly processes: A meta-analysis. (J. Chase, ed.)Ecology Letters 25:1711–1724.

O'Hara, R. B. 2005. The anarchist's guide to ecological theory. Or, we don't need no stinkin' laws. Oikos 110:390–393.

O'Neill, R. V., C. T. Hunsaker, S. P. Timmins, B. L. Jackson, K. B. Jones, K. H. Riitters, and J. D. Wickham. 1996. Scale problems in reporting landscape pattern at the regional scale. Landscape Ecology 11:169–180.

O'Sullivan, J. D., J. Christopher, R. Wilson, and A. G. Rossberg. 2023. Community composition exceeds area as a predictor of long-term conservation value. PLoS Computational Biology 19:1–15.

Oksanen, J., F. G. Blanchet, M. Friendly, R. Kindt, P. Legendre, D. McGlinn, P. R. Minchin, et al. 2020. vegan: Community Ecology Package.

Olden, J. D., and N. L. R. Poff. 2003. Toward a Mechanistic Understanding and Prediction of Biotic Homogenization. American Naturalist 162:442–460.

Orrock, J. L., and J. I. Watling. 2010. Local Community size mediates ecological drift and competition in metacommunities. Proceedings of the Royal Society B: Biological Sciences 277:2185–2191.

Ovaskainen, O., J. Rybicki, and N. Abrego. 2019. What can observational data reveal about metacommunity processes? Ecography 42:1877–1886.

Partel, M., M. Zobel, K. Zobel, and E. Van Der Maarel. 1996. The Species Pool and Its Relation to Species Richness : Evidence from Estonian Plant Communities. Oikos 75:111–117.

Peres-Neto, P.-R., M. Leibold, and S. Dray. 2012a. Assessing the effects of spatial contingency and environmental filtering on metacommunity phylogenetics. Ecology 93:14–30.

Peres-Neto, P. R., and P. Legendre. 2010. Estimating and controlling for spatial structure in the study of ecological communities. Global Ecology and Biogeography 19:174–184.

Peres-Neto, P. R., P. Legendre, S. Dray, and D. Borcard. 2006. Variation partitioning of species data matrices: estimation and comparison of fractions. Ecology 87:2614–25.

Peres-Neto, P. R., M. Leibold, and S. Dray. 2012b. Assessing the effects of spatial contingency and environmental filtering on metacommunity phylogenetics. Ecology 93:14–30.

Peres-Neto, P. R., J. D. Olden, and D. a Jackson. 2001. Environmentally constrained null models : site suitability as occupancy criterion. Oikos 93:110–120.

Perez, M., T. J. Morris, K. Cottenie, and A. N. Schwalb. 2023. Limitations of beta diversity in conservation site selection. Ecological Indicators 154:110732.

Pilotto, F., I. Kühn, R. Adrian, R. Alber, A. Alignier, C. Andrews, J. Bäck, et al. 2020. Meta-analysis of multidecadal biodiversity trends in Europe. Nature Communications 11.

Pinheiro, J., D. Bates, S. DebRoy, D. Sarkar, and R Core Team. 2019. nlme: Linear and Nonlinear Mixed Effects Models. R-package ver 3.1-140.

Poethke, H. J., A. Kubisch, O. Mitesser, and T. Hovestadt. 2016. The evolution of density-dependent dispersal under limited information. Ecological Modelling 338:1–10.

Presley, S. J., L. M. Cisneros, B. D. Patterson, and M. R. Willig. 2012. Vertebrate metacommunity structure along an extensive elevational gradient in the tropics: a comparison of bats, rodents and birds. Global Ecology and Biogeography 21:968–976.

Preston, F. W. 1960. Time and Space and the Variation of Species. Ecology 41:611–627.

Qian, H., and R. E. Ricklefs. 2012. Disentangling the effects of geographic distance and environmental dissimilarity on global patterns of species turnover. Global Ecology and Biogeography 21:341–351.

R Core Team. 2023. R: A Language and Environment for Statistical Computing. Vienna, Austria.

Rahbek, C., M. K. Borregaard, A. Antonelli, R. K. Colwell, B. G. Holt, D. Nogues-bravo, C. M. Ø. Rasmussen, et al. 2019. Building mountain biodiversity: Geological and evolutionary processes 1119:1114–1119.

Record, S., N. M. Voelker, P. L. Zarnetske, N. I. Wisnoski, J. D. Tonkin, C. Swan, L. Marazzi, et al. 2021. Novel Insights to Be Gained From Applying Metacommunity Theory to Long-Term, Spatially Replicated Biodiversity Data. Frontiers in Ecology and Evolution 8.

Resetarits, E. J., S. E. Cathey, and M. A. Leibold. 2018. Testing the keystone community concept: effects of landscape, patch removal, and environment on metacommunity structure. Ecology 99:57–67.

Ricklefs, R. E. 2008. Disintegration of the ecological community. The American naturalist 172:741–750.

Ron, R., O. Fragman-Sapir, and R. Kadmon. 2018. Dispersal increases ecological selection by increasing effective community size. Proceedings of the National Academy of Sciences of the United States of America 115:11280–11285.

Roswell, M., J. Dushoff, and R. Winfree. 2021. A conceptual guide to measuring species diversity. Oikos 130:321–338.

Roughgarden, J. 2009. Is there a general theory of community ecology? Biology and Philosophy 24:521–529.

Ruhí, A., T. Datry, and J. L. Sabo. 2017a. Interpreting beta-diversity components over time to conserve metacommunities in highly dynamic ecosystems. Conservation Biology 31:1459–1468.

———. 2017b. Interpreting beta-diversity components over time to conserve metacommunities in highly dynamic ecosystems. Conservation Biology 31:1459–1468.

Sheard, C., M. H. C. Neate-Clegg, N. Alioravainen, S. E. I. Jones, C. Vincent, H. E. A. MacGregor, T. P. Bregman, et al. 2020. Ecological drivers of global gradients in avian dispersal inferred from wing morphology. Nature Communications 11.

Sheldon, K. S., R. B. Huey, M. Kaspari, and N. J. Sanders. 2018. Fifty Years of Mountain Passes: A Perspective on Dan Janzen's Classic Article. The American Naturalist 191:553–565.

Sheldon, K. S., and J. J. Tewksbury. 2014. The impact of seasonality in temperature on thermal tolerance and elevational range size. Ecology 95:2134–2143.

Shoemaker, L. G., L. M. Hallett, L. Zhao, D. C. Reuman, S. Wang, K. L. Cottingham, R. J. Hobbs, et al. 2022. The long and the short of it: Mechanisms of synchronous and compensatory dynamics across temporal scales. Ecology 1–14.

Shoemaker, L. G., and B. A. Melbourne. 2016. Linking metacommunity paradigms to spatial coexistence Mechanisms. Ecology 97:2436–2446.

Shoemaker, L. G., L. L. Sullivan, I. Donohue, J. S. Cabral, R. J. Williams, M. M. Mayfield, J. M. Chase, et al. 2020a. Integrating the underlying structure of stochasticity into community ecology. Ecology 101:1–17.

Shoemaker, L. G., L. L. Sullivan, I. Donohue, J. S. Cabral, R. J. Williams, M. M. Mayfield, J. M. Chase, et al. 2020b. Integrating the underlying structure of stochasticity into community ecology. Ecology 101:1–17.

Simberloff, D. 2004. Community Ecology: Is It Time to Move On? The American Naturalist 163:787–799.

Siqueira, T., V. S. Saito, L. M. Bini, A. S. Melo, D. K. Petsch, V. L. Landeiro, K. T. Tolonen, et al. 2020. Community size can affect the signals of ecological drift and niche selection on biodiversity. Ecology 101:0–3.

Snell Taylor, S., J. R. Coyle, E. P. White, and A. H. Hurlbert. 2020. A simulation study of the use of temporal occupancy for identifying core and transient species. PLOS ONE 15:e0241198.

Snell Taylor, S. J., B. S. Evans, E. P. White, and A. H. Hurlbert. 2018. The prevalence and impact of transient species in ecological communities. Ecology 99:1825–1835.

Soininen, J. 2014. A quantitative analysis of species sorting across organisms and ecosystems. Ecology 95:3284–3292.

Sokol, E. R., J. E. Barrett, T. J. Kohler, D. M. Mcknight, M. R. Salvatore, and L. F. Stanish. 2020. Evaluating Alternative Metacommunity Hypotheses for Diatoms in the McMurdo Dry Valleys Using Simulations and Remote Sensing Data 8:1–18.

Spalding, V. M. 1909. Distribution and movements of desert plants. Carnegie institution of Washington.

Stegen, J. C., A. L. Freestone, T. O. Crist, M. J. Anderson, J. M. Chase, L. S. Comita, H. V Cornell, et al. 2013. Stochastic and deterministic drivers of spatial and temporal turnover in breeding bird communities. (K. Evans, ed.)Global Ecology and Biogeography 22:202–212.

Stevens, G. C. 1989. The Latitudinal Gradient in Geographical Range : How so Many Species Coexist in the Tropics. American Naturalist 133:240–256.

Sullivan, A. J. D. O., J. C. D. Terry, R. Wilson, and A. G. Rossberg. 2021. Predictors of the long-term conservation value of sites in a mechanistic metacommunity model.

Sunday, J. M., A. E. Bates, and N. K. Dulvy. 2011. Global analysis of thermal tolerance and latitude in ectotherms. Proceeding of the Royal Society of London B: Biological Science 278:1823–1830.

Suzuki, Y., and E. P. Economo. 2021. From species sorting to mass effects: spatial network structure mediates the shift between metacommunity archetypes. Ecography 44:715–726.

Tatsumi, S., R. Iritani, and M. W. Cadotte. 2021. Temporal changes in spatial variation: partitioning the extinction and colonisation components of beta diversity. (M. Anderson, ed.)Ecology Letters 24:1063–1072.

———. 2022. Partitioning the temporal changes in abundance-based beta diversity into loss and gain components. Methods in Ecology and Evolution 13:2042–2048.

Taylor, D. R., L. W. Aarssen, C. Loehle, and C. On. 1990. On the Relationship between r/K Selection and Environmental Carrying Capacity: A New Habitat Templet for Plant Life History Strategies. Oikos 58:239.

Thioulouse, J., S. Dray, A.--B. Dufour, A. Siberchicot, T. Jombart, and S. Pavoine. 2018. Multivariate Analysis of Ecological Data with {ade4}. Springer.

Thompson, P. L., and E. A. Fronhofer. 2019. The conflict between adaptation and dispersal for maintaining biodiversity in changing environments. Proceedings of the National Academy of Sciences of the United States of America 116:21061–21067.

Thompson, P. L., L. M. Guzman, L. De Meester, Z. Horváth, R. Ptacnik, B. Vanschoenwinkel, D. S. Viana, et al. 2020. A process-based metacommunity framework linking local and regional scale community ecology. Ecology Letters 23:1314–1329.

Tonkin, J. D., M. T. Bogan, N. Bonada, B. Rios-Touma, and D. A. Lytle. 2017. Seasonality and predictability shape temporal species diversity. Ecology 98:1201–1216.

Tscharntke, T., J. M. Tylianakis, T. A. Rand, R. K. Didham, L. Fahrig, P. Batáry, J. Bengtsson, et al. 2012. Landscape moderation of biodiversity patterns and processes - eight hypotheses. Biological Reviews 87:661–685.

Tucker, C. M., L. G. Shoemaker, K. F. Davies, D. R. Nemergut, and B. A. Melbourne. 2016. Differentiating between niche and neutral assembly in metacommunities using null models of β-diversity. Oikos 125:778–789.

Umaña, M. N., C. Zhang, M. Cao, L. Lin, and N. G. Swenson. 2017. A core-transient framework for trait-based community ecology: an example from a tropical tree seedling community.

Ecology Letters.

Urban, M. C., G. Bocedi, A. P. Hendry, J. B. Mihoub, G. Pe'er, A. Singer, J. R. Bridle, et al. 2016. Improving the forecast for biodiversity under climate change. Science 353.

URBAN, M., M. LEIBOLD, P. AMARASEKARE, L. DEMEESTER, R. GOMULKIEWICZ, M. HOCHBERG, C. KLAUSMEIER, et al. 2008. The evolutionary ecology of metacommunities. Trends in Ecology & Evolution 23:311–317.

Van Allen, B. G., N. L. Rasmussen, C. J. Dibble, P. A. Clay, and V. H. W. Rudolf. 2017. Top predators determine how biodiversity is partitioned across time and space. (J. Chase, ed.)Ecology Letters.

Van Der Kooi, C. J., D. G. Stavenga, K. Arikawa, G. Belušič, and A. Kelber. 2021. Evolution of Insect Color Vision: From Spectral Sensitivity to Visual Ecology. Annual Review of Entomology 66:435–461.

van der Plas, F. 2019. Biodiversity and ecosystem functioning in naturally assembled communities. Biological Reviews 94:1220–1245.

Vaz, S., S. Manes, D. Gama-Maia, L. Silveira, G. Mattos, P. C. Paiva, M. Figueiredo, et al. 2021. Light pollution is the fastest growing potential threat to firefly conservation in the Atlantic Forest hotspot. Insect Conservation and Diversity 14:211–224.

Vellend, M. 2010. Conceptual Synthesis in Community Ecology. The Quarterly Review of Biology 85:183–206.

———. 2018. The Theory of Ecological Communities. (S. A. Levin & H. S. Horn, eds.)Monographs in Population Biology (Vol. 57). Princeton University Press.

Vellend, M., D. S. Srivastava, K. M. Anderson, C. D. Brown, J. E. Jankowski, E. J. Kleynhans, N. J. B. Kraft, et al. 2014. Assessing the relative importance of neutral stochasticity in ecological communities. Oikos 123:1420–1430.

Venevskaia, I., S. Venevsky, and C. D. Thomas. 2013. Projected latitudinal and regional changes in vascular plant diversity through climate change: short-term gains and longer-term losses. Biodiversity and Conservation 22:1467–1483.

Viana, D. S., and J. M. Chase. 2019. Spatial scale modulates the inference of metacommunity assembly processes. Ecology 100:1–9.

———. 2022. Increasing climatic decoupling of bird abundances and distributions. Nature Ecology and Evolution 6:1299–1306.

Ward, E. J., J. E. Jannot, Y. W. Lee, K. Ono, A. O. Shelton, and J. T. Thorson. 2015. Using spatiotemporal species distribution models to identify temporally evolving hotspots of species co-occurrence. Ecological Applications 25:2198–2209.

Wardhaugh, C. W., M. J. Stone, and N. E. Stork. 2018. Seasonal variation in a diverse beetle assemblage along two elevational gradients in the Australian Wet Tropics. Scientific Reports 8:8559.

Wardle, D. A., O. Zackrisson, G. Hörnberg, and C. Gallet. 1997. The influence of island area on

ecosystem properties. Science 277:1296–1299.

Webb, C. O., D. D. Ackerly, M. a. McPeek, and M. J. Donoghue. 2002. Phylogenies and Community Ecology. Annual Review of Ecology and Systematics 33:475–505.

Weiher, E., D. Freund, T. Bunton, A. Stefanski, T. Lee, and S. Bentivenga. 2011. Advances, challenges and a developing synthesis of ecological community assembly theory. Philosophical Transactions of the Royal Society B: Biological Sciences 366:2403–2413.

White, E. P., S. K. Morgan Ernest, P. B. Adler, A. H. Hurlbert, and S. Kathleen Lyons. 2010. Integrating spatial and temporal approaches to understanding species richness. Philosophical Transactions of the Royal Society B: Biological Sciences.

White, J. W., A. Rassweiler, J. F. Samhouri, A. C. Stier, and C. White. 2014. Ecologists should not use statistical significance tests to interpret simulation model results. Oikos 123:385–388.

Whittaker, R. J., M. B. Araújo, P. Jepson, R. J. Ladle, J. E. M. Watson, and K. J. Willis. 2005. Conservation Biogeography: assessment and prospect. Diversity and Distributions 11:3–23.

Wiens, J. a. 1989. Spatial Scaling in Ecology. Functional Ecology 3:385–397.

Willig, M. R., and S. J. Presley. 2016. Biodiversity and metacommunity structure of animals along altitudinal gradients in tropical montane forests. Journal of Tropical Ecology 32:421–436.

Winegardner, A. K., B. K. Jones, I. S. Y. Ng, T. Siqueira, and K. Cottenie. 2012. The terminology of metacommunity ecology. Trends in Ecology and Evolution 27:253–254.

Wisnoski, N. I., M. A. Leibold, and J. T. Lennon. 2019. Dormancy in metacommunities. American Naturalist 194:135–151.

Wisnoski, N. I., and L. G. Shoemaker. 2022. Seed banks alter metacommunity diversity: The interactive effects of competition, dispersal and dormancy. Ecology Letters 25:740–753.

Yang, X., J. Tan, K. H. Sun, and L. Jiang. 2020. Experimental demonstration of the importance of keystone communities for maintaining metacommunity biodiversity and ecosystem functioning. Oecologia 193:437–447.

Zhang, H., D. Bearup, I. Nijs, S. Wang, G. Barabás, Y. Tao, and J. Liao. 2021. Dispersal network heterogeneity promotes species coexistence in hierarchical competitive communities. Ecology Letters 24:50–59.

Zobel, M. 2016. The species pool concept as a framework for studying patterns of plant diversity. Journal of Vegetation Science 27:8–18.

Zobel, M., R. Otto, L. Laanisto, A. Naranjo-Cigala, M. Pärtel, and J. M. Fernández-Palacios. 2011. The formation of species pools: Historical habitat abundance affects current local diversity. Global Ecology and Biogeography 20:251–259.

Zuloaga, J., and J. T. Kerr. 2017. Over the top: do thermal barriers along elevation gradients limit biotic similarity? Ecography 40:478–486.