

Washington University in St. Louis

Washington University Open Scholarship

Arts & Sciences Electronic Theses and
Dissertations

Arts & Sciences

7-28-2023

Biotic Interactions and the Maintenance of Biodiversity across a Tropical Elevational Gradient

David Henderson

Washington University in St. Louis

Follow this and additional works at: https://openscholarship.wustl.edu/art_sci_etds

Recommended Citation

Henderson, David, "Biotic Interactions and the Maintenance of Biodiversity across a Tropical Elevational Gradient" (2023). *Arts & Sciences Electronic Theses and Dissertations*. 3000.

https://openscholarship.wustl.edu/art_sci_etds/3000

This Dissertation is brought to you for free and open access by the Arts & Sciences at Washington University Open Scholarship. It has been accepted for inclusion in Arts & Sciences Electronic Theses and Dissertations by an authorized administrator of Washington University Open Scholarship. For more information, please contact digital@wumail.wustl.edu.

WASHINGTON UNIVERSITY IN ST. LOUIS

Division of Biology and Biomedical Sciences
Evolution, Ecology and Population Biology

Dissertation Examination Committee:

Dr. Jonathan A. Myers, Chair

Dr. Rachel Penczykowski

Dr. Brian E. Sedio

Dr. Adam B. Smith

Dr. J. Sebastian Tello

Biotic Interactions and the Maintenance of Biodiversity Across a Tropical Elevational Gradient

by

David Henderson

A dissertation presented to
Washington University in St. Louis
in partial fulfillment of the
requirements for the degree
of Doctor of Philosophy

August 2023
St. Louis, Missouri

© 2023, David Henderson

Table of Contents

List of Figures	iv
List of Tables	v
Acknowledgments.....	vi
Abstract.....	vii
Chapter 1: Introduction.....	1
1.1 Background.....	1
1.2 Chapter Overview	6
1.3 Large Scale Diversity Patterns: Elevational Gradients.....	7
1.4 Introduction References.....	9
Chapter 2: Disentangling Determinism and Stochasticity in Local Tree Neighborhoods Across a Tropical Elevational Gradient.....	14
2.1 Introduction.....	14
2.2 Methods.....	17
2.3 Results.....	23
2.4 Discussion	25
2.5 References.....	31
2.6 Figures.....	36
Chapter 3: An Ecological Metabolomics Window on Plant Chemistry and Diversity Gradients: Chemical Dissimilarity within Tree Communities Increases with Species Diversity and Decreases with Elevation in the Central Andes	40
3.1 Introduction.....	40
3.2 Methods.....	44
3.3 Results.....	49
3.4 Discussion	51
3.5 References.....	59
3.6 Tables.....	63
3.7 Figures.....	65
Chapter 4: Functional Trait Patterns Reveal Differential Mechanisms of Tree Community Assembly Along a Tropical Elevational Gradient	70
4.1 Introduction.....	70

4.2 Methods.....	75
4.3 Results.....	80
4.4 Discussion	81
4.5 References.....	86
4.6 Figures.....	91
Chapter 5: Conclusion.....	94
5.1 Conclusion	94
5.2 Conclusion References.....	96
Appendix A:.....	98
Appendix B:	112
Appendix C:.....	114
Appendix D: R code.....	120

List of Figures

Figure 2.1: Stem mapped plots, Calculation of neighborhood dissimilarity	36
Figure 2.2: Mean taxonomic dissimilarity of tree neighborhoods	37
Figure 2.3: Mean functional dissimilarity of tree neighborhoods	38
Figure 3.1: Chemical dendrogram	65
Figure 3.2: Climate PCA biplot	66
Figure 3.3: Variation in median chemical dissimilarity	67
Figure 3.4: Variation in chemical phylogenetic signal	68
Figure 4.1: Map of Bolivian Andes	91
Figure 4.2: SES relationships of functional dispersion of morphological	92
Figure 4.3: SES relationships of functional dispersion of morphological traits	93

List of Tables

Table 3.1: Plot data.....	63
Table 3.2: Phylogenetic Signal of Metabolites.....	64

Acknowledgments

It would be difficult for me to express my gratitude to all of the people who have helped me throughout this experience.

There was Sebastian Moreno and Cristian, who acted as gracious hosts and guides. There was Gary and Barbra who did the same and additionally let us use part of their personal property as a field site.

My committee. Dr. Rachel Penczykowski for being a constant friend and role model during my years here. Dr. Adam Smith for his gracious support and for designing a project for me that opened up new worlds. Dr. Brian Sedio for introducing me to metabolomics, for his advice in analysis and writing, and for a great visit in the lab in Texas. And of course, my advisors Dr. Jonathan Myers and Dr. Sebastian Tello, for their unwavering support and for giving me enough runway to finish with a chance of landing safely.

Professor Stanton Braude was not only a mentor but a friend who provided guidance through difficult times.

My friends and labmates who have walked much of this journey with me. This is especially true of my cohort, with whom I shared so much of these last years.

And I thank my family for helping me examine my options and supporting my decision.

David Henderson

Washington University in St. Louis

August 2023

ABSTRACT OF THE DISSERTATION

Biotic Interactions and the Maintenance of Biodiversity across a Tropical Elevational Gradient

by

David Henderson

Doctor of Philosophy in Biology and Biomedical Sciences

Evolution, Ecology and Population Biology

Washington University in St. Louis, 2023

Dr. Jonathan A. Myers, Chair

A major goal of community ecology is to elucidate the processes that create patterns of biodiversity. Specifically, the discipline focuses on the processes that influence community assembly, diversity gradients, and/or species coexistence. Coexistence mechanisms have been traditionally thought of in terms of resource competition, whereby species are able to coexist by having different niche requirements, both biotic and abiotic, that maximize their fitness. Much of the previous work in this arena focused on the abiotic factors that influence species coexistence. However, many of the seminal ideas about the forces that drive abundance and coexistence patterns hypothesized that it was differences in the nature of biotic interactions between those species that were responsible for the observed patterns. However, due to the variability and complexity of biotic interactions, competition was generally the only interaction to be extensively considered. Further research gave evidence to the idea that antagonistic interactions with natural enemies are a possible niche axis that can determine coexistence of the hosts.

The vast number of natural enemies and the even greater number of defensive chemical metabolites provide a probable mechanism of niche differentiation, where a species niche is defined by those enemies that they avoid through their defenses. Advances in analytical chemical metabolomics techniques are used in this dissertation to examine the role that more complex interaction types play in structuring communities. More recent research utilizing these techniques have found evidence of the role that plant chemical defenses play in mediating interactions between plant hosts and their natural enemies. Theoretical and technical advances are combined here to investigate the factors that drive species diversity patterns and that maintain species coexistence in 3 principal ways: 1) by testing the relative importance of competitive interactions among tree species across an elevational-diversity gradient in the tropical Andes (Chapters 2, 3); 2) by testing the relative importance of enemy mediated interactions via chemically-mediated niche differences across the same elevational-diversity gradient (Chapters 3, 4); and 3) by exploring the effects of the biotic and abiotic environments on trait patterns and community assembly processes.

Chapter 2 investigated whether competitive interactions among tropical tree species vary systematically across a large-scale biodiversity gradient. Using tree species data collected from a network of permanent plots in the Bolivian Andes, I quantified the taxonomic and functional differences of neighboring species of individuals of each species within a plot. The results showed that the taxonomic composition of tree neighborhoods becomes more stochastic with increasing species diversity. This suggests that competitive interactions appear to be more unpredictable among neighborhoods of the same species in higher-diversity tree communities. Interestingly, similar patterns were not detected in the functional trait composition of tree

neighborhoods. Taken together, the two patterns suggest that species traits may be influenced by the abiotic environment and that biotic interactions other than competition could be at play.

Chapter 3 investigated how enemy-mediated interactions varied over a gradient of climate, diversity, and elevation using an ecological metabolomics approach. Defensive chemical traits were measured from leaf compounds and used to gauge the strength of enemy mediated interactions within communities. The results showed that chemical dissimilarity among tree species increased with increasing community diversity and toward more benign conditions and that evolution of chemical defenses is more rapid along these gradients. This implies that natural enemies impose a stronger selective pressure on plant chemical defenses in more diverse communities and in more productive climates, and that the defensive chemical compounds that hosts use are less phylogenetically conserved in higher diversity communities. The differential strength of biotic interactions and the greater evolutionary lability of defensive metabolites across communities implies that these enemy mediated interactions likely play a role in the maintenance and origin of biodiversity gradients.

Chapter 4 tested how community assembly is affected by both biotic and the abiotic environments, by examining the relationships between two suites of plant functional traits (chemical and morphological), and abiotic site characteristics (climate and soils). Chemical traits were used to approximate the biotic environment while morphological traits approximated the abiotic. The results showed systematic variation in morphological trait patterns but no similar patterns among the chemical traits. The abiotic environment was found to have a stronger than expected influence on chemical trait patterns, where trees may utilize chemical defenses differently in wet and dry environments. A focus on individual chemical defensive compounds to

gauge their influences on community assembly and dynamics would be a fruitful continuation of this work.

Overall, this thesis sheds light on how local biotic interactions scale up to affect larger-scale biodiversity patterns in tropical montane ecosystems. Understanding the link between local scale interactions and large-scale patterns is key for finding ties between complementary research and for protecting and conserving biodiversity by accurately predicting future changes between species.

Chapter 1: Introduction

1.1 Background

The biodiversity found on the planet is both tremendous and variable (Darwin 1859, Schemske et al. 2009). The causes of this variation in diversity have led biologists and biogeographers to propose many different theories (Hutchinson 1959, Tilman and Pacala 1993, McGill 2010) in the search for the community assembly mechanisms (i.e., “assembly rules”; Diamond 1975) that organize and maintain these natural ecosystems (Diamond 1975, Ricklefs 2003, Krefl and Jetz 2007). Understanding the mechanisms that govern spatial patterns and the maintenance of biodiversity has widespread implications for conservation and management in a rapidly changing world (Lomolino 2001).

Changes in mechanisms across gradients

There is some disagreement as to whether ecological communities are more strongly structured by large-scale regional forces, such as speciation and dispersal, or by more local interactions, such as competition, predation and disturbance (Ricklefs 2008). The use of a gradient (environmental/ diversity) can provide insight about the relative roles of regional and local processes over large scales (Lomolino and Brown 2006, Tello et al. 2015), and resolve this long-standing difference of viewpoints.

Changes in diversity across gradients are hypothesized to be driven by changes in the mechanisms that assemble natural communities (Myers et al. 2013, Brown 2014, Tello et al.

2015). This variation in community assembly mechanisms is thought to be the result of sampling effects from species pool variation along the gradient (Kraft et al. 2011) as well as local community assembly mechanisms, such as environmental filtering, dispersal limitation, facilitation, priority effects or any number of local biotic interactions (Tello et al. 2015). However, the scale of most gradients and the complexity of the communities within them make the study of these mechanisms extremely difficult. Despite evidence of scale dependent (Tello et al. 2015), geographic (Myers et al. 2013) and abiotic (McFadden et al. 2019) variation in community assembly mechanisms, there is still great uncertainty about the forces that generate diversity gradients, as the patterns detected could be the result of many different combinations of different assembly mechanisms acting in concert.

Among the sometimes conflicting (though non-mutually exclusive) ideas about community assembly, one common thread is that most are influenced, either directly or indirectly, by biotic interactions (Schemske 2009). Biotic interactions, described as, “complex influences of organism upon organism”, by Alfred Wallace, are indeed complex, variable, and thus, difficult to quantify. This is perhaps why there has been considerable research at the community level for other proposed community assembly mechanisms (such as environmental filtering, competitive interactions, dispersal limitation and neutral assembly mechanisms; Spasojevic and Sunding 2012, Gerhold et al. 2015), but most studies of biotic interactions focus on specific species pairs, for example in plant/ herbivore systems (Dyer et al. 2007) or two species competition studies. There has been far less research investigating the community (or larger) level patterns of biotic interaction (Schmeske et al. 2009).

Biotic interactions

Biotic interactions can influence community assembly by means of predation (top down, keystone effects), competition (competitive exclusion, niche partitioning, character displacement) or through more specialized interactions with natural enemies. However, evidence suggests that their influence varies regularly across environmental gradients (Callaway et al. 2002, Schemske et al. 2009, Wiegand et al. 2012). This regular variation has helped to define some of the seminal theories (Darwin 1859, Wallace 1878, Lomolino et al. 2006) about the mechanisms underlying the major, general patterns of life on the planet. Examples include hypotheses that link the latitudinal diversity gradient to an assumed gradient in the importance of biotic interactions (Dobzhansky 1950, Schemske et al. 2009) as well as prominent ideas focused on species coexistence (Janzen 1970, Connell 1971, Kraft et al. 2015).

Competitive interactions

There has long been a focus on competition between species as a driver of occurrence patterns among members of a community (HilleRisLambers et al. 2012). Two competing species cannot coexist if both rely on the same resource (Lotka 1927, Volterra 1927). The classic belief was that for competing species to coexist, they must show niche differentiation, or maximize their fitness in different abiotic and biotic environments, in order to reduce competition (Gause 1934, Chesson 2000). This generally involves tradeoffs (Silvertown 2004), such that intraspecific competition is greater than interspecific competition (Chesson 2000, 2003).

Niche differentiation focused on competitive interactions has been supported in studies of animal communities, but is less compelling for organisms such as plants, that rely upon similar resources and show little trophic differentiation (Chase and Liebhold 2003, Silvertown 2004, but see Adler et al. 2010). In highly diverse communities, especially, the number of coexisting

species, make the idea of each one occupying its own resource-based niche untenable, despite evidence that plants do compete for vital resources (Fowler 1986, Aarssen and Epp 1990, Goldberg and Barton 1992, Casper and Jackson 1997).

Along a gradient, biotic interactions might be expected to shift from highly competitive in resource-rich, climatically benign areas to less competitive, or even mutualistic interactions in areas that impose harsh abiotic selection pressures. This expected shift is based on the assumption that abiotic factors such as wind, temperature and seasonality are less limiting to plant growth in the more benign areas (i.e. the tropics, lowland elevations), allowing plants to grow until growth or reproduction is limited by resources, while in abiotically harsh areas (i.e. the north temperate zone, high elevations) factors such as wind and temperature may limit plant growth more than resource availability (Callaway et al. 2002).

Enemy mediated interactions

Recent (Terborgh 2012, Sedio et al. 2018, Levi et al. 2019), as well as more classical studies (Dobzhansky 1950, Janzen 1970, Connell 1971) suggest that host-enemy interactions play an important role in defining community structure and maintaining the coexistence of the species found within those communities.

The way that biotic interactions influence community assembly relies upon the role that chemical defenses play in enemy selection for host defenses. Differences in chemical defenses among (plant host) species may make it possible for those species to ‘escape’, from the natural enemies of neighboring heterospecific species, giving them a fitness advantage. In contrast to other plant strategies (e.g. extrafloral nectaries, rapid leaf expansion, trichomes, synchrony and timing of leaf production), there are thousands upon thousands of unique defensive chemical

compounds, leading to a nearly infinite number of chemical niches (Coley & Kusar 2014, Coley et al. 2018). This expanded potential for niche differentiation along chemical axes, may provide a more accurate explanation of how species are able to coexist in high diversity communities.

Along a gradient, enemy mediated biotic interactions would be expected to shift from stronger and more specialized interactions at lower elevation, more resource-rich, climatically benign areas to weaker and less specialized interactions, in areas that impose stronger abiotic selection pressures (Schemske et al. 2009, Sam et al. 2020). These effects would be manifest in the changes in the diversity and abundances of the chemical defensive metabolites used by plant hosts within those areas.

Recent advances in chemical analytical techniques (Sedio 2017, Sedio et al. 2018) now allow researchers to measure community-level, chemical similarity, between and within species to predict how interactions with natural enemies affect community structure. When measured across a gradient, systematic changes in community chemical composition, can give clues about how biotic interactions change in tandem with other variables (environmental, demographic, seasonal/temporal). These clues into how biotic interactions change systematically can be used to more accurately predict the mechanisms that drive the patterns of community structure across gradients.

1.2 Chapter Overview

That there are large systematic patterns of diversity across the planet has been well established (Darwin 1859, Shemske et al. 2009). The mechanisms that maintain diversity are less well characterized. It could be that the same mechanisms that maintain diversity in an upper latitude, deciduous forest are the same, and operate by the same principles as those that maintain diversity in an incredibly species-rich tropical forest. Alternatively, the same mechanisms may operate differently or with differing strength in different regions, or it could be that there are completely different mechanisms operating among regions of the world. This dissertation will help to fill gaps in our knowledge of the mechanisms that maintain diversity.

Chapter 2 investigated whether biotic interactions vary systematically across large scale biodiversity gradients. This was done by examining the tendency of biotic interactions within communities to be more diffuse compared to communities where there are few species that consistently interact. The idea of diffuse interactions was conceived by Hubbell and Foster in 1986 and was a major tenet of the later neutral theory. A situation where diffuse interactions might occur is where in communities with many species, an individual of a given species may interact with a group of very different species than other individuals of its species. This can create a situation where selection would act on the species as a whole by the average over time of all the different neighboring species that individuals have interacted with. I tested this idea by quantifying the taxonomic and functional differences of neighboring species of individuals of each species within a plot. The plots tested covered a range of climactic conditions and species diversities.

Chapter 3 tested the hypotheses that stronger and more specialized biotic interactions contribute to higher species diversity at lower elevations and latitudes. This was done by quantifying the effect of multi-trophic, plant-enemy interactions by measuring the chemical traits used defensively. A metric that describes the pairwise chemical distance between species within a plot was tested against plot diversity and environmental variables to determine the strength of enemy-mediated biotic interactions operating within the plots.

Chapter 4 investigated how community assembly is affected by both the biotic and the abiotic environment by examining the relationships between two suites of plant functional traits: chemical traits (chemical defensive compounds) that are defined by biotic interactions, and morphological traits (resource acquisition) that are defined by abiotic site characteristics. This was accomplished by examining the functional distance, measured as functional dispersion, of the species with subplots of 5 – 20 m. The functional dispersion scores of each plot, composed of numerous subplots, were tested against climatic and environmental variables. A null model that simulated community trait patterns was used to understand the effect of both the biotic and the abiotic environments on community structure.

1.3 Large Scale Diversity Patterns: Elevational Gradients

Perhaps the most well-known pattern of biodiversity on the planet is the latitudinal diversity gradient (Forster 1778, Turner 2004). There are many explanations for the pattern of increasing diversity from the poles to the equator, but, some of the most well-accepted explanations attribute the gradient to evolutionary history, area or increased intensity of biotic

interactions (Mittelbach et al. 2007, Schemske 2009). Although this global pattern has been investigated for centuries, there is still no consensus as to its cause. However, as many studies have concluded, (Ricklefs & Schluter 1993, Lawton 1996, Brown & Lomolino 1998), general patterns in nature may result from the combined effects of many non-mutually exclusive processes, rather than the assumed independent effects of a single dominant force (Lomolino 2001). The potentially complex nature of general patterns of diversity, combined with the fact that there is only one latitudinal gradient (on two hemispheres), makes an already difficult search for answers even more daunting. Elevational gradients, however, may provide a system where it is possible to find greater insight into the workings of large-scale, general biodiversity patterns (MacArthur 1972, McCain and Grytnes 2010).

The elevational gradients found along the slopes of mountains have long acted as a testing system critical to the development of central theories of biology (Lomolino 2001). Historical examples abound, from Linnaeus's (1743) supposed Mount Ararat, in Turkey, where Noah landed the Ark, to von Humbolt's (1849) studies on Ecuador's Chimborazo volcano, and even including Wallace (1876) and Darwin's (1839) studies in Indonesia and Chile, respectively. Studies that have utilized elevational gradients in their research have found that elevational gradients on mountains roughly mirror the latitudinal diversity gradient from equator to poles as one ascends up to the peak (Stevens 1992, Brown 2001). Unlike the latitudinal diversity gradient, however, elevational gradients are replicated many times over the surface of the planet. These repeated natural experiments can provide an arena in which to separate different mechanisms operating at different spatial scales (Tello et al. 2015), by using comparative analyses (McCain 2005). Furthermore, unlike the latitudinal gradient, where biodiversity turnover/changes occur over large degrees of spatial area, the biodiversity transitions on elevational gradients are rapid,

sometimes with lowland species having dispersal ranges into the highlands and vice versa. This means that dispersal limitation may be unlikely at the edges of a species elevational range (Gworek et al. 2006). This makes for a great system in which to test different community assembly mechanisms such as abiotic filtering, dispersal limitation, competitive exclusion, neutral assembly theories.

Although it is true that elevational gradients are replicated many times over, the mountains that they occur on are not the same (Rahbek et al. 2019). The widely assumed pattern of a monotonic decrease in richness with increasing elevation is by no means consistent across every mountain range (Rahbek 1995, Grytnes 2002, Körner 2007). Nor does the assumption that mountain land area steadily decreases with elevation hold for every mountain range (Körner 2007). Some mountain ranges such as the Rockies of North America actually have larger areas slightly above the base of the mountain, and others, such as the Himalayas of Asia, contain the greatest surface area on flat plateaus found at very high altitudes. However, differences notwithstanding, elevational gradients offer a unique opportunity to gain insight into the large, general drivers of patterns organizing the life on Earth.

1.4 Introduction References

Aarssen, L. W., & Epp, G. A. (1990). Neighbour manipulations in natural vegetation a review. *Journal of Vegetation Science*, 1(1), 13-30.

Adler, P. B., Ellner, S. P., & Levine, J. M. (2010). Coexistence of perennial plants: an embarrassment of niches. *Ecology letters*, 13(8), 1019-1029.

Brown, J. H. (2014). Why are there so many species in the tropics?. *Journal of biogeography*, 41(1), 8-22.

Brown, J. H., & Lomolino, M. V. (2000). Concluding remarks: historical perspective and the future of island biogeography theory. *Global Ecology and Biogeography*, 9(1), 87-92.

- Callaway, R. M., Brooker, R. W., Choler, P., Kikvidze, Z., Lortie, C. J., Michalet, R., ... & Armas, C. (2002). Positive interactions among alpine plants increase with stress. *Nature*, 417(6891), 844-848.
- Casper, B. B., & Jackson, R. B. (1997). Plant competition underground. *Annual review of ecology and systematics*, 28(1), 545-570.
- Chase, J. M., & Leibold, M. A. (2003). *Ecological niches: linking classical and contemporary approaches*. University of Chicago Press.
- Chesson, P. (2000). Mechanisms of maintenance of species diversity. *Annual review of Ecology and Systematics*, 31(1), 343-366.
- Chesson, P. (2003). Quantifying and testing coexistence mechanisms arising from recruitment fluctuations. *Theoretical population biology*, 64(3), 345-357.
- Coley, P. D., & Kursar, T. A. (2014). On tropical forests and their pests. *Science*, 343(6166), 35-36.
- Coley, P. D., Endara, M. J., & Kursar, T. A. (2018). Consequences of interspecific variation in defenses and herbivore host choice for the ecology and evolution of *Inga*, a speciose rainforest tree. *Oecologia*, 187(2), 361-376.
- Connell, J. H. (1971). On the role of natural enemies in preventing competitive exclusion in some marine animals and in rain forest trees. *Dynamics of populations*, 298(312).
- Darwin, C. (Ed.). (1839). *The Zoology of the Voyage of HMS Beagle: Under the Command of Captain Fitzroy, RN, During the Years 1832 to 1836: Published with the Approval of the Lords Commissioners of Her Majesty's Treasury*. Smith, Elder and Company.
- Darwin, C. (1859). On the origin of species. *published on*, 24.
- Diamond, J. M. (1975). The island dilemma: lessons of modern biogeographic studies for the design of natural reserves. *Biological conservation*, 7(2), 129-146.
- Dobzhansky, T. (1950). Evolution in the tropics. *American scientist*, 38(2), 209-221.
- Dyer, L. A., Singer, M. S., Lill, J. T., Stireman, J. O., Gentry, G. L., Marquis, R. J., ... & Coley, P. D. (2007). Host specificity of Lepidoptera in tropical and temperate forests. *Nature*, 448(7154), 696-699.
- Forster, J. R. (1778). Excerpts from remarks on the organic bodies, Ch. 5 in Observations made during a voyage round the world. *Foundations of biogeography*, 19-27.
- Fowler, N. (1986). The role of competition in plant communities in arid and semiarid regions. *Annual review of ecology and Systematics*, 17(1), 89-110.
- Gause, G. F. (1934). Experimental analysis of Vito Volterra's mathematical theory of the struggle for existence. *Science*, 79(2036), 16-17.
- Gerhold, P., Cahill, J. F., Winter, M., Bartish, I. V., & Prinzing, A. (2015). Phylogenetic patterns are not proxies of community assembly mechanisms (they are far better). *Functional Ecology*, 29(5), 600-614.
- Goldberg, D. E., & Barton, A. M. (1992). Patterns and consequences of interspecific competition in natural communities: a review of field experiments with plants. *The American Naturalist*, 139(4), 771-801.

- Grytnes, J. A., & Vetaas, O. R. (2002). Species richness and altitude: a comparison between null models and interpolated plant species richness along the Himalayan altitudinal gradient, Nepal. *The American Naturalist*, 159(3), 294-304.
- Gworek, J. R., Vander Wall, S. B., & Brussard, P. F. (2007). Changes in biotic interactions and climate determine recruitment of Jeffrey pine along an elevation gradient. *Forest Ecology and Management*, 239(1-3), 57-68.
- HilleRisLambers, J., Adler, P. B., Harpole, W. S., Levine, J. M., & Mayfield, M. M. (2012). Rethinking community assembly through the lens of coexistence theory. *Annual review of ecology, evolution, and systematics*, 43.
- Hubbell, S. P. Foster R.B. (1986) Biology, chance and history and the structure of tropical rain forest tree communities. *Community ecology*"(ed. Diamond, J. & Case, TJ), 314-329.
- Hutchinson, G. E. (1959). Homage to Santa Rosalia or why are there so many kinds of animals?. *The American Naturalist*, 93(870), 145-159.
- Janzen, D. H. (1970). Herbivores and the number of tree species in tropical forests. *The American Naturalist*, 104(940), 501-528.
- Kreft, H., & Jetz, W. (2007). Global patterns and determinants of vascular plant diversity. *Proceedings of the National Academy of Sciences*, 104(14), 5925-5930.
- Kraft, N. J., Comita, L. S., Chase, J. M., Sanders, N. J., Swenson, N. G., Crist, T. O., ... & Cornell, H. V. (2011). Disentangling the drivers of β diversity along latitudinal and elevational gradients. *Science*, 333(6050), 1755-1758.
- Kraft, N. J., Godoy, O., & Levine, J. M. (2015). Plant functional traits and the multidimensional nature of species coexistence. *Proceedings of the National Academy of Sciences*, 112(3), 797-802.
- Körner, C. (2007). The use of 'altitude' in ecological research. *Trends in ecology & evolution*, 22(11), 569-574.
- Lawton, J. H. (1996). Patterns in ecology. *Oikos*, 75, 145-147.
- Lomolino, M. V. (2001). Elevation gradients of species-density: historical and prospective views. *Global Ecology and biogeography*, 10(1), 3-13.
- Lomolino, M. V., Riddle, B. R. C., & Brown, J. H. C. (2006). *Biogeography* (No. 578.09 L839b Ej. 1). Sinauer Associates, Inc.
- Lotka, A. J. (1927). Fluctuations in the abundance of a species considered mathematically. *Nature*, 119(2983), 12-12.
- Levi, T., Barfield, M., Barrantes, S., Sullivan, C., Holt, R. D., & Terborgh, J. (2019). Tropical forests can maintain hyperdiversity because of enemies. *Proceedings of the National Academy of Sciences*, 116(2), 581-586.
- McCain, C. M. (2005). Elevational gradients in diversity of small mammals. *Ecology*, 86(2), 366-372.
- McCain, C. M., & Grytnes, J. A. (2010). Elevational gradients in species richness. *eLS*.
- MacArthur, R. H., Diamond, J. M., & Karr, J. R. (1972). Density compensation in island faunas. *Ecology*, 53(2), 330-342.

- McFadden, I. R., Sandel, B., Tsirogiannis, C., Morueta-Holme, N., Svenning, J. C., Enquist, B. J., & Kraft, N. J. (2019). Temperature shapes opposing latitudinal gradients of plant taxonomic and phylogenetic β diversity. *Ecology letters*, *22*(7), 1126-1135.
- Mittelbach, G. G., Schemske, D. W., Cornell, H. V., Allen, A. P., Brown, J. M., Bush, M. B., ... & McCain, C. M. (2007). Evolution and the latitudinal diversity gradient: speciation, extinction and biogeography. *Ecology letters*, *10*(4), 315-331.
- Myers, J. A., Chase, J. M., Jiménez, I., Jørgensen, P. M., Araujo-Murakami, A., Paniagua-Zambrana, N., & Seidel, R. (2013). Beta-diversity in temperate and tropical forests reflects dissimilar mechanisms of community assembly. *Ecology letters*, *16*(2), 151-157.
- Rahbek, C. (2005). The role of spatial scale and the perception of large-scale species-richness patterns. *Ecology letters*, *8*(2), 224-239.
- Rahbek, C., Borregaard, M. K., Colwell, R. K., Dalsgaard, B., Holt, B. G., Morueta-Holme, N., ... & Fjeldså, J. (2019). Humboldt's enigma: What causes global patterns of mountain biodiversity?. *Science*, *365*(6458), 1108-1113.
- Ricklefs, R. E. (2003). A comment on Hubbell's zero-sum ecological drift model. *Oikos*, *100*(1), 185-192.
- Ricklefs, R. E. (2008). Disintegration of the ecological community: American Society of Naturalists Sewall Wright award winner address. *The American Naturalist*, *172*(6), 741-750.
- Ricklefs, R. E., & Schluter, D. (1993). Species diversity: regional and historical influences. *Species diversity in ecological communities*, *350*, 363.
- Sam, K., Koane, B., Sam, L., Mrazova, A., Segar, S., Volf, M., ... & Novotny, V. (2020). Insect herbivory and herbivores of *Ficus* species along a rain forest elevational gradient in Papua New Guinea. *Biotropica*, *52*(2), 263-276.
- Schemske, D. W., Mittelbach, G. G., Cornell, H. V., Sobel, J. M., & Roy, K. (2009). Is there a latitudinal gradient in the importance of biotic interactions?. *Annu. Rev. Ecol. Evol. Syst.*, *40*, 245-269.
- Sedio, B. E. (2017). Recent breakthroughs in metabolomics promise to reveal the cryptic chemical traits that mediate plant community composition, character evolution and lineage diversification. *New Phytologist*, *214*(3), 952-958.
- Silvertown, J. (2004). Plant coexistence and the niche. *Trends in Ecology & evolution*, *19*(11), 605-611.
- Sedio, B. E., Boya P, C. A., & Rojas Echeverri, J. C. (2018). A protocol for high-throughput, untargeted forest community metabolomics using mass spectrometry molecular networks. *Applications in plant sciences*, *6*(3), e1033.
- Spasojevic, M. J., & Suding, K. N. (2012). Inferring community assembly mechanisms from functional diversity patterns: the importance of multiple assembly processes. *Journal of Ecology*, *100*(3), 652-661.
- Stevens, G. C. (1992). The elevational gradient in altitudinal range: an extension of Rapoport's latitudinal rule to altitude. *The American Naturalist*, *140*(6), 893-911.
- Tello, J. S., Myers, J. A., Macía, M. J., Fuentes, A. F., Cayola, L., Arellano, G., ... & Jørgensen, P. M. (2015). Elevational gradients in β -diversity reflect variation in the strength of local community assembly mechanisms across spatial scales. *PLoS One*, *10*(3), e0121458.
- Terborgh, J. (2012). Enemies maintain hyperdiverse tropical forests. *The American Naturalist*, *179*(3), 303-314.

Tilman, D., & Pacala, S. (1993). The maintenance of species richness in plant communities. *Unknown Journal*, 13-25.

Turner, J. R. (2004). Explaining the global biodiversity gradient: energy, area, history and natural selection. *Basic and Applied Ecology*, 5(5), 435-448.

Volterra, V. (1927). Fluctuations in the abundance of a species considered mathematically. *Nature*, 119(2983), 12-13.

Von Humboldt, A. (1849). *Aspects of Nature, in Different Lands and Different Climates: With Scientific Elucidations. in 2 Vols.* Lea and Blanchard.

Wallace, A. R. (1876). *Die geographische Verbreitung der Thiere: nebst einer studie über die Verwandtschaften der Lebenden und ausgestobenen Faunen in ihrer Beziehung zu den früheren Veränderungen der Erdoberfläche* (Vol. 1). R. v. Zahn.

Wallace, A. R. (1878). *Tropical nature, and other essays.* Macmillan and Company.

Wiegand, T., Huth, A., Getzin, S., Wang, X., Hao, Z., Gunatilleke, C. S., & Gunatilleke, I. N. (2012). Testing the independent species' arrangement assertion made by theories of stochastic geometry of biodiversity. *Proceedings of the Royal Society B: Biological Sciences*, 279(1741), 3312-3320.

Chapter 2: Disentangling Determinism and Stochasticity in Local Tree Neighborhoods Across a Tropical Elevational Gradient

2.1 Introduction

Geographic variation in biotic interactions has been hypothesized to shape the assembly and diversity of ecological communities from local to biogeographic scales (Wallace 1878, Dobzhansky 1950, MacArthur 1972, Schemske et al. 2009, Fine 2015). Within communities, patterns of diversity reflect the net outcomes of negative species interactions including resource competition and predation, and positive species interactions including facilitation and mutualisms. Among communities, the nature and strength of these interactions may change systematically across ecological gradients. Prominent examples include changes in the strength of interspecific competition across productivity gradients (Huston 1979), changes in the strength and specialization of plant-enemy interactions across latitudinal gradients (Janzen 1970, Connell 1971), latitudinal and elevational gradients in the relative importance of abiotic and biotic controls on biodiversity (Schemske et al. 2009, HilleRisLambers et al. 2013), and changes in the relative importance of facilitative interactions across abiotic stress gradients (Callaway et al. 2002). Despite ample evidence that biotic interactions, particularly competitive interactions, have a strong effect on the composition of local communities (Goldberg & Barton 1992, Choler et al. 2001), far less is known about how the nature or strength of competitive interactions differs between communities or regions (Diamond 1975, Schemske et al. 2009). Understanding how competitive biotic interactions vary across gradients is increasingly important for predicting how

species ranges, species diversity, and ecosystems respond to environmental change (Urban et al. 2016).

Theory suggests that competitive biotic interactions may shift from more deterministic outcomes in low-diversity communities to more stochastic outcomes in high-diversity communities (Hubbell and Foster 1986, Hurtt and Pacala 1995, Chisholm and Pacala 2011). Based on their observations in a tropical tree community, Hubbell and Foster (1986) hypothesized that chance and history are more important determinants of community structure in species-rich communities than either competition or abiotic factors. They observed a high degree of variability in the taxonomic identities of neighboring trees from tree neighborhood to tree neighborhood among individuals of the same species. In this way, high species diversity could lead to unpredictable and inconsistent species-species interactions, which leads to diffuse evolution by weakening selection for niche differentiation. Ecological interactions among species in these communities should be weak, and result in stochastic community assembly. In a spatial analysis of three large forest plots with different tree species richness (two tropical and one temperate), Wiegand et al. (2012) found support for the hypothesis that the spatial arrangement of species becomes increasingly stochastic in more species-rich communities. These findings suggest that stochastic outcomes of diluted species interactions can overpower effects of more predictable (deterministic) species interactions in species-rich communities.

The Hubbell-Foster hypothesis makes two key predictions about how the composition of biotic neighborhoods should vary across biodiversity gradients. First, it predicts that neighborhoods of individuals surrounding a given focal species (conspecific neighborhoods) are more variable (high compositionally dissimilarity) in high-diversity communities than in low-diversity communities. This corresponds to the original observation by Hubbell and Foster (1986)

and represents the proposed effects of high diversity that leads to diffuse evolution and weak interspecific interactions. Second, it predicts that compositional differences among conspecific neighborhoods are closer to random (more stochastic) in high-diversity communities than in low-diversity communities. This reflects the proposed outcome of weak species interactions on local species distributions and community structure. Both predictions can be tested by analyzing patterns in the taxonomic and functional-trait composition of local neighborhoods. If functionally similar species compete more strongly for shared resources than functionally dissimilar species, functionally similar species are less likely to coexist at local scales (Adler et al. 2013), resulting in non-random patterns in the trait composition of conspecific neighborhoods. Therefore, the composition of functional traits at neighborhood scales can provide insights into how biotic and abiotic processes structure communities (e.g., Kraft et al. 2008, Fortunel et al. 2016), especially in communities with functionally redundant species that share similar traits (Fukami et al. 2005). Research suggests that functional trait patterns may reflect abiotic factors of the environment (Dubuis et al. 2013). These abiotic factors are taxon independent and may have value in prediction of community arrangements (Webb et al. 2010, Cadotte et al. 2015). However, identification of functional trait patterns at biogeographic scales remains elusive (Lamanna et al. 2014) and the generality of these processes in shaping large-scale diversity gradients remains unknown. Moreover, empirical tests of these predictions have largely focused on either the taxonomic or functional composition of neighborhoods within single communities (Wiegand et al. 2017), resulting in key gaps in our understanding of how neighborhood-scale patterns and processes vary across larger-scale diversity gradients.

In this study, we tested the Hubbell and Foster hypothesis by examining if determinism and stochasticity in local tree neighborhoods change in relative importance across a tropical

elevational gradient. As trees are sessile organisms, their immediate biotic neighborhood has a strong influence on their fitness (Hubbell et al. 2001, Canham & Uriarte 2006, Uriarte et al. 2010), making them an ideal system to study how competitive interactions affect the structure of local neighborhoods. Using 31 1-ha forest plots distributed across an ~3000 m elevational-diversity gradient in the Bolivian Andes (15-137 tree species per plot), we analyzed patterns in the taxonomic and functional composition of conspecific neighborhoods at three neighborhood scales (neighborhood radii of 10, 15, and 20 m). To test the prediction that compositional differences among conspecific neighborhoods increase from low-diversity to high-diversity communities, we first calculated the taxonomic and functional dissimilarity among neighborhoods of all individuals of each species within a plot. We then tested whether the mean dissimilarity across species' neighborhoods increased with forest-plot diversity. To test the prediction that compositional differences among conspecific neighborhoods are more random in high-diversity communities than in low-diversity communities, we used a null model to randomize the taxonomic identity and functional traits of both all (conspecific and heterospecific) and heterospecific only neighboring trees. We then tested whether mean deviations from the null model decreased with forest-plot diversity. Larger deviations from the null model would indicate a stronger effect of local ecological processes that diversify or homogenize the species taxonomic or trait composition of neighbors across conspecific neighborhoods.

2.2 Methods

Forest plot network in the Bolivian Andes

Our study utilized data collected by the Madidi Project, a 20-year collaboration between the Herbario Nacional de Bolivia (La Paz, Bolivia) and the Missouri Botanical Garden, (St. Louis, USA) to document the flora of the Madidi region in the Andean Mountains of Bolivia (Friedman-Rudovsky 2012, Tello et al. 2015). The elevational gradient encompasses different forest types and a broad range of abiotic conditions (Rafiqpoor & Ibish 2004). The Madidi Project dataset includes a network of 50, 1-ha forest plots in which all woody plants (hereafter trees) with a diameter at breast height (DBH) ≥ 10 have been tagged, mapped, measured, and identified to species or morphospecies. A small fraction of individuals ($<0.5\%$) could not be identified in this way and were removed from all analyses. For this study, we used 31 plots in which 8 plant functional traits were measured on tagged trees throughout 3-4 censuses, roughly every 5 years (Table S1). The 31 forest plots range in elevation from 724-3334 m and species richness from 15 species to 137 species (Table S2).

Functional trait data

For each species in each forest plot, we measured eight functional traits including leaf, stem, and whole-plant traits associated with plant life-history strategies (Table S1; Figure S1). Traits were chosen to encompass plant life history and overall growth strategies: maximum height and maximum diameter at breast height (DBH), which are associated positively with the position of the species in the vertical light gradient of the vegetation, competitive vigor, reproductive size, potential lifespan, and whether a species is able to establish and attain reproductive size between disturbance events; specific leaf area (SLA), which is associated with higher photosynthetic rates and grow faster, but lower resistance to damage; relative growth rate

(RGR), which is associated with tree life history strategy with respect to environmental productivity (soil nutrients, light moisture), where a higher RGR indicates faster growth but less defenses; leaf thickness, which is associated with abiotic environment, where sun leaves tend to be thicker than shade leaves as well as lower in N%, slower in CO₂ diffusion and subject to more internal shading of chloroplasts; leaf size, which represents a compromise between growth and resource use efficiency, where larger leaf sizes generally grow faster, but are less efficient; twig specific density is positively associated with stability, defense, architecture, hydraulics, C gain and growth potential of plants; relative twig bark thickness represents a tradeoff between growth and survival as thicker bark protects buds and tissues against attack by pathogens, herbivores, frost or drought, but is associated with slower growth (Perez-Harguindeguy et al. 2016).

Functional traits were collected over multiple censuses from 5 individuals of each species, when possible, per plot. The field collection and laboratory methods used in the determination of functional trait values were adapted from (Cornelissen et al. 2003) and can be read in full detail in the Madidi Project methods manual (Jørgensen et al. 2015).

The trait values assigned to each species were the average trait values for each species in each plot. The traits used were for the most part not strongly correlated ($r^2 = < 0.27$), however traits such as SLA and leaf thickness ($r^2 = -0.65$), and maximum height and DBH ($r^2 = 0.68$) were more strongly correlated. We calculated the mean pairwise distance (MPD) of trait values both as a multivariate measure, which includes all eight functional traits combined, as well as univariate measures for single traits (Table S1; Figure S1).

Taxonomic and functional composition of tree neighborhoods

We calculated the taxonomic and functional composition of tree neighborhoods in three steps. First, we defined a focal tree's biotic neighborhood (hereafter neighborhood) as all the individuals (both conspecific and heterospecific) that occur within a circle of a given radius drawn around a focal tree. Neighborhoods were built this way for all individuals of each species in a plot. Individuals were excluded if their distance to the edge of the plot was smaller than the neighborhood radius. This prevented using trees with artificially incomplete neighborhoods. We performed all analyses at three neighborhood sizes: 10, 15 and 20 m radii. In addition, we performed all analyses by defining neighborhoods including only heterospecific neighbors. These analyses yielded similar results to analyses including both conspecific and heterospecific neighbors. For simplicity, we therefore present results including both conspecific and heterospecific neighbors.

Second, for each species in each forest plot, we calculated the taxonomic dissimilarity among all intra-specific pairs of neighborhoods (Figure 1). We calculated pairwise taxonomic dissimilarity using an abundance-based dissimilarity metric (Bray-Curtis). We calculated functional dissimilarity using the mean pairwise distance (MPD) of trait values with the `comdist` function (`picante` package) in R.

Third, for each species in each forest plot, we calculated mean pairwise dissimilarity of each neighborhood, both taxonomically and functionally, which represents the typical variability in neighborhood composition from one individual of that species to the next. We aggregated this data to the community level by calculating an abundance weighted mean for the whole forest plot. This plot-wide dissimilarity metric represents the typical neighborhood dissimilarity for all of the species in a plot.

Null model of random assembly

To evaluate how deterministic the taxonomic and functional composition of neighborhoods are, we compared the observed compositional neighborhood dissimilarities to values produced by a randomized null model. The null model accounts for the mechanisms responsible for plot-level species abundances and richness but eliminates the effects of the local ecological processes that determine the composition of individual tree neighborhoods. In this way, empirical deviations from null model expectations can be used to quantify the relative effects of those local ecological interactions (Chase & Myers 2011; Kraft et al. 2011). High deviations, in this context, suggest a stronger role for local deterministic interactions.

For the null model, we defined a species pool for each forest plot as the total number and relative abundances of species observed in that plot. In each iteration of the null model, null neighborhoods were created by randomly sampling individuals from the plot without replacement from the species pool. In this way, the null model constrained the plot-level species abundance distribution (SAD), as well as the spatial distribution of all trees, to be the same in null and observed datasets, but switches the identities of the species contained within. Next, we calculated null neighborhood dissimilarities using the same methods as for the observed data. We ran a fixed number of iterations of the null model ($n = 1000$). This produced a distribution of null values of taxonomic and trait dissimilarities that were expected due to sampling from plots of variable diversities along the gradient, but in the absence of local ecological interactions. Based on this frequency distribution, we calculated a standard effect size (SES; Kraft et al. 2011) for pairwise SES values between all conspecific neighborhoods and then averaged all SES values for

each species in each plot, and across all species in each plot, as described for the observed data. The SES was calculated as:

$$\text{SES} = \frac{\text{obs. dissimilarity} - \text{mean exp. dissimilarity}}{\text{SD of expected dissimilarities}}$$

Where the **mean exp. dissimilarity** and the **SD of expected dissimilarities** are the average and standard deviation of the frequency distribution of null values for a plot and the *obs. dissimilarity* are the plot dissimilarity values calculated from the observed data.

Hypothesis testing

In order to investigate the effects that community diversity has on community assembly, we quantified diversity in several ways: species richness is an intuitive measure but does not take into account relative abundances or evenness of species in communities, the Shannon diversity index which does take into account community evenness but not abundance, and the inverse Simpson's index which accounts for both community relative abundances and evenness. Results are shown for the inverse Simpson's measure of diversity throughout. We obtained quantitatively similar results using the Shannon diversity index and using observed plot species richness.

To test our predictions, we regressed the plot-wide mean observed neighborhood dissimilarities and plot-wide mean SES values against plot species diversity. If, as proposed by the Hubbell and Foster hypothesis, competitive interactions among tree species shift from more deterministic outcomes in low-diversity communities to more stochastic outcomes in high-diversity communities, we would expect the mean observed dissimilarities of taxonomic and

functional neighborhoods to increase with plot species diversity. In addition, we would expect mean SES values of neighborhood taxonomic and functional dissimilarities to vary systematically with plot species diversity, from larger deviations from the null model in lower-diversity plots (i.e. less stochastic) to smaller deviations from the null model in higher-diversity plots (i.e. more stochastic).

2.3 Results

Taxonomic dissimilarity of tree neighborhoods

Across forest plots, mean observed taxonomic dissimilarity among conspecific tree neighborhoods increased significantly with plot species diversity at all three neighborhood sizes (Figure 2). This relationship was strongest at the smallest neighborhood size (10-m neighborhood radius: $P = 1.1e-09$, $r^2 = 0.74$; Figure 2A), very similar at the intermediate neighborhood size (15-m neighborhood radius; $P = 1.052e-09$, $r^2 = 0.74$; Figure 2C), and weaker at the largest neighborhood size (20-m neighborhood radius: $P = 4.7e-09$, $r^2 = 0.72$; Figure 2E).

Null-model deviations of taxonomic dissimilarity (mean standardized effect sizes; SES) decreased significantly with plot species diversity at the all neighborhood sizes (Figure 2). This relationship was strongest at the smallest neighborhood size (10-m neighborhood radius: $P = 0.0003$, $r^2 = 0.37$; Figure 2B), weaker at the intermediate neighborhood size (15-m neighborhood radius; $P = 0.0008$, $r^2 = 0.32$; Figure 2D), and weakest, though significant at the largest neighborhood size (20-m neighborhood radius: $P = 0.03$, $r^2 = 0.12$; Figure 2F). At all three

neighborhood sizes, null-model deviations were positive in most forest plots, indicating higher taxonomic dissimilarity among conspecific neighborhoods than expected by chance.

Functional dissimilarity of tree neighborhoods

In contrast to taxonomic dissimilarity, mean observed functional dissimilarity of conspecific neighborhoods and null-model deviations of functional dissimilarity were generally unrelated to plot species diversity at all three neighborhood sizes (Figure 3; Table S2). Similarly, there were few significant patterns observed for all traits combined and individual traits, with a few exceptions. First, mean observed functional dissimilarity of DBH was positively related to plot species diversity at the 15-m neighborhood scale ($P = 0.034$, $r^2 = 0.13$; Figure 3E; SES: $P = 0.84$, $r^2 = -0.035$; Figure 3F), but this relationship was not significant for null-model deviations of DBH (Figure 3F). This relationship also held at the 10m ($P = 0.032$, $r^2 = 0.13$; Figure S3E; SES: $P = 0.96$, $r^2 = -0.037$; Figure S3F) and 20m ($P = 0.035$, $r^2 = 0.12$; Figure S3E; SES: $P = 0.84$, $r^2 = -0.035$; Figure S3F) scales. Also, mean observed dissimilarity in RGR varied positively with plot species diversity at the 10m neighborhood size ($P = 0.034$, $r^2 = 0.13$; Figure S3C), and the null deviations varied negatively with plot diversity (SES: $P = 0.012$, $r^2 = 0.19$; Figure S3D). At the 20m scale, the observed relationship of RGR with plot diversity was not significant, while there was a significant negative relationship with the SES comparisons ($P = 0.11$, $r^2 = 0.70$; Figure S3C; SES: $P = 0.012$, $r^2 = 0.19$; Figure S3D). For the majority of forest plots, null-model deviations of all traits combined, and specific leaf area (SLA) were negative (Figure 3B, D), indicating lower mean functional dissimilarity among conspecific neighborhoods than expected by chance. Null-model deviations for other individual traits (e.g., maximum tree

diameter & twig specific wood density; showed more variable patterns, with a combination of positive, zero, and negative SES (Figure 3F, H).

2.4 Discussion

Overall, we found strong support taxonomically but no support functionally for the Hubbell-Foster hypothesis. In support of our predictions, we found that the dissimilarity in neighborhood species composition increased with plot species diversity at all three neighborhood scales (Prediction 1). We also found that null-model deviations of taxonomic dissimilarity decreased with plot species diversity at small to intermediate neighborhood scales (10–15 m neighborhood radii) (Prediction 2), suggesting more random differences in species composition among local neighborhoods in higher-diversity communities. We found that the relationship between null-model deviations of taxonomic composition and forest plot diversity was strongest at the smallest neighborhood size (10-m radius) and became insignificant at the largest neighborhood size (20-m radius). This result is in accord with other studies that have observed the strongest effects of neighboring trees on focal individuals at smaller neighborhood scales (Hubbell et al. 2001, Comita et al. 2010, Wiegand et al. 2012, Wiegand & Moloney 2014, Murphy et al. 2017). However, neither of the predictions were upheld for most tests when the functional composition of tree neighborhoods were analyzed. Together, these findings suggest that competitive interactions play a stronger role in shaping the taxonomic composition of local biotic neighborhoods, but not necessarily the composition of species' traits, in lower-diversity tree communities. In contrast, competitive interactions appear to be more unpredictable among neighborhoods of the same species in higher-diversity tree communities.

This tendency towards more random competitive interactions among species in higher-diversity communities can help to explain differences in species diversity across ecological gradients. In high diversity communities, where individuals of a single species can have dissimilar neighbors, the selective strength of a single pairwise interaction is, by necessity, reduced, and so this benign abiotic environment may lead to a community where competitive interactions among pairs of species are largely diffuse. A feedback loop of a slow accumulation of species could occur as new species disperse to or originate within a local community, but are not selectively excluded over time, resulting in the maintenance of higher-diversity communities. In lower-diversity communities, in contrast, more frequent, predictable competitive interactions would lead to species being excluded and thus lower diversity over time.

We found that the taxonomic dissimilarity of tree neighborhoods varied systematically across the diversity gradient, from less variable and non-random in lower-diversity forest plots to more variable and random in higher-diversity forest plots (Figure 2). Previous studies investigating the effects of competitive interactions on community assembly have largely focused on environmentally driven, deterministic niche differences (Weiher & Keddy 1995, Chesson 2000, Cavender-Bares et al. 2006, Mayfield & Levine 2010). However, the unpredictable identity of neighbors among individuals of a focal species in high-diversity communities may increase the unpredictability of neighborhood interactions. The tendency of species-rich communities to have few dominant and many rare species (Hubbell 2015, Silk et al. 2015) further increases the unpredictability of individual biotic neighborhoods and further limits opportunities for pairwise species interactions, as there may be less individuals of each species with which to interact (Gaston 2000). This variability among biotic neighborhoods also has the potential to slow rates of competitive exclusion, when recruitment of individuals of a species in

favorable biotic neighborhoods counteracts competitive exclusion in less favorable biotic neighborhoods (Hurtt & Pacala 1995, Mayfield & Levine 2010, Wang et al. 2012). These effects projected over evolutionary timescales could result in selection for ecological (or functional) equivalence (Hubbell 2006, Wang et al. 2016), where many species with similar life-histories and functional traits co-occur, and the composition of biotic neighborhoods is more strongly influenced by stochastic birth-death processes than by deterministic niche differences among species.

Previous studies have suggested that seemingly random (neutral) patterns in the composition of biotic neighborhoods could result from a combination of stochastic and deterministic processes (Hurtt and Pacala 1995, Volkov et al. 2009, McGill 2010, Wiegand et al. 2012, Wang et al. 2016). Consistent with this idea, we found that the taxonomic composition of tree neighborhoods in high-diversity plots still deviated, if slightly in some cases, from the null model. Similar to previous studies in tree communities (Wiegand et al. 2012), our results suggest that stochastic species interactions are relatively more important in higher-diversity communities, and possibly overwhelm the signals of predictable interactions/ associations between individuals of species located near to each other (Wiegand et al. 2012, Parmentier et al. 2014, Perry et al. 2014, Wang et al. 2016). In low-diversity plots, where there are fewer species and where there may be an increasing dominance of certain species, the strength of deterministic, pairwise species interactions may increase, leading to greater neighborhood predictability and thus, greater deviations from null models of random community assembly, as observed in this and previous studies involving plant communities along ecological gradients (Wiegand et al. 2012, Perry et al. 2014, Wang et al. 2016, but see Callaway et al. 2002).

In contrast to the taxonomic structure of tree neighborhoods, the functional structure of tree neighborhoods did not vary systematically across the diversity gradient. This suggests that overall, community diversity does not influence how predictable or unpredictable the functional neighborhoods of trees are along the diversity gradient and that from a functional perspective, the dissimilarity of neighborhoods is similar for species in high-diversity communities or in low-diversity communities (prediction 1). Moreover, the strength of local deterministic processes in shaping variation in functional composition among neighborhoods is the same along the diversity gradient (prediction 2). The lack of a systematic relationship between the functional composition of neighborhoods and species diversity may reflect a high degree of functional redundancy among co-occurring species. If species' traits are highly redundant within a community, then high diversity introduces variation in species composition, but the redundancy causes the same functional composition despite different taxonomic compositions. If trees within a community are separated into similar life-history guilds, such as shade-tolerant tree species in tropical forests (Hubbell & Foster 1986, Hubbell 2006), then this tendency toward similarity in functional composition will be greater (Hubbell 2001, Webb et al. 2006, Uriarte et al. 2010, Paine et al. 2012, Lebrija-Trejos et al. 2014). Another potential explanation could be dispersal limitation (Hurtt & Pacala 1995, Muller-Landau et al. 2002, Uriarte et al. 2010), where a high proportion of conspecifics increases high functional redundancy within tree neighborhoods. However, we obtained similar results when conspecifics were removed from the analyses, so this explanation is less likely. A more probable explanation for the discrepancy between taxonomic and functional patterns is that the elevational gradient is environmentally/abiotically correlated to the observed functional patterns (Read et al. 2014). For example, abiotic filtering would cause species with similar functional traits to co-occur more frequently (Keddy 1992, Weiher et al.

1998, Kraft et al. 2008, Swenson & Enquist 2009), as observed in many functional trait-based analyses (Vellend 2016). Another factor may be convergent evolution, where similar functional traits evolve across different taxonomic groups over evolutionary timescales (Clausen et al. 1940, Ackerly & Reich 1999, Swenson & Enquist 2007). The lack of functional patterns within local communities could be strongly influenced by the abiotic conditions found over the elevational gradient that may overwhelm any effects of the competitive biotic interactions on the functional arrangement of the neighborhoods.

Future studies of tree neighborhoods across diversity gradients can be expanded to consider several additional processes. First, studies can evaluate the relative roles of competitive interactions among different life stages. Our dataset only included older trees ≥ 10 cm DBH. Analyses of younger life stages, including individuals in the seedling and sapling stages, may reveal stronger competitive interactions (Hubbell and Foster 1986, Givnish 2010, Swamy & Terborgh 2010). Second, some functional traits may be more strongly associated with abiotic niche requirements than competitive ability. For example, smaller leaf size (leaf area) is strongly associated with heat stress, cold stress, and high-radiation stress (Perez-Harguindeguy et al. 2013). Other traits, particularly chemical or physical plant defenses against natural enemies (Terborgh 2012, Sedio et al. 2018) may well show different patterns. Finally, diversity gradients are often correlated with abiotic factors (temperature, precipitation, seasonality), potentially confounding the observed relationships. The shift in the strength or importance of stochastic biotic interactions along the diversity gradient might be further clarified by identifying consistent, pairwise relationships between species in a community.

Although geographic variation in biotic interactions has been a central theme in ecology and biogeography for decades, surprisingly little is known about how the nature and strength of

competitive interactions varies among local communities across ecological gradients. Our findings suggest that competitive interactions play a stronger role in shaping the taxonomic composition of local biotic neighborhoods, but not necessarily the composition of species' traits, in lower-diversity tree communities. Understanding how competitive biotic interactions vary across gradients is increasingly important for predicting how species ranges, species diversity, and ecosystems respond to environmental change. Overall, our results suggest that the role of competitive biotic interactions in the structuring of biotic communities should be quantified and evaluated in the consideration of large-scale biodiversity gradients.

Acknowledgements

We thank the Dirección General de Biodiversidad, the Bolivian Park Service (SERNAP), the Madidi National Park and local communities for permits, access and collaboration in Bolivia, where fieldwork was supported by the National Science Foundation (DEB 0101775, DEB 0743457, DEB 1836353). Additional financial support to the Madidi Project was provided by the Missouri Botanical Garden, the National Geographic Society (NGS 7754-04 and NGS 8047-06), The German Embassy in Bolivia, International Center for Advanced Renewable Energy and Sustainability (I-CARES) at Washington University in St. Louis, the Comunidad de Madrid (Spain), Consejo Superior de Investigaciones Científicas (Spain), Centro de Estudios de América Latina (Banco Santander and Universidad Autónoma de Madrid, Spain), and the Taylor and Davidson families. This work was developed in part during the working group 'A Synthesis of Patterns and Mechanisms of Diversity and Forest Change in the Andes' funded by the Living Earth Collaborative at Washington University in St. Louis. We thank all the researchers, students

and local guides who were involved in the collection of the data, particularly Maritza Cornejo, Alejandro Araujo, Narel Paniagua and Peter Jørgenson. David Henderson was supported by a Catharine M. Lieneman Botany Fellowship at Washington University in St. Louis.

2.5 References

- Ackerly, D. D., & Reich, P. B. (1999). Convergence and correlations among leaf size and function in seed plants: a comparative test using independent contrasts. *American Journal of Botany*, 86(9), 1272-1281.
- Adler, P. B., Fajardo, A., Kleinhesselink, A. R., & Kraft, N. J. (2013). Trait-based tests of coexistence mechanisms. *Ecology letters*, 16(10), 1294-1306.
- Cadotte, M. W., Arnillas, C. A., Livingstone, S. W., & Yasui, S. L. E. (2015). Predicting communities from functional traits. *Trends in Ecology & Evolution*, 30(9), 510-511.
- Callaway, R. M., Brooker, R. W., Choler, P., Kikvidze, Z., Lortie, C. J., Michalet, R., ... & Cook, B. J. (2002). Positive interactions among alpine plants increase with stress. *Nature*, 417(6891), 844-848.
- Canham, C. D., Papaik, M. J., Uriarte, M., McWilliams, W. H., Jenkins, J. C., & Twery, M. J. (2006). Neighborhood analyses of canopy tree competition along environmental gradients in New England forests. *Ecological applications*, 16(2), 540-554.
- Cavender-Bares, J., Keen, A., & Miles, B. (2006). Phylogenetic structure of Floridian plant communities depends on taxonomic and spatial scale. *Ecology*, 87(sp7), S109-S122.
- Chase, J. M., & Myers, J. A. (2011). Disentangling the importance of ecological niches from stochastic processes across scales. *Philosophical transactions of the Royal Society B: Biological sciences*, 366(1576), 2351-2363.
- Chesson, P. (2000). General theory of competitive coexistence in spatially-varying environments. *Theoretical population biology*, 58(3), 211-237.
- Chisholm, R. A., & Pacala, S. W. (2011). Theory predicts a rapid transition from niche-structured to neutral biodiversity patterns across a speciation-rate gradient. *Theoretical Ecology*, 4, 195-200.
- Choler, P., Michalet, R., & Callaway, R. M. (2001). Facilitation and competition on gradients in alpine plant communities. *Ecology*, 82(12), 3295-3308.
- Clausen, J., Keck, D. D., & Hiesey, W. M. (1940). Experimental studies on the nature of species. I. Effect of varied environments on western North American plants. *Experimental studies on the nature of species. I. Effect of varied environments on western North American plants.*, (520).
- Comita, L. S., Muller-Landau, H. C., Aguilar, S., & Hubbell, S. P. (2010). Asymmetric density dependence shapes species abundances in a tropical tree community. *Science*, 329(5989), 330-332.
- Connell, J. H. (1971). On the role of natural enemies in preventing competitive exclusion in some marine animals and in rain forest trees. *Dynamics of populations*, 298(312).

- Cornelissen, J. H. C., Lavorel, S., Garnier, E., Díaz, S., Buchmann, N., Gurvich, D. E., ... & Poorter, H. (2003). A handbook of protocols for standardised and easy measurement of plant functional traits worldwide. *Australian journal of Botany*, 51(4), 335-380.
- Diamond, J. M. (1975). The island dilemma: lessons of modern biogeographic studies for the design of natural reserves. *Biological conservation*, 7(2), 129-146.
- Dobzhansky, T. (1950). Evolution in the tropics. *American scientist*, 38(2), 209-221.
- Dubuis, A., Rossier, L., Pottier, J., Pellissier, L., Vittoz, P., & Guisan, A. (2013). Predicting current and future spatial community patterns of plant functional traits. *Ecography*, 36(11), 1158-1168.
- Fine, P. V. (2015). Ecological and evolutionary drivers of geographic variation in species diversity. *Annual Review of Ecology, Evolution, and Systematics*, 46, 369-392.
- Fortunel, C., Valencia, R., Wright, S. J., Garwood, N. C., & Kraft, N. J. (2016). Functional trait differences influence neighbourhood interactions in a hyperdiverse Amazonian forest. *Ecology letters*, 19(9), 1062-1070.
- Friedman-Rudovsky, J. (2012). Taking the Measure of Madidi. *Science*. 337: 285–287.
- Fuentes, A. (2005). Una introducción a la vegetación de la región de Madidi. *Ecología en Bolivia*, 40(3), 1-31.
- Fukami, T., Martijn Bezemer, T., Mortimer, S. R., & van der Putten, W. H. (2005). Species divergence and trait convergence in experimental plant community assembly. *Ecology letters*, 8(12), 1283-1290.
- Gaston, K. J. (2000). Global patterns in biodiversity. *Nature*, 405(6783), 220-227.
- Givnish, T. J. (1999). On the causes of gradients in tropical tree diversity. *Journal of ecology*, 87(2), 193-210.
- Givnish, T. J. (2010). Ecology of plant speciation. *Taxon*, 59(5), 1326-1366.
- Goldberg, D. E., & Barton, A. M. (1992). Patterns and consequences of interspecific competition in natural communities: a review of field experiments with plants. *The American Naturalist*, 139(4), 771-801.
- HilleRisLambers, J., Harsch, M. A., Ettinger, A. K., Ford, K. R., & Theobald, E. J. (2013). How will biotic interactions influence climate change–induced range shifts?. *Annals of the New York Academy of Sciences*, 1297(1), 112-125.
- Hubbell, S. P. (2006). Neutral theory and the evolution of ecological equivalence. *Ecology*, 87(6), 1387-1398.
- Hubbell, S. P. (2015). Estimating the global number of tropical tree species, and Fisher’s paradox. *Proceedings of the National Academy of Sciences*, 112(24), 7343-7344.
- Hubbell, S. P., Ahumada, J. A., Condit, R., & Foster, R. B. (2001). Local neighborhood effects on long-term survival of individual trees in a neotropical forest. *Ecological Research*, 16, 859-875.
- Hubbell, S. P. Foster RB (1986) Biology, chance and history and the structure of tropical rain forest tree communities. *Community ecology* (ed. Diamond, J. & Case, TJ), 314-329.
- Hurtt, G. C., & Pacala, S. W. (1995). The consequences of recruitment limitation: reconciling chance, history and competitive differences between plants. *Journal of theoretical biology*, 176(1), 1-12.

- Huston, M. (1979). A general hypothesis of species diversity. *The American Naturalist*, 113(1), 81-101.
- Janzen, D. H. (1970). Herbivores and the number of tree species in tropical forests. *The American Naturalist*, 104(940), 501-528.
- Jørgensen, P. M., Fuentes, A. F., Miranda, T., & Cayola, L. Manual de Trabajo.
- Keddy, P. A. (1992). Assembly and response rules: two goals for predictive community ecology. *Journal of vegetation science*, 3(2), 157-164.
- Kraft, N. J., Comita, L. S., Chase, J. M., Sanders, N. J., Swenson, N. G., Crist, T. O., ... & Myers, J. A. (2011). Disentangling the drivers of β diversity along latitudinal and elevational gradients. *Science*, 333(6050), 1755-1758.
- Kraft, N. J., Valencia, R., & Ackerly, D. D. (2008). Functional traits and niche-based tree community assembly in an Amazonian forest. *Science*, 322(5901), 580-582.
- Lamanna, C., Blonder, B., Violle, C., Kraft, N. J., Sandel, B., Šímová, I., ... & Enquist, B. J. (2014). Functional trait space and the latitudinal diversity gradient. *Proceedings of the National Academy of Sciences*, 111(38), 13745-13750.
- Lebrija-Trejos, E., Wright, S. J., Hernández, A., & Reich, P. B. (2014). Does relatedness matter? Phylogenetic density-dependent survival of seedlings in a tropical forest. *Ecology*, 95(4), 940-951.
- MacArthur, R. H., Diamond, J. M., & Karr, J. R. (1972). Density compensation in island faunas. *Ecology*, 53(2), 330-342.
- Mayfield, M. M., & Levine, J. M. (2010). Opposing effects of competitive exclusion on the phylogenetic structure of communities. *Ecology letters*, 13(9), 1085-1093.
- McGill, B. J. (2010). Towards a unification of unified theories of biodiversity. *Ecology letters*, 13(5), 627-642.
- Muller-Landau, H. C., Wright, S. J., Calderón, O., Hubbell, S. P., & Foster, R. B. (2002). Assessing recruitment limitation: concepts, methods and case-studies from a tropical forest. In *Seed dispersal and frugivory: ecology, evolution and conservation. Third International Symposium-Workshop on Frugivores and Seed Dispersal, São Pedro, Brazil, 6-11 August 2000* (pp. 35-53). Wallingford UK: CABi Publishing.
- Murphy, S. J., Wiegand, T., & Comita, L. S. (2017). Distance-dependent seedling mortality and long-term spacing dynamics in a neotropical forest community. *Ecology letters*, 20(11), 1469-1478.
- Myers, J. A., Tello, J. S., Jørgensen, P., Araujo-Murakami, A., Cayola-Pérez, L., Cornejo-Mejía, M., ... & Loza-Rivera, M. I. Untangling determinism and stochasticity within local species neighborhoods across a tropical biodiversity gradient.
- Paine, C. T., Norden, N., Chave, J., Forget, P. M., Fortunel, C., Dexter, K. G., & Baraloto, C. (2012). Phylogenetic density dependence and environmental filtering predict seedling mortality in a tropical forest. *Ecology Letters*, 15(1), 34-41.
- Parmentier, I., Réjou-Méchain, M., Chave, J., Vleminckx, J., Thomas, D. W., Kenfack, D., ... & Hardy, O. J. (2014). Prevalence of phylogenetic clustering at multiple scales in an African rain forest tree community. *Journal of ecology*, 102(4), 1008-1016.

- Perez-Harguindeguy, N., Diaz, S., Garnier, E., Lavorel, S., Poorter, H., Jaureguiberry, P., ... & Cornelissen, J. H. C. (2016). Corrigendum to: New handbook for standardised measurement of plant functional traits worldwide. *Australian Journal of botany*, *64*(8), 715-716.
- Perry, G. L., Miller, B. P., Enright, N. J., & Lamont, B. B. (2014). Stochastic geometry best explains spatial associations among species pairs and plant functional types in species-rich shrublands. *Oikos*, *123*(1), 99-110.
- Rafiqpoor, D., & Ibsch, P. (2004). The physical environment of Bolivia. *Biodiversity: The richness of Bolivia. State of knowledge and conservation. Editorial Fundación Amigos de la Naturaleza, Santa Cruz de la Sierra*, 4-45.
- Read, Q. D., Moorhead, L. C., Swenson, N. G., Bailey, J. K., & Sanders, N. J. (2014). Convergent effects of elevation on functional leaf traits within and among species. *Functional ecology*, *28*(1), 37-45.
- Schemske, D. W., Mittelbach, G. G., Cornell, H. V., Sobel, J. M., & Roy, K. (2009). Is there a latitudinal gradient in the importance of biotic interactions?. *Annu. Rev. Ecol. Evol. Syst.*, *40*, 245-269.
- Sedio, B. E., Parker, J. D., McMahon, S. M., & Wright, S. J. (2018). Comparative foliar metabolomics of a tropical and a temperate forest community.
- Slik, J. F., Arroyo-Rodríguez, V., Aiba, S. I., Alvarez-Loayza, P., Alves, L. F., Ashton, P., ... & Hurtado, J. (2015). An estimate of the number of tropical tree species. *Proceedings of the National Academy of Sciences*, *112*(24), 7472-7477.
- Swamy, V., & Terborgh, J. W. (2010). Distance-responsive natural enemies strongly influence seedling establishment patterns of multiple species in an Amazonian rain forest. *Journal of Ecology*, *98*(5), 1096-1107.
- Swenson, N. G., & Enquist, B. J. (2007). Ecological and evolutionary determinants of a key plant functional trait: wood density and its community-wide variation across latitude and elevation. *American Journal of Botany*, *94*(3), 451-459.
- Swenson, N. G., & Enquist, B. J. (2009). Opposing assembly mechanisms in a Neotropical dry forest: implications for phylogenetic and functional community ecology. *Ecology*, *90*(8), 2161-2170.
- Terborgh, J. (2012). Enemies maintain hyperdiverse tropical forests. *The American Naturalist*, *179*(3), 303-314.
- Tello, J. S., Myers, J. A., Macía, M. J., Fuentes, A. F., Cayola, L., Arellano, G., ... & Jørgensen, P. M. (2015). Elevational gradients in β -diversity reflect variation in the strength of local community assembly mechanisms across spatial scales. *PLoS One*, *10*(3), e0121458.
- Urban, M. C., Bocedi, G., Hendry, A. P., Mihoub, J. B., Pe'er, G., Singer, A., ... & Travis, J. M. (2016). Improving the forecast for biodiversity under climate change. *Science*, *353*(6304), aad8466.
- Uriarte, M., Swenson, N. G., Chazdon, R. L., Comita, L. S., John Kress, W., Erickson, D., ... & Thompson, J. (2010). Trait similarity, shared ancestry and the structure of neighbourhood interactions in a subtropical wet forest: implications for community assembly. *Ecology letters*, *13*(12), 1503-1514.
- Vellend, M. (2016). *The Theory of Ecological Communities*. Princeton University Press, Princeton, NJ.
- Volkov, I., Banavar, J. R., Hubbell, S. P., & Maritan, A. (2009). Inferring species interactions in tropical forests. *Proceedings of the National Academy of Sciences*, *106*(33), 13854-13859.
- Wallace, A. R. (1878). *Tropical nature, and other essays*. Macmillan and Company.

- Wang, W., Peng, C., Kneeshaw, D. D., Larocque, G. R., & Luo, Z. (2012). Drought-induced tree mortality: ecological consequences, causes, and modeling. *Environmental Reviews*, 20(2), 109-121.
- Wang, J. F., Zhang, T. L., & Fu, B. J. (2016). A measure of spatial stratified heterogeneity. *Ecological indicators*, 67, 250-256.
- Webb, C. O., Gilbert, G. S., & Donoghue, M. J. (2006). Phylodiversity-dependent seedling mortality, size structure, and disease in a Bornean rain forest. *Ecology*, 87(sp7), S123-S131.
- Webb, C. T., Hoeting, J. A., Ames, G. M., Pyne, M. I., & LeRoy Poff, N. (2010). A structured and dynamic framework to advance traits-based theory and prediction in ecology. *Ecology letters*, 13(3), 267-283.
- Weihner, E., Clarke, G. P., & Keddy, P. A. (1998). Community assembly rules, morphological dispersion, and the coexistence of plant species. *Oikos*, 309-322.
- Weihner, E., & Keddy, P. A. (1995). Assembly rules, null models, and trait dispersion: new questions from old patterns. *Oikos*, 159-164.
- Wiegand, T., Huth, A., Getzin, S., Wang, X., Hao, Z., Gunatilleke, C. S., & Gunatilleke, I. N. (2012). Testing the independent species' arrangement assertion made by theories of stochastic geometry of biodiversity. *Proceedings of the Royal Society B: Biological Sciences*, 279(1741), 3312-3320.
- Wiegand, T., & Moloney, K. A. (2013). *Handbook of spatial point-pattern analysis in ecology*. CRC press.
- Wiegand, T., Uriarte, M., Kraft, N. J., Shen, G., Wang, X., & He, F. (2017). Spatially explicit metrics of species diversity, functional diversity, and phylogenetic diversity: Insights into plant community assembly processes. *Annual Review of Ecology, Evolution, and Systematics*, 48, 329-351.

2.6 Figures

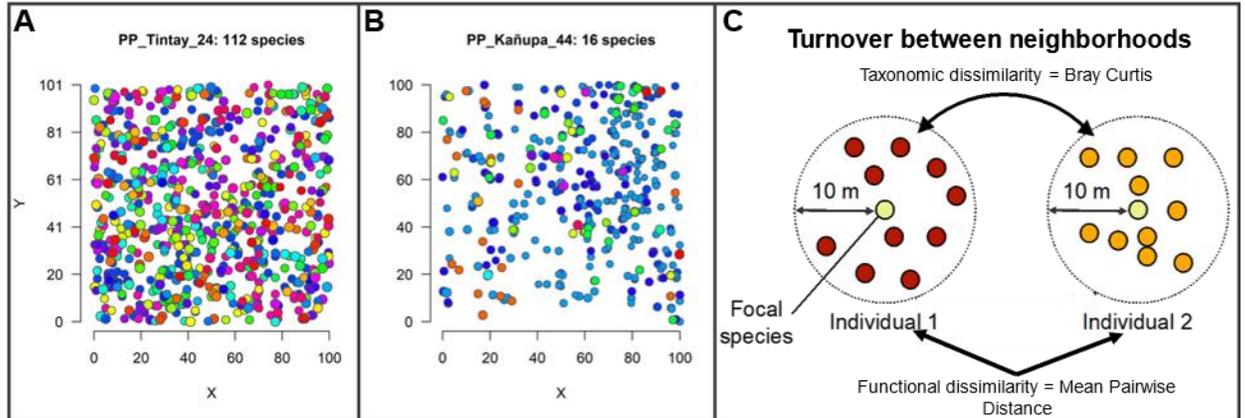


Figure 1 – Examples of two stem-mapped, 1-ha forest plots, showing spatial distributions of individual trees (circles) and tree species (circle colors). A) High-diversity forest plot containing 112 tree species. B) Low diversity plot containing 16 tree species. The x and y axes show distances in meters from the bottom-left corner of the plot. C) Illustration of methods used to calculate the taxonomic dissimilarity (Bray Curtis dissimilarity) and functional dissimilarity (Mean Pairwise Distance of one or more functional traits) between 10-m neighborhoods of two conspecific individuals of a focal tree species.

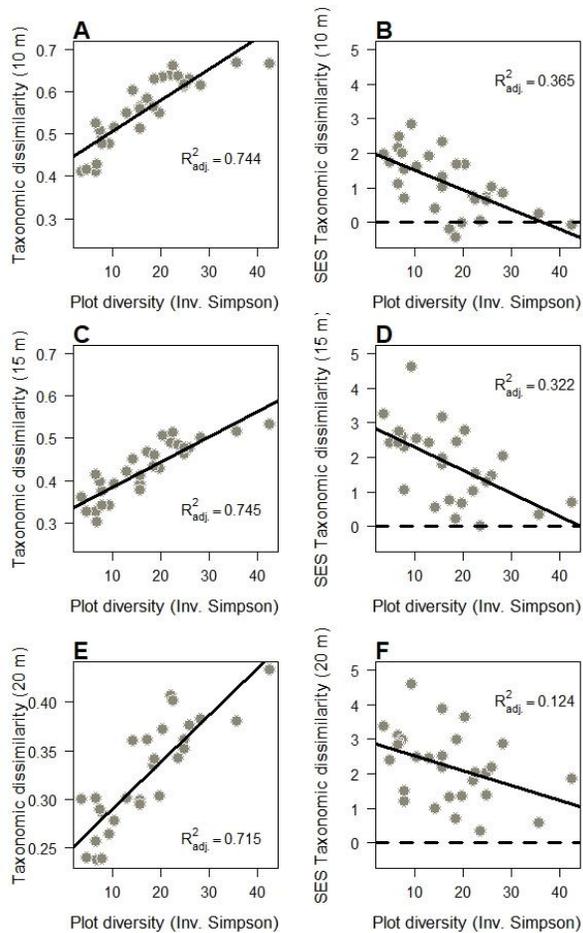


Figure 2 – Mean taxonomic dissimilarity of tree neighborhoods across the diversity gradient. The left panels show relationships between the mean observed taxonomic dissimilarity (Bray Curtis dissimilarity) of conspecific tree neighborhoods and species diversity (Inverse Simpson Index; $N = 31$ forest plots) at A) 10-m, C) 15-m, and E) 20-m neighborhood radii. The right panels show the same relationships for null-model deviations (mean standardized effects sizes, SES). The dashed line shows the null expectation ($SES = 0$); positive and negative SES indicate higher and lower taxonomic dissimilarity than expected from random sampling of individuals from the plot species pool, respectively. Solid trend lines indicate significant linear relationships ($P < 0.05$).

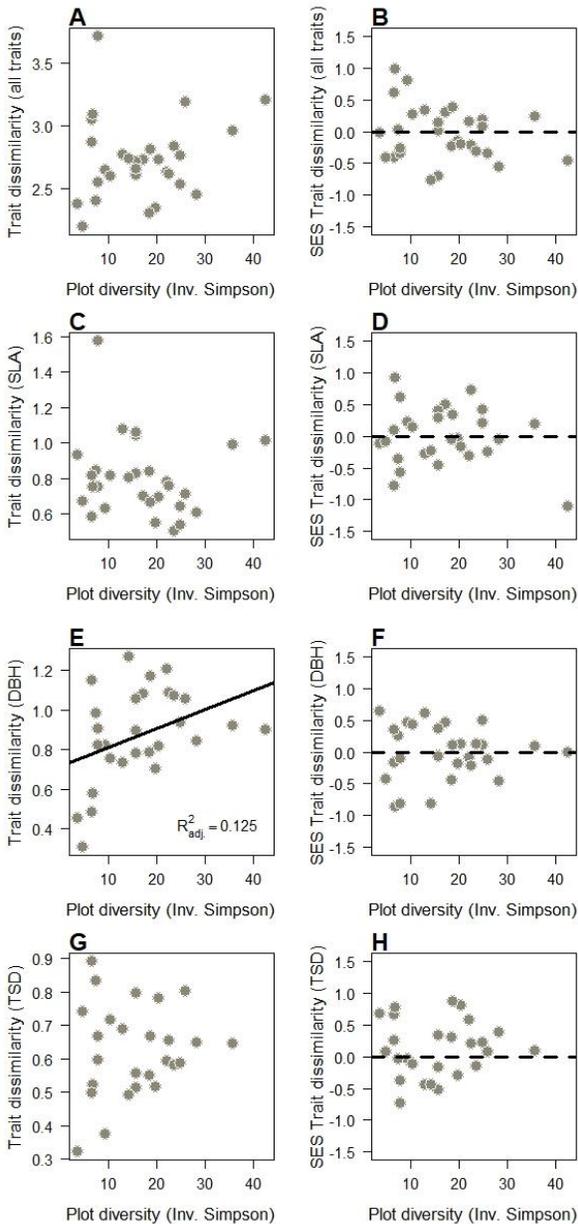


Figure 3 – Mean functional dissimilarity of tree neighborhoods across the diversity gradient. The left panels show the relationships between the mean observed functional dissimilarity of (trait mean pairwise distances) and species diversity (Inverse Simpson Index; $N = 29$ forest plots) at the at 15m neighborhood scale for A) all eight traits combined, and three selected individual traits including C) specific leaf area (SLA), E) maximum tree diameter at breast height (max. DBH), and G) Twig specific wood density. The right panels show the same relationships for null-model deviations (mean standardized effect sizes, SES). The dashed line shows the null

expectation ($SES = 0$); positive and negative SES indicate higher and lower functional dissimilarity than expected from random sampling of traits from the plot species pool, respectively. Solid trend lines indicate significant linear relationships ($P < 0.05$). Results were quantitatively similar for 10 and 20 m neighborhood sizes (Figures S2 & S3).

Chapter 3: An ecological metabolomics window on plant chemistry and diversity gradients: Chemical dissimilarity within tree communities increases with species diversity and decreases with elevation in the Central Andes

3.1 Introduction

Foundational hypotheses in ecology and evolution posit that stronger and more specialized biotic interactions contribute to large-scale gradients in biological diversity (Schemske et al. 2009). Wallace (1878) and Dobzhansky (1950) proposed that biotic interactions comprise a stronger selective force than the abiotic environment in the tropics. However, the mechanisms by which tropical forests may facilitate the ecological coexistence of hundreds to thousands of tree species remain unclear (Wright 2002). Unlike animals, which can exploit distinct resources, nearly all plants require light, water, CO₂, and a few shared micronutrients, so opportunities for resource-based niche differentiation are few (Hubbell 2001). In contrast to resource-based niche axes, the nearly infinite variety of insect herbivores and microbial pathogens provides a highly multidimensional space within which plant species can carve out a distinct niche defined by the enemies they support, and by those they avoid. Specialized natural enemies can maintain species-rich plant communities by attacking their host plants where they are abundant, impeding their fitness relative to competitors that avoid the enemy (Janzen 1970, Connell 1971, Bever et al. 2015). Hence, large-scale gradients in biodiversity may be attributed

to greater pressure from specialized herbivores and pathogens at lower elevations and latitudes with warmer, wetter, and less-seasonal climates (Schemske et al. 2009, Comita et al. 2014, Terborgh 2012, Levi et al. 2019). A major impediment in understanding these complex interactions has been the difficulty in the study of chemical ecology within communities and across large spatial and taxonomic scales.

Plant-secondary metabolites are organic molecules that do not function in the primary, resource-acquisitive metabolism but mediate plant responses to abiotic or biotic stress or function as defenses against the ability of herbivores or pathogens to identify or digest the plant host, or through acute toxicity. The natural enemies of the plant host do evolve counters to these defenses, though often at the cost of generality (Schemske et al. 2009).

Plant-chemical defenses mediate biotic interactions, defining the host-use relationships between plants and their natural enemies (Becerra 1997, Kursar et al. 2009, Salazar et al. 2016). In response, herbivores and pathogens can evolve counters to plant chemical defenses, but often at the cost of generality (Ehrlich & Raven 1964). Plant-enemy coevolution can result in host-use patterns that track plant secondary metabolites, promote chemical diversity and species richness in plant communities (Sedio and Ostling 2013), and mediate selection for chemical divergence among closely related plants (Becerra 1997, Kursar et al. 2009, Endara et al. 2017, Salazar et al. 2016). Biotic interactions are presumed to exert greater selective pressure in more stable, benign climates that allow for a greater range of adaptations (Dobzhansky 1950, Fischer 1960), or allowed an escape from glaciation on larger timelines (Wallace 1878, Schemske et al. 2009). Plant-secondary metabolites are organic molecules that do not function in the primary, resource-acquisitive metabolism, and have been shown to be more evolutionarily labile than other traits, where even closely related species can have very different metabolomes (Wink 2003). At

evolutionary timescales this evolutionary lability is expected to result in patterns of less phylogenetic conservation of secondary metabolites among species found in warmer, wetter, more stable climates, where biotic interactions are presumably stronger. Along an elevational gradient, the abundance of herbivores and pathogens tends to decrease with elevation (Pellissier et al., 2014; Sam et al., 2020), while abiotic stress tends to increase with elevation. This can result in a tradeoff (Coley 1985) where high-elevation plants may be expected to invest more in chemical defenses because compensatory regrowth of biomass lost to natural enemies under unfavorable abiotic conditions and low nutrients is relatively more costly to high-elevation plants (Defosse et al., 2018; Salgado et al. 2016). Furthermore, abiotic stress itself may select for investment in, and optimization of, specialized secondary metabolites that mediate plant stress response or protect against damage, such as from ultraviolet light (Volf et al. 2020). Yet unlike plant-enemy interactions that may undergo reciprocal coevolution, abiotic stress ought to select for convergence on shared, optimal traits (Asplund et al. 2022; Bakhtiari et al., 2021). At extremes, abiotic stress may even result in the collapse of some metabolic pathways (Pellissier et al. 2014). On the other hand, high elevations may select for unique metabolites not found in lowland plants (Defosse et al. 2021). Perhaps because of such discordant selection, several studies have found non-linear, hump-shaped relationships between herbivory, plant secondary metabolite dissimilarity and elevational gradients (Sam et al., 2020; Volf et al., 2020).

Despite the importance of plant chemistry in mediating community dynamics, key gaps remain in our understanding of the chemical ecology of plant communities across elevational gradients. First, although a few studies have examined chemical variation over elevational gradients in herbaceous grassland communities (Defosse et al., 2018, 2021), previous studies of woody plants have focused on single genera (Sam et al. 2019, Volf et al. 2020, 2022). Second,

the role that plant secondary metabolites play in generating biodiversity patterns has until recently been limited by their overwhelming diversity and the lack of untargeted approaches to study them at macroecological scales. Here, we overcome this obstacle by taking advantage of recent innovations in untargeted metabolomics based on mass spectrometry (Wang et al. 2016, Dührkop et al. 2019) that enable the study of chemical ecology at the scale of species-rich ecological communities (Sedio 2017, Sedio et al. 2018).

In this study we explored the hypothesis that stronger selection by natural enemies at lower elevations shape gradients in the diversity and evolution of plant secondary metabolites. We utilized data from 16 1-ha forest plots distributed across an ~3000-m elevational-diversity gradient in the Bolivian Andes (17-137 tree species per plot). Using recent advances in large-scale chemical-metabolomic analytical techniques (Wang et al. 2016, Sedio 2017), we compared patterns of primary and secondary foliar metabolites in 473 species (906 unique species-plot combinations) to plot diversity, elevation, climate, and phylogeny across the gradient to test four predictions: 1) Interspecific differences in plant-secondary metabolites will increase with species diversity; 2) Interspecific differences in plant-secondary metabolites will increase towards warmer, wetter, less-seasonal climates; 3) Plant species exhibit faster evolution of secondary metabolites (i.e., less phylogenetic signal) in high diversity communities; and 4) Plant species exhibit faster evolution of secondary metabolites (i.e., less phylogenetic signal) in warmer, wetter, and less-seasonal locations. Evidence in favor of these predictions would lend support to the hypothesis that variation in the strength of selection for interspecific divergence in secondary metabolites associated with climatic gradients contributes to the widespread elevational diversity gradient in trees.

3.2 Methods

Floristic Data

Floristic data were collected as part of the Madidi project (Jørgensen et al. 2015), a collaboration of more than two decades between the Herbario Nacional de Bolivia and the Missouri Botanical Garden to document the flora of the Madidi region in the Andes of Bolivia (Tello et al. 2015). The region ranges in elevation from lowland rainforests located at around 200 m above sea level (a.s.l.) to high mountains above 6,000 m a.s.l., above the tree line (Fuentes 2005). The elevational gradient is covered by different forest types and encompasses a broad range of abiotic (climatic and environmental) conditions (Rafiqpoor & Ibish 2004, Friedman-Rudovsky 2012). The Madidi Project includes a total of 50 1-ha permanent plots ranging in elevation from 212 m to 3334 m above sea level. For this study, we selected a subset of 16 1-ha permanent plots in which leaves were sampled for chemical analyses and which spanned gradients in elevation (662-3324 m above sea level), climate, and tree-species richness (17-137 species per plot) (Table 1). The 16 plots include three seasonally dry, low-elevation forest plots, and 13 moist, montane forest plots (Table 1). Some of the most abundant genera in the low elevation moist plots include: *Miconia*, *Sloanea*, *Ocotea* and at high elevations (> 2500m): *Weinmannia*, *Hedyosmum*, *Clethra*, and in the seasonally dry plots: *Weinmannia*, *Hedyosmum*, *Clethra*. The species richness exhibits the typical negative relationship with elevation among the 13 moist forest plots, but that the 3 seasonally dry forest exhibit a unique pattern of low species diversity at low elevations. Within each plot, all woody plants (hereafter trees) with a diameter at breast height of at least 10 cm were spatially mapped, measured, and identified to a valid species

or morphospecies. The majority of these plots have been censused a minimum of two times. The 16 plots represented a total of 473 species, and 906 unique species-by-plot.

Chemical Analysis

Analytical chemistry and untargeted metabolomics

Leaves of up to five individual trees per species per plot, that were collected after 2010, and dried on silica gel were used for metabolite sampling. When a species had fewer than 5 individuals in a plot, we sampled leaves from all individuals. These samples were extracted and analyzed following Sedio et al. (2021). Briefly, 50 mg of dry leaf tissue was ground to a fine powder and 10 mg weighed for extraction in 1800 μ L 90:10 methanol:water at pH 5 overnight at 4 °C. We used this solvent to extract metabolites over a wide range of polarity; mild acidity improves the solubility of most alkaloids, an important class of defensive secondary metabolites (Sedio et al. 2017). Extracts of five individuals per species per plot were pooled to create 906 extract pools representing each unique species-by-plot for subsequent analysis.

We analyzed filtered extract pools using ultra-high performance liquid chromatography-tandem mass spectrometry (UHPLC-MS/MS) using a Thermo Fisher Scientific (Waltham, MA, USA) Vanquish UHPLC with a C18 column and a Thermo QExactive quadrupole-orbitrap MS. Separation of metabolites by UHPLC was followed by heated electrospray ionization (HESI) in positive mode using full scan MS1 and data-dependent acquisition of MS2. Detailed instrumental methods are described by Sedio et al. (2021). Spectra for all 906 species-by-plot is curated as a public MassIVE dataset on the Global Natural Products Social (GNPS) Molecular Networking server (<ftp://massive.ucsd.edu/MSV000090549>)

Raw LC-MS data were centroided and processed for peak detection, peak alignment, and filtering using MZmine2 (Pluskal et al. 2010). Aligned chromatograms were used to create a ‘feature-based molecular network’ (FBMN; Nothias et al. 2020) using GNPS (Wang et al. 2016). The resulting network was used to create a dendrogram in which the structural similarities of all metabolites were reflected in one phylogeny-like dendrogram using Qemistree (Tripathi et al. 2021). Metabolites were annotated by predicting molecular formulae using Sirius (Dührkop et al. 2015), predicting molecular structures using CSI:FingerID (Dührkop et al. 2019), and classifying compounds chemically using ClassyFire (Djoumbou Feunang et al. 2016) and NPCClassifier (Kim et al. 2022).

NPCClassifier is a deep-learning artificial intelligence that classifies metabolites according to basic biosynthetic pathways, in addition to superclasses and classes (Kim et al. 2022). We used the “pathway”-level classifications of NPCClassifier to group metabolites into “primary” and “secondary” metabolites as follows: primary metabolites were those classified in the “Carbohydrates” and “Fatty acids” pathways, whereas secondary metabolites were those classified in the “Alkaloids”, “Amino acids and Peptides”, “Polyketides”, “Shikimates and Phenylpropanoids”, and “Terpenoids” as well as compounds likely to be products of multiple core biosynthetic pathways that included one of these secondary-metabolite pathways. Glycosides were classified based on their non-carbohydrate moieties; nucleotides and nucleosides were classified among the carbohydrates at the “pathway” level (Kim et al. 2022). Our classification scheme was based on the broad likelihood of a metabolite being associated with anti-herbivore or antimicrobial defense. For example, amino acids and peptides include many primary metabolites, but may also include defensive compounds.

Chemical structural and compositional similarity (CSCS)

Sedio et al. (2017) developed a metric that quantifies chemical structural-compositional similarity (CSCS) over all compounds among species pairs. Conventional distance or similarity indices such as Bray-Curtis incorporate shared compounds but ignore structural similarity of unique compounds, and hence underestimate the similarity of species with distinct but very structurally similar, and perhaps functionally redundant, metabolites (Sedio et al. 2017). For each pair of the 906 species-by-plot, we calculated CSCS using i) the whole metabolome, using all metabolites in the data, ii) primary metabolites, and iii) secondary metabolites. We transformed CSCS matrices into dissimilarity matrices by calculating 1-CSCS. We calculated the abundance-weighted median 1-CSCS for the species assemblages represented by each of the 16 forest plots.

To disentangle the community chemical dissimilarity from the effect of diversity *per se*, we carried out rarefaction based on 12 species, the [sampled] species richness of Kañupa, the most species-poor plot. We calculated rarefied CSCS by taking a random sample of 12 species and calculating the median chemical dissimilarity values at the plot-level. This operation was performed 1000 times for each plot and the mean of the distribution was taken as the rarefied median chemical dissimilarity value. We obtained qualitatively similar results using observed and rarefied CSCS values. For simplicity, we therefore focus on results for observed CSCS and include results for rarefied CSCSS in the supplementary material.

Climate Data

Climate variables were selected to represent the variation in plot temperature, precipitation, and seasonality over the elevational gradient. The temperature variables annual mean temperature and temperature annual range were used from WorldClim Version 2.1 (Fick &

Hijmans 2017). The precipitation variables annual precipitation and precipitation seasonality were used from the Tropical Rainfall Measuring Mission (TRMM) because of the greater accuracy compared to WorldClim data in the Madidi region. The 4 variables were scaled and centered, and a principal components analysis was run, of which the first two axes were used in the following analyses.

Phylogenetic Signal

In order to quantify phylogenetic relationships among species, a phylogenetic tree was constructed using the V.Phylomaker package (Jin & Qian 2019) in R. As inputs, V.Phylomaker requires family, genus and species information, which is then referenced against two combined mega trees (Zanne et al. 2014, Smith & Brown 2018) to generate the phylogenetic tree. The resulting tree was generated from all 50 of the Madidi permanent plots and had 1123 unique species as tips. The tree was then rooted and transformed into a distance matrix using the ‘cophenetic’ function, in the stats package in R, in order to be directly comparable to the chemical distance matrices. The tree was pruned to include only the species recorded in the 16 plots for all phylogenetic signal analyses, which included 892 species-by-plot.

For each plot, we calculated Adams’ (2014) K_{mult} metric of phylogenetic signal for multivariate trait data. This technique compares an explicit model of evolution in multivariate trait space to the observed trait data, accounting for the topology and branch lengths of the phylogeny. When $K_{mult} < 1$, taxa are less chemically similar to one another than expected by Brownian motion evolution on the observed phylogeny, whereas $K_{mult} > 1$ indicates that species are more chemically similar to each other than expected by Brownian motion. The K_{mult} test is an improvement over the Mantel test for matrix correlation between chemical distance and

phylogenetic distance matrices, which does not consider an explicit model of trait evolution underlying the expected relationship between phylogenetic and trait distance (Swenson 2014).

Hypothesis Testing: Plot-Level Regressions

We tested our predictions using linear regression. For each plot, we calculated tree species diversity as the inverse Simpson's index using the R *vegan* package (Hurlbert 1971). We chose the inverse Simpson's index because it is scale-independent and is insensitive to differences in numbers of individuals (Chase et al., 2018). To test Prediction 1 (P1), we regressed median chemical dissimilarity versus diversity (inverse Simpson's index) for the 16 forest plots. To test P2, we regressed median chemical dissimilarity versus elevation and the first two principal component (PC) axes of climatic variation. To test P3, we regressed K_{mult} versus diversity, and to test P4, we regressed K_{mult} versus elevation and the first two PC axes of climatic variation among the 16 forest plots. The three seasonally dry, low-elevation forest plots appeared to exhibit distinct relationships to other variables not represented by the 13 moist, montane forest plots. Hence, all regressions were repeated with the 13 moist montane forests, excluding the seasonally dry forests.

3.3 Results

Overview of metabolomics and climate data

We detected 22,576 unique metabolites in foliar extracts from 473 species in 16 1-ha forest plots (Figure 1). Metabolites ranged in mass from 116.0704 to 599.4789 Daltons (Da). We

generated a predicted molecular structure and biosynthetic classification for 18,364 and 19,844 of the metabolites, respectively. Metabolites classified at the level of biosynthetic pathway of origin (“pathway” in NPClassifier, Kim et al. 2022) were represented by 4,448 Alkaloids, 458 Amino acids and Peptides, 262 Carbohydrates, 928 Fatty acids, 584 Polyketides, 6,131 Shikimates and Phenylpropanoids, 6,871 Terpenoids, and 153 metabolites derived from more than one major pathway.

The first two principal components of climatic variation represented 71.2% and 15.8% of the variation among the 16 forest plots, respectively. The first principal component of climatic variation (PC1) was composed of temperature and precipitation, whereas the second principal component of climatic variation (PC2) was primarily composed of precipitation seasonality and temperature annual range (Figure 2).

Chemical dissimilarity and phylogenetic signal across gradients

Chemical dissimilarity with respect to all detected metabolites increased with species diversity ($R^2 = 0.40$, $p < 0.01$), decreased with elevation ($R^2 = 0.30$, $p = 0.02$), and increased along Climate PC1 ($R^2 = 0.46$, $p < 0.01$) but was unrelated to Climate PC2 among the 16 forest plots (Figure 3a-d). The relationship between chemical dissimilarity and elevation was much stronger when three seasonally dry, low-elevation forest plots were excluded (Figure 3b, dashed line excluding open circles, $R^2 = 0.54$, $p < 0.01$).

Chemical dissimilarity with respect to secondary metabolites increased with species diversity (Figure 3e; $R^2 = 0.43$, $p < 0.01$) and Climate PC1 (Figure 3g; $R^2 = 0.35$, $p < 0.01$) among the 16 forest plots. This measure decreased with elevation when three lowland dry forests were excluded (Figure 3f; $R^2 = 0.45$, $p < 0.01$). In contrast, plot median chemical dissimilarity

with respect to primary metabolites decreased with climate PC2 (Figure 3l; $R^2 = 0.25$, $p = 0.03$) but was unrelated to diversity, elevation, or climate PC1 (Figure 3i-k).

Phylogenetic signal was exceedingly low for all, secondary, and primary metabolites, as none of the plots approached the Brownian motion expectation ($K_{\text{mult}} = 1$) for any of the three metabolite classes (Table 1). Phylogenetic signal appeared greatest for low-elevation, low-species diversity dry forests and the highest-elevation, low-species diversity moist montane forests in the gradient (Table 1).

Phylogenetic signal with respect to all detected metabolites decreased with species diversity (Figure 4a; $R^2 = 0.27$, $p = 0.2$) but was unrelated to elevation or Climate PC1 or PC2 (Figure 4b-d). Phylogenetic signal with respect to secondary metabolites decreased with species diversity (Figure 4e; $R^2 = 0.35$, $p = 0.01$). When the three seasonally dry, low-elevation dry forest plots were excluded, phylogenetic signal with respect to secondary metabolites increased with elevation (Figure 4f; $R^2 = 0.26$, $p = 0.04$) and decreased with Climate PC1 (Figure 4g; $R^2 = 0.24$, $p = 0.05$). Phylogenetic signal with respect to primary metabolites decreased with species diversity (Figure 4i; $R^2 = 0.28$, $p = 0.02$) but was unrelated to elevation or Climate PC1 or PC2 (Figure 4j-l).

3.4 Discussion

Elevational diversity gradients are a striking feature of our planet and have inspired the development of ideas in ecology for centuries (von Humboldt & Bonpland 1807, Lomolino 2001, Rahbek 2005). The 16 tropical forest plots we examined here represented a wide range of variation in elevation, species diversity, and climate within a regional biodiversity hotspot in the

central Andes Mountains. Hypotheses propose that biodiversity gradients are driven by variation in the relative strength and nature of selection imposed by the abiotic and biotic environment (Wallace 1878, Dobzhansky 1950, Schemske et al. 2008, Lim et al. 2015), which are predicted to be reflected in the interspecific dissimilarity and phylogenetic signal of secondary-metabolite profiles among co-occurring tree species. Our results broadly support four specific predictions concerning the relationships between chemical dissimilarity and phylogenetic signal and underlying gradients in species diversity, elevation, and climate, which we discuss below.

High-diversity communities are composed of species with more dissimilar secondary metabolites

In Prediction 1, we predicted a positive relationship between chemical dissimilarity and community species diversity, based on the hypothesis that diversity itself is increased by antagonistic biotic interactions that select for chemical divergence among species (Dobzhansky 1950, Ehrlich and Raven 1964), reduce natural-enemy overlap among species (Becerra 1997, Endara et al. 2017), and promote competitive coexistence (Sedio and Ostling 2013). Our results supported Prediction 1. We found that chemical dissimilarity increased with species diversity (Figure 3a,e), whether measured in terms of the whole metabolome or metabolites we classified as secondary metabolites based on their biosynthetic pathways of origin as predicted by NPClassifier (Kim et al. 2022). These patterns emerged in models based on median chemical dissimilarity of all co-occurring species at the forest-plot scale, and models based on rarified subsamples of co-occurring species that served to control for variation in species richness (Figure S1). These results are consistent with hypotheses that attribute community-scale variation in species diversity to variation in the strength of mechanisms that promote chemical

divergence among species, such as pressure from relatively host-specific, but oligophagous herbivores and pathogens (Schemske et al. 2009, Lim et al. 2015). Furthermore, dissimilarity with respect to secondary metabolites strongly increased with species diversity (Figure 3e), while dissimilarity with respect to primary metabolites did not (Figure 3i). This contrast suggests that species differences in secondary metabolites, which include alkaloids, phenolics, polyketides, terpenoids (including steroid, or cardiac, glycosides), and non-protein amino acids that function as anti-herbivore and/or antimicrobial defenses, contribute to the diversity gradient.

Communities in warmer, wetter, and less-seasonal climates are composed of species with more dissimilar secondary metabolites

In Prediction 2, we predicted that chemical dissimilarity of co-occurring species would decrease with elevation and increase in warmer, wetter, and less-seasonal climates, based on the hypothesis that chemically mediated plant-enemy coevolution that selects for chemical divergence among plants plays a greater role in these abiotically benign climates (Wallace 1878, Dobzhansky 1950). Our results supported Prediction 2. We found that chemical dissimilarity decreased with elevation and increased with temperature and precipitation as reflected in climatic PC1. For the metabolome considered as a whole, these patterns were true whether three lowland tropical dry forests were included or not (Figure 3b,c). For secondary metabolites, the positive relationship between chemical dissimilarity and PC1 was significant (Figure 3 g), whereas the negative relationship between chemical dissimilarity and elevation was significant only if three lowland tropical forests were excluded (Figure 3f).

Recent investigations of metabolomic variation associated with elevational gradients have considered single genera of woody plants, such as *Ficus* in Papua New Guinea (Volf et al.

2020) and *Salix* in Europe (Volf et al. 2022, in press), and communities of herbaceous plants in Europe (Defosse et al. 2018, 2021). However, few of these studies have explicitly quantified chemical dissimilarity of co-occurring species at points along an elevational gradient. Volf et al. (2022, in press) found that low-elevation *Salix* were more dissimilar with respect to salicinoids, an important class of phenolic chemical defenses, than high-elevation willows, a result consistent with ours.

Studies of chemical dissimilarity of co-occurring species have more frequently asked whether communities exhibit chemical overdispersion, wherein species that co-occur at the local scale are less chemically similar than expected by chance given the regional species pool. This result has been reported for species-rich tree and shrub genera in the lowland Neotropics, including *Bursera* (Burseraceae) in Mexico (Becerra 2007), *Inga* (Fabaceae) in Panama and Peru (Kursar et al. 2009), *Piper* in Costa Rica (Salazar et al. 2016), and *Protium* (Burseraceae) in Peru (Vleminckx et al. 2018). Similar patterns have been found in *Ficus* in Papua New Guinea (Volf et al. 2018), Euphorbiaceae (principally *Macaranga*) in Yunnan, China (Wang et al. 2022) and in an assessment of seven species-rich genera in Panama (Sedio et al. 2017). In addition to studies focused on single lineages, community-wide studies have found that plants that co-occur within meters are more chemically dissimilar than expected from a community-wide sample (Wang et al. in rev) and the chemical dissimilarity of co-occurring species tends to decrease with latitude in the temperate and tropical zones (Sedio et al. 2018) and with temperature and precipitation within the boreal and temperate zones (Sedio et al. 2021).

Our study contrasts with these previous studies focused on single lineages and comparative studies focused on variation among whole plant communities at continental scales (Sedio et al. 2018, 2021) in that we focused on variation along a local elevational gradient within

the same biogeographic region (Central Andes). Nevertheless, our results are consistent with the widely observed chemical dissimilarity of tree species in low-elevation tropical forests. However, it is worth noting that Sedio et al. (2018, 2021) found differences in chemical dissimilarity among geographically distant plant communities with very different biogeographic histories and little possibility for dispersal over ecological timescales. Our findings suggest that variation in temperature, precipitation, and seasonality over distances of kilometers may generate variation in community chemical dissimilarity comparable to that of climatic gradients on a continental scale, and hence that the underlying mechanisms that link climate to chemical evolution and competitive coexistence are likely general and operate over a wide range of variation in spatial scale.

In stark contrast with our results concerning the whole metabolome or secondary metabolites, we observed a positive relationship between chemical dissimilarity with respect to primary metabolites and Climate PC2 (Figure 31). As Climate PC2 was primarily defined by temperature range and precipitation seasonality, this result suggests that species that occur in more seasonal climates are less dissimilar chemically than species in climates that are stable throughout the year.

Chemical divergence among closely related species is greater in high-diversity communities, and in warmer, wetter, and less-seasonal climates

Phylogenetic signal was much lower than expected based on a model of Brownian motion drift without selection, even in high-elevation plots with comparatively higher phylogenetic signal than wetter, low-elevation plots (Table 1). Plants do exhibit phylogenetic signal with respect to broad chemical classes that tend to occur in certain plant families or genera

(quinolizidine alkaloids in some lineages of legumes, Wink 2003). However, our result is consistent with other recent studies that have concluded that plant metabolite composition can be highly evolutionarily labile, especially in tropical climates when phylogenetic signal is measured among confamilial species (Becerra 1997, Kursar et al. 2009, Salazar et al. 2018, Volf et al. 2018, Wang et al. 2022) and this can degrade phylogenetic signal when measured in the context of a species-rich forest community characterized by many co-occurring congeneric and confamilial species (Sedio et al. 2018, 2021, Wang et al. in rev).

In Predictions 3 and 4, we predicted that chemical phylogenetic signal among co-occurring species would decrease with species diversity, increase with elevation, and decrease in warmer, wetter, and less-seasonal climates, based on the hypothesis that selection by natural enemies for chemical divergence among closely related species (Becerra 1997, Kursar et al. 2009) is relatively stronger in such environments. Our results supported Prediction 3 for secondary metabolites as well as both all and primary metabolites (Figure 4a,e,i). Prediction 4, that phylogenetic signal increases with elevation and decreases with temperature and precipitation reflected in Climate PC1, was supported only when we excluded the three seasonally dry, low-elevation forest plots (Figure 4 f,g). These and other results (Figure 3b,f) suggest that the climatic data we used in our PCA may not completely represent the climates experienced by the three seasonally dry forests. This may be because climate varies dramatically over a finer spatial scale in the topographically heterogeneous Andes and their foothills than that captured by the TRMM satellite, and/or the climatic variables we used did not reflect the source of abiotic stress experienced by the dry forests, such as, for example, if short periods of intense drought or strong interannual variation stresses plants in a manner that is not reflected in the TRMM precipitation seasonality variable. However, the relationship between phylogenetic

signal with respect to secondary metabolites and elevation and Climate PC1, respectively, when the three seasonally dry forests were excluded suggests that the rate of chemical divergence among tree species varies with temperature along a wet-forest elevational gradient. This result supports our Prediction 4, which follows from Dobzhansky's (1950) that biotic interactions are stronger forces of natural selection in warmer, wetter, and less-seasonal climates.

Conclusions

Biodiversity gradients are a striking feature of our planet. While latitudinal biodiversity gradients have generally attracted more attention from biologists (Schemske et al. 2009), elevational gradients present perhaps a better opportunity to test hypotheses regarding proposed mechanisms, as key climatic variables vary over short distances, permitting experimentation or comparative study within a single system in which interacting species could plausibly disperse. Hence, we suggest that future research should take advantage of elevational gradients to test basic hypotheses concerning the intensity of plant interactions with insect herbivores and microbial pathogens, their effects on rates of chemical evolution in tree communities, and their contribution to the maintenance of species diversity.

Our results support the hypothesis that chemically mediated species interactions shape elevational diversity gradients by imposing stronger selection for interspecific divergence in plant chemical defenses in warmer, wetter, and less seasonal climates. Abiotic stress associated with high elevations may select for secondary metabolite evolution distinct from that imposed by biotic stressors (Defosse et al. 2021, Volf et al. 2020, 2022), but competitive interactions among plants mediated by shared herbivores and pathogens are expected to select for chemical

divergence. Our results suggest that the strength of this mechanism varies with climate in a manner that affects character evolution and diversity in plant communities. Our study also illustrates the promise of ecological metabolomics in the study of biogeography, community ecology, and complex species interactions in high-diversity ecosystems.

Acknowledgements

We thank J. Dwenger for assistance in the lab and M. Volf for valuable discussion. We thank the Dirección General de Biodiversidad, the Bolivian Park Service (SERNAP), the Madidi National Park and local communities for permits, access and collaboration in Bolivia, where fieldwork was supported by the National Science Foundation (DEB 0101775, DEB 0743457, DEB 1836353). Additional financial support to the Madidi Project was provided by the Missouri Botanical Garden, the National Geographic Society (NGS 7754-04 and NGS 8047-06), The German Embassy in Bolivia, International Center for Advanced Renewable Energy and Sustainability (I-CARES) at Washington University in St. Louis, the Comunidad de Madrid (Spain), Consejo Superior de Investigaciones Científicas (Spain), Centro de Estudios de América Latina (Banco Santander and Universidad Autónoma de Madrid, Spain), and the Taylor and Davidson families. Financial support for this study was provided by a grant from the Living Earth Collaborative at Washington University in St. Louis and the National Science Foundation (DEB Award 1557094 to JAM, DEB Award 2240431 to JAM, DEB Award 2240430 to BES). DH was supported by a Catharine M. Lieneman Botany Fellowship at Washington University in St. Louis.

3.5 References

- Aide, T. M. (1993). Patterns of leaf development and herbivory in a tropical understory community. *Ecology*, *74*(2), 455-466.
- Becerra, J. X. (1997). Insects on plants: macroevolutionary chemical trends in host use. *Science*, *276*(5310), 253-256.
- Bever, J. D., Mangan, S. A., & Alexander, H. M. (2015). Maintenance of plant species diversity by pathogens. *Annual review of ecology, evolution, and systematics*, *46*, 305-325.
- Cavender-Bares, J., Keen, A., & Miles, B. (2006). Phylogenetic structure of Floridian plant communities depends on taxonomic and spatial scale. *Ecology*, *87*(sp7), S109-S122.
- Close, D. C., & McArthur, C. (2002). Rethinking the role of many plant phenolics—protection from photodamage not herbivores?. *Oikos*, *99*(1), 166-172.
- Coley, P. D. (1991). Comparison of herbivory and plant defenses in temperate and tropical broad-leaved forests. *Plant-animal interactions: evolutionary ecology in tropical and temperate regions*, 25-49.
- Comita, L. S., Queenborough, S. A., Murphy, S. J., Eck, J. L., Xu, K., Krishnadas, M., ... & Zhu, Y. (2014). Testing predictions of the Janzen–Connell hypothesis: a meta-analysis of experimental evidence for distance- and density-dependent seed and seedling survival. *Journal of Ecology*, *102*(4), 845-856.
- Connell, J. H. (1971). On the role of natural enemies in preventing competitive exclusion in some marine animals and in rain forest trees. *Dynamics of populations*, *298*(312).
- Dobzhansky, T. (1950). Evolution in the tropics. *American scientist*, *38*(2), 209-221.
- Dührkop, K., Fleischauer, M., Ludwig, M., Aksenov, A. A., Melnik, A. V., Meusel, M., ... & Böcker, S. (2019). SIRIUS 4: a rapid tool for turning tandem mass spectra into metabolite structure information. *Nature methods*, *16*(4), 299-302.
- Ehrlich, P. R., & Raven, P. H. (1964). Butterflies and plants: a study in coevolution. *Evolution*, 586-608.
- Fick, S. E., & Hijmans, R. J. (2017). WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. *International journal of climatology*, *37*(12), 4302-4315.
- Fischer, A. G. (1960). Latitudinal variations in organic diversity. *Evolution*, *14*(1), 64-81.
- Forrister, D. L., Endara, M. J., Younkin, G. C., Coley, P. D., & Kursar, T. A. (2019). Herbivores as drivers of negative density dependence in tropical forest saplings. *Science*, *363*(6432), 1213-1216.
- Friedman-Rudovsky, J. (2012). Taking the measure of Madidi.
- Fuentes, A. (2005). Una introducción a la vegetación de la región de Madidi. *Ecología en Bolivia*, *40*(3), 1-31.

- Graham, C. H., Carnaval, A. C., Cadena, C. D., Zamudio, K. R., Roberts, T. E., Parra, J. L., ... & Sanders, N. J. (2014). The origin and maintenance of montane diversity: integrating evolutionary and ecological processes. *Ecography*, *37*(8), 711-719.
- Hurlbert, S.H. (1971). The nonconcept of species diversity: a critique and alternative parameters. *Ecology* **52**, 577–586.
- Janzen, D. H. (1967). Why mountain passes are higher in the tropics. *The American Naturalist*, *101*(919), 233-249.
- Janzen, D. H. (1970). Herbivores and the number of tree species in tropical forests. *The American Naturalist*, *104*(940), 501-528.
- Janzen, D. H. (1980). When is it coevolution?.
- Jin, Y., & Qian, H. (2019). V. PhyloMaker: an R package that can generate very large phylogenies for vascular plants. *Ecography*, *42*(8), 1353-1359.
- Jørgensen, P. M., Fuentes, A. F., Miranda, T., & Cayola, L. (2015). Manual de Trabajo.
- Kerr, J. T., Kharouba, H. M., & Currie, D. J. (2007). The macroecological contribution to global change solutions. *science*, *316*(5831), 1581-1584.
- Kim, S., Lim, S. W., & Choi, J. (2022). Drug discovery inspired by bioactive small molecules from nature. *Animal Cells and Systems*, 1-12.
- Kursar, T. A., Dexter, K. G., Lokvam, J., Pennington, R. T., Richardson, J. E., Weber, M. G., ... & Coley, P. D. (2009). The evolution of antiherbivore defenses and their contribution to species coexistence in the tropical tree genus *Inga*. *Proceedings of the National Academy of Sciences*, *106*(43), 18073-18078.
- Levi, T., Barfield, M., Barrantes, S., Sullivan, C., Holt, R. D., & Terborgh, J. (2019). Tropical forests can maintain hyperdiversity because of enemies. *Proceedings of the National Academy of Sciences*, *116*(2), 581-586.
- Lomolino, M. V. (2001). Elevation gradients of species-density: historical and prospective views. *Global Ecology and biogeography*, *10*(1), 3-13.
- Nogués-Bravo, D., Araújo, M. B., Romdal, T., & Rahbek, C. (2008). Scale effects and human impact on the elevational species richness gradients. *Nature*, *453*(7192), 216-219.
- Pellissier, L., Roger, A., Bilat, J., & Rasmann, S. (2014). High elevation *Plantago lanceolata* plants are less resistant to herbivory than their low elevation conspecifics: is it just temperature?. *Ecography*, *37*(10), 950-959.
- Qian, H., Chen, S., Mao, L., & Ouyang, Z. (2013). Drivers of β -diversity along latitudinal gradients revisited. *Global Ecology and Biogeography*, *22*(6), 659-670.
- Rahbek, C. (2005). The role of spatial scale and the perception of large-scale species-richness patterns. *Ecology letters*, *8*(2), 224-239.
- Rafiqpoor, D., & Ibsch, P. (2004). The physical environment of Bolivia. *Biodiversity: The richness of Bolivia. State of knowledge and conservation. Editorial Fundación Amigos de la Naturaleza, Santa Cruz de la Sierra*, 4-45.
- Salazar, D., Jaramillo, A., & Marquis, R. J. (2016). The impact of plant chemical diversity on plant–herbivore interactions at the community level. *Oecologia*, *181*, 1199-1208.
- Salgado, A. L., Suchan, T., Pellissier, L., Rasmann, S., Ducrest, A. L., & Alvarez, N. (2016). Differential phenotypic and genetic expression of defence compounds in a plant–herbivore interaction along elevation. *Royal Society open science*, *3*(9), 160226.

- Salminen, J. P., & Karonen, M. (2011). Chemical ecology of tannins and other phenolics: we need a change in approach. *Functional ecology*, 25(2), 325-338.
- Sam, K., Koane, B., Sam, L., Mrazova, A., Segar, S., Volf, M., ... & Novotny, V. (2020). Insect herbivory and herbivores of *Ficus* species along a rain forest elevational gradient in Papua New Guinea. *Biotropica*, 52(2), 263-276.
- Sanders, N. J., & Rahbek, C. (2012). The patterns and causes of elevational diversity gradients. *Ecography*, 35(1), 1.
- Schemske, D. W., Mittelbach, G. G., Cornell, H. V., Sobel, J. M., & Roy, K. (2009). Is there a latitudinal gradient in the importance of biotic interactions?. *Annu. Rev. Ecol. Evol. Syst.*, 40, 245-269.
- Sedio, B. E., & Ostling, A. M. (2013). How specialised must natural enemies be to facilitate coexistence among plants?. *Ecology letters*, 16(8), 995-1003.
- Sedio, B. E. (2017). Recent breakthroughs in metabolomics promise to reveal the cryptic chemical traits that mediate plant community composition, character evolution and lineage diversification. *New Phytologist*, 214(3), 952-958.
- Sedio, B. E., Rojas Echeverri, J. C., Boya P, C. A., & Wright, S. J. (2017). Sources of variation in foliar secondary chemistry in a tropical forest tree community.
- Sedio, B. E., Parker, J. D., McMahon, S. M., & Wright, S. J. (2018). Comparative foliar metabolomics of a tropical and a temperate forest community.
- Sedio, B. E., Boya P, C. A., & Rojas Echeverri, J. C. (2018). A protocol for high-throughput, untargeted forest community metabolomics using mass spectrometry molecular networks. *Applications in Plant Sciences*, 6(3), e1033.
- Sedio, B. E., Spasojevic, M. J., Myers, J. A., Wright, S. J., Person, M. D., Chandrasekaran, H., ... & Vandermeer, J. H. (2021). Chemical similarity of co-occurring trees decreases with precipitation and temperature in North American forests. *Frontiers in Ecology and Evolution*, 9, 679638.
- Senior, J. K., Potts, B. M., Davies, N. W., Wooliver, R. C., Schweitzer, J. A., Bailey, J. K., & O'Reilly-Wapstra, J. M. (2016). Phylogeny explains variation in the root chemistry of *Eucalyptus* species. *Journal of Chemical Ecology*, 42, 1086-1097.
- Smith, S. A., & Brown, J. W. (2018). Constructing a broadly inclusive seed plant phylogeny. *American journal of botany*, 105(3), 302-314.
- Sundqvist, M. K., Sanders, N. J., & Wardle, D. A. (2013). Community and ecosystem responses to elevational gradients: processes, mechanisms, and insights for global change. *Annual review of ecology, evolution, and systematics*, 44, 261-280.
- Terborgh, J. (2012). Enemies maintain hyperdiverse tropical forests. *The American Naturalist*, 179(3), 303-314.
- Tello, J. S., Myers, J. A., Macía, M. J., Fuentes, A. F., Cayola, L., Arellano, G., ... & Jørgensen, P. M. (2015). Elevational gradients in β -diversity reflect variation in the strength of local community assembly mechanisms across spatial scales. *PLoS One*, 10(3), e0121458.
- Terborgh, J. (2012). Enemies maintain hyperdiverse tropical forests. *The American Naturalist*, 179(3), 303-314.
- Von Humboldt, A., & Bonpland, A. (2010). *Essay on the Geography of Plants*. University of Chicago Press.
- Wallace, A. R. (1878). *Tropical nature, and other essays*. Macmillan and Company.

Wang, M., Carver, J. J., Phelan, V. V., Sanchez, L. M., Garg, N., Peng, Y., ... & Bandeira, N. (2016). Sharing and community curation of mass spectrometry data with Global Natural Products Social Molecular Networking. *Nature biotechnology*, 34(8), 828-837.

Wink, M. (2003). Evolution of secondary metabolites from an ecological and molecular phylogenetic perspective. *Phytochemistry*, 64(1), 3-19.

Wright, J. S. (2002). Plant diversity in tropical forests: a review of mechanisms of species coexistence. *Oecologia*, 130, 1-14.

Zanne, A. E., Tank, D. C., Cornwell, W. K., Eastman, J. M., Smith, S. A., FitzJohn, R. G., ... & Beaulieu, J. M. (2014). Three keys to the radiation of angiosperms into freezing environments. *Nature*, 506(7486), 89-92.

3.6 Tables

Table 1 - Variation in tree species richness, elevation, mean annual precipitation (MAP), and mean annual temperature (MAT) among 16 1-ha forest plots in the Madidi Project, Bolivia. Dry forest plots *italicized*.

Plot Name	Species richness	Elevation (m)	MAP (mm)	MAT (°C)	N trees sampled	N Spp. Metabolomes
Chaqui 32	35	3116	975	11.9	85	22
Fuerte 27	80	1900	1350	18	236	72
Kanupa 44	17	3324	961	10.7	52	12
Lomaka 40	137	1242	1332	21	248	109
Lomasa 39	95	1054	1383	21.5	250	81
<i>Pintat 5</i>	48	880	1684	22	164	42
<i>Resina 12</i>	50	662	1858	23.1	141	45
Sumpul 34	103	1223	1500	20.4	245	79
Tapuri 45	43	2697	1013	15.2	121	35
Tintay 24	117	1400	1358	19.7	283	87
Tintay 25	94	1468	1357	20	253	73
Titiri 42	34	2859	945	13.1	106	29
Tocoaq 29	83	2407	1202	15.6	170	71
Tocoaq 30	62	2510	1221	15.7	133	48
Tocoaq 28	81	2200	1288	16.8	209	67
<i>Yarimi 9</i>	40	850	1698	22.2	129	33

Table 2 – Phylogenetic signal (Kmult) for all metabolites, secondary metabolites, and primary metabolites for each forest plot. Dry forest plots *italicized*.

Plot Name	Kmult total metabolites	Kmult defensive metabolites	Kmult primary metabolites
Chaqui 32	0.480754529	0.484029572	0.411885738
Fuerte 27	0.053069475	0.051800002	0.097807575
Kanupa 44	0.160817612	0.202938517	0.189504321
Lomaka 40	0.080496417	0.084059673	0.195930742
Lomasa 39	0.04575673	0.045695698	0.092983475
<i>Pintat 5</i>	0.485211584	0.480105384	0.513034931
<i>Resina 12</i>	0.317434922	0.317895271	0.342970918
Sumpul 34	0.082048672	0.081485901	0.142247847
Tapuri 45	0.046396496	0.045138257	0.130433359
Tintay 24	0.055756469	0.05618906	0.073947205
Tintay 25	0.071480487	0.070726057	0.071011916
Titiri 42	0.103655099	0.1031501	0.290799712
Tocoaq 29	0.220003432	0.038435117	0.014061188
Tocoaq 30	0.045213988	0.044291159	0.092308025
Tocoaq 28	0.214260359	0.217448637	0.092308025
<i>Yarimi 9</i>	0.591126666	0.596744641	0.573387071

3.7 Figures

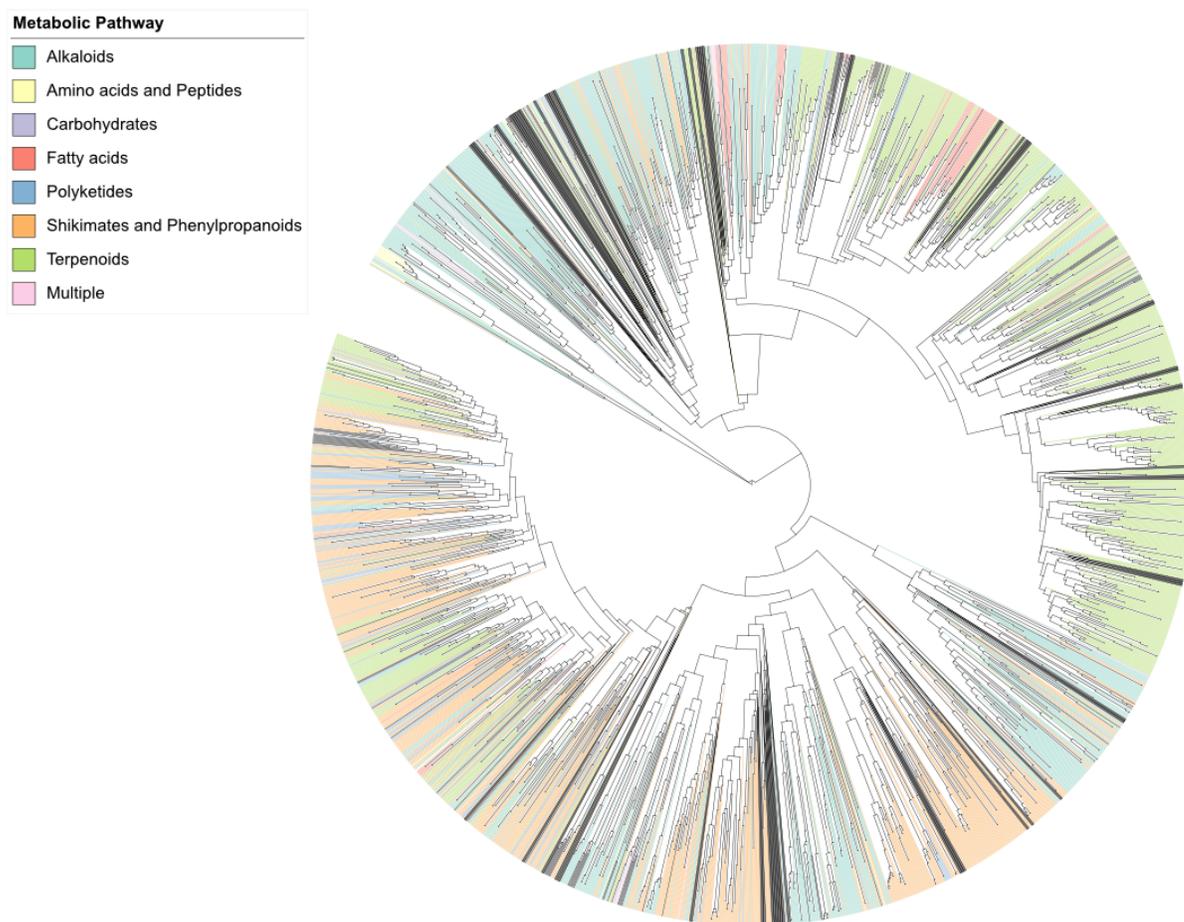


Figure 1 - Qemistree dendrogram showing relationships of the 18,364 classified metabolites found within the 473 species sampled.

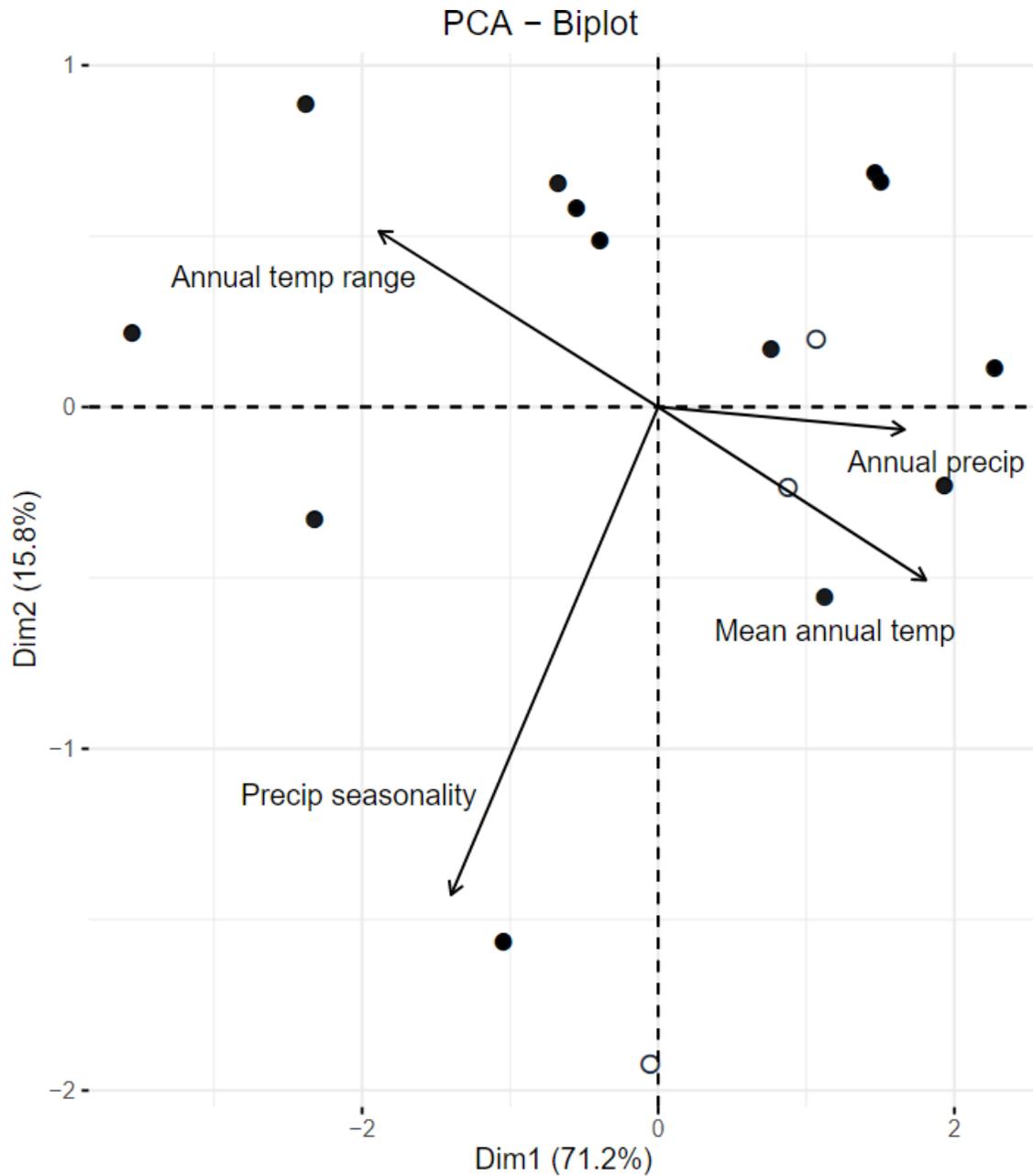


Figure 2 - shows the loadings for climate PCA included, for each plot, Mean Annual Temperature, Mean Annual Temperature Range, Total Annual Precipitation, and Precipitation Seasonality. Closed dots represent the moist forest plots, while open dots represent the dry, seasonal forest plots.

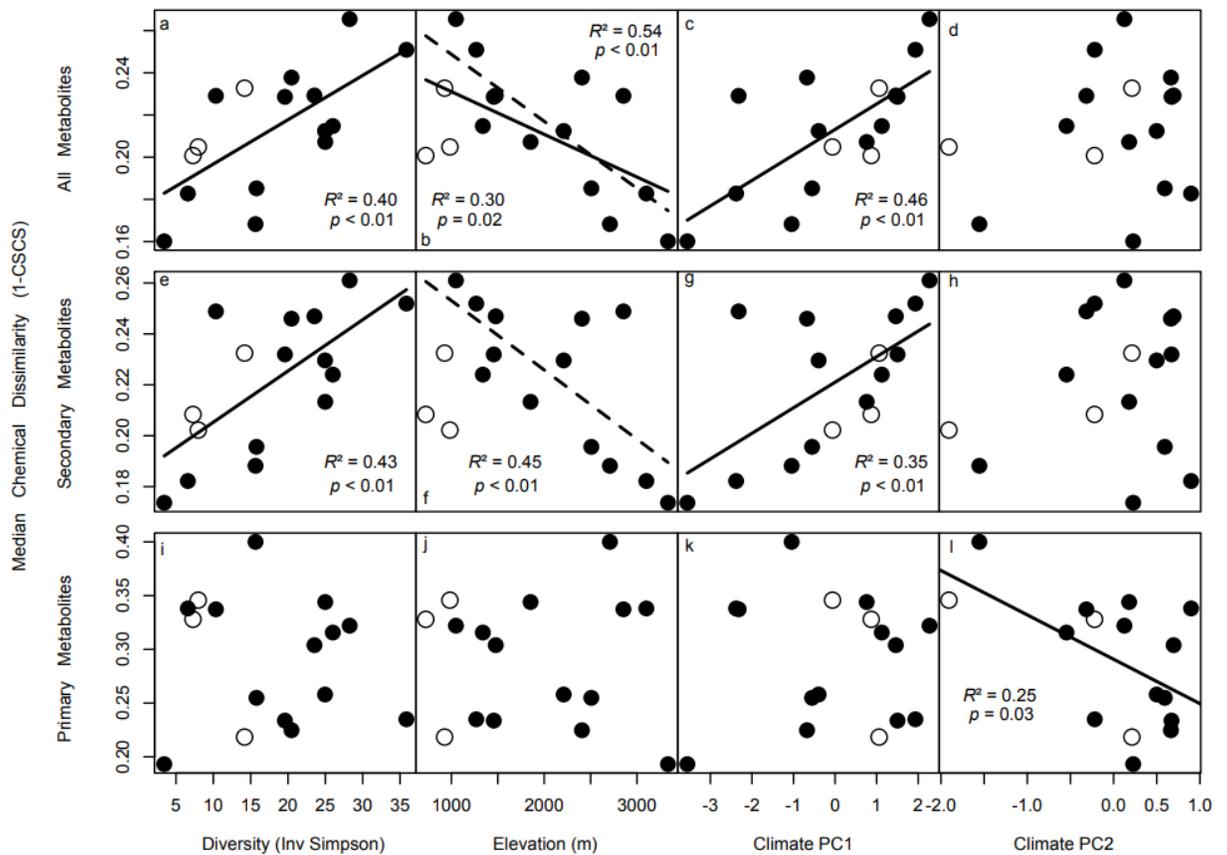


Figure 3 – Variation in median chemical dissimilarity (1-CSCS) vs species diversity (inverse Simpson index), elevation (m), and climate among 16 forest plots in Madidi, Bolivia. Panels a-d represent linear regressions between median chemical dissimilarity among co-occurring species with respect to the whole metabolite and (a) species diversity, (b) elevation, (c) Climate PC1, and (d) Climate PC2, respectively. Panels e-h represent linear regressions between chemical dissimilarity with respect to secondary metabolites and (e) species diversity, (f) elevation, (g) Climate PC1, and (h) Climate PC2, respectively. Panels i-l represent linear regressions between chemical dissimilarity with respect to primary metabolites and (i) species diversity, (j) elevation, (k) Climate PC1, and (l) Climate PC2, respectively. Secondary metabolites are defined as those derived from the Alkaloids, Amino acid and Peptides, Polyketides, Shikimates and Phenylpropanoids, and Terpenoids biosynthetic pathways using NPClassifier (Kim et al. 2022). Primary metabolites are defined as those derived from the Carbohydrates and Fatty acids pathways. Three seasonally dry forests are represented by open circles. Regressions using all 16 forest plots are represented by solid lines; regressions excluding three seasonally dry forests are represented by dashed lines. Adjusted R^2 and p -values are presented for significant ($p < 0.05$) regressions only.

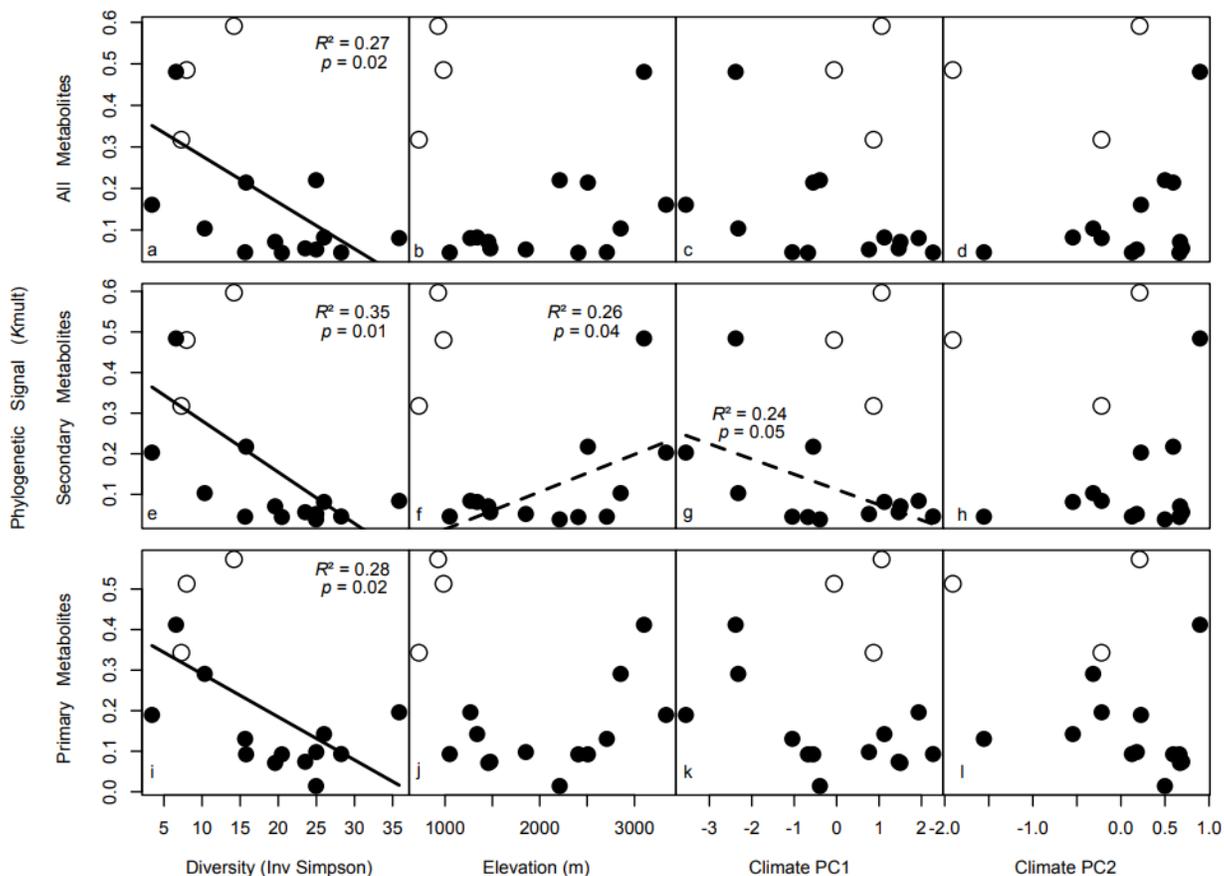


Figure 4 – Variation in chemical phylogenetic signal (K_{mult} ; Adams 2014) among co-occurring species vs species diversity (inverse Simpson index), elevation (m), and climate among 16 forest plots in Madidi, Bolivia. Panels a-d represent linear regressions between phylogenetic signal among co-occurring species with respect to the whole metabolite and (a) species diversity, (b) elevation, (c) Climate PC1, and (d) Climate PC2, respectively. Panels e-h represent linear regressions between phylogenetic signal with respect to secondary metabolites and (e) species diversity, (f) elevation, (g) Climate PC1, and (h) Climate PC2, respectively. Panels i-l represent linear regressions between phylogenetic signal with respect to primary metabolites and (i) species diversity, (j) elevation, (k) Climate PC1, and (l) Climate PC2, respectively. Secondary metabolites are defined as those derived from the Alkaloids, Amino acid and Peptides, Polyketides, Shikimates and Phenylpropanoids, and Terpenoids biosynthetic pathways using NPClassifier (Kim et al. 2022). Primary metabolites are defined as those derived from the Carbohydrates and Fatty

acids pathways. Three seasonally dry forests are represented by open circles. Regressions using all 16 forest plots are represented by solid lines; regressions excluding three seasonally dry forests are represented by dashed lines. Adjusted R^2 and p -values are presented for significant ($p < 0.05$) regressions only.

Chapter 4: Functional Trait Patterns Reveal Differential Mechanisms of Tree Community Assembly Along a Tropical Elevational Gradient

4.1 Introduction

An enduring motivation of ecologists is to find and define overarching/ general mechanisms by which communities of different species assemble (Vellend 2010, Mittlebach & Schemske 2015). A second longstanding motivation that ecologists seek is to uncover how those species' interactions affect community assembly (Schemske et al. 2009, Vellend 2010). These two aims are intertwined, and advances in one motivation may engender advances in the other. The existing view of community assembly often separates the complex dynamics of community assembly and species interactions into separate or nested processes (HilleRisLambers et al. 2012) that act as a result of differing forces, both biotic and abiotic (McGill et al. 2006). The biotic and the abiotic environments are hypothesized to have different effects on the resulting communities, but little research has been done on the interplay between (McGill et al. 2006, Kraft et al. 2015).

Abiotic forces are associated with attributes of the physical environment that influence a species' fitness. The extension of these ideas at the community scale is often termed abiotic filtering; a situation in which only certain species can survive and maintain in a given environment, such that the environment selects for the species found within (HilleRisLambers et

al. 2012, Kraft et al. 2015). In the case of functional traits, this would mean that only a subset of traits can survive there, so the species within would show similar, or clustered, trait patterns.

Much of the previous research relied upon morphological traits to quantify the effect of the abiotic environment, and indeed, morphological traits give valuable information about how species acquire resources, withstand cold, drought and other attributes of their abiotic environment. For example, the specific leaf area (SLA), leaf size, and relative growth rate (RGR) would decrease in cooler and dryer locations, but the leaf thickness would increase. However, the abiotic environment might also have an effect on chemical traits as well. In seasonal locations, especially where trees are deciduous, there could be a reduced investment in foliar chemical defenses and a greater reliance on other defensive strategies (Sam et al. 2020), such as synchronous leaf growth (Forrester et al. 2019), than in less seasonal locations where leaves may remain on trees for several years.

The biotic environment, on the other hand, can be conceptualized as the total effect of the biotic interactions operating in a community (McGill et al. 2006, Švamberková et al. 2017), often in regard to fitness. Biotic forces that affect community assembly have, in the past, been considered only as competitive interactions, the result of which is that the more competitive species are predicted to thrive or even to exclude their competitors from the community (Cadotte et al. 2015). Another view of the biotic landscape considers the interactions between species of different trophic levels (Janzen 1970, Connell 1971), often referred to as enemy mediated interactions. In contrast to standard ideas of competition for resources, enemy mediated interactions hypothesizes that specialized interactions between plant hosts and their natural enemies play a strong role in community assembly, with the hosts ability to defend against or avoid their natural enemies largely defining the host species' niche. The strength of such enemy

mediated interactions is believed to be stronger in the tropics (Schemske et al. 2009). This view has explanatory potential, as there are an almost infinite variety of natural enemies that can potentially provide an equal number of niches by which plant hosts can differentiate (Erlich & Raven 1964, Bever et al. 2015). This mufti-trophic view of the biotic environment has been given more consideration in recent years as the methods for testing the ideas have become available (Wang et al. 2016, Sedio 2017).

If the hypothesized biotic environment can be viewed as being built upon the back-and-forth interactions between plant hosts and their natural enemies, then the plant traits that are used to quantify the biotic environment are the chemical defensive metabolites that the hosts employ to escape, evade, or deter their enemies (Erlich & Raven 1964, Wetzels & Whitehead 2020). While plant-host-defenses can be morphological in nature (Forrister et al. 2019), the great variety of plant-chemical metabolites, many of which are employed defensively (Kursar et al. 2009, Salazar et al. 2016), have been shown to be a hypothetically feasible way to quantify these interactions, resulting in an increasing number of analyses of enemy mediated interactions in recent years (Kursar et al. 2009, Salazar et al. 2016, Forrister et al. 2019, Sedio et al. 2021, Endara et al. 2022, Volf et al. 2022). The conclusions from this body of research have supported the idea that species differentiate along the lines of their chemical-defensive traits, whereby species that have similar defenses suffer attack from shared enemies resulting in decreased fitness, while those that do not share defenses avoid enemies, resulting in increased fitness and or survival of those chemically dissimilar species (Coley et al. 2018, Sedio et al. 2018).

Whether considering chemical-defensive traits in relation to the biotic environment, or morphological traits in relation to the abiotic environment, the expected relationships can be measured similarly. Overdispersion of traits occurs when traits are more dissimilar than expected

by chance. In the case of morphological traits, overdispersed trait patterns can occur from either competition for resources leading to character displacement (MacArthur & Levins 1967) or niche differentiation (Kraft et al. 2008, Laughlin et al. 2020). Overdispersion can also occur from strong pressure from enemy mediated interactions where species with similar chemical profiles will be more vulnerable to attack from shared enemies, in the case of chemical-defensive traits. Underdispersion of traits is the opposite of overdispersion, where traits are more similar than expected, as a result of abiotic filtering and/ or resource limitation in the abiotic environment. Harsher environments, whether having more extreme climates or poorer nutrient resources can exert stronger selective pressure, resulting in only a subset of potential morphological and chemical forms able to exist there (May et al. 2013), and can result in underdispersed trait patterns in both suites of traits. Morphological traits would show a clustered pattern, with traits that reflect the abiotic challenges present (such as thicker smaller leaves in colder environments or deciduous leaves in more seasonal environments), while chemical defensive traits may show greater similarity simply because in harsh or nutrient-poor environments individuals are less able to invest in defensive chemicals (Carmona et al. 2011, Mithöfer et al. 2012).

Despite the longstanding attention of ecologists, there is still uncertainty regarding how the biotic and abiotic environments affect community assembly and how their relative strengths may vary (Kraft et al. 2015). Many previous studies have assumed that functional trait patterns are driven by the abiotic environment without considering at all the effects of the biotic (McGill et al. 2006, May et al. 2013, Kraft et al. 2015). Most of the existing studies that do consider the biotic environment in relation to the effects of biotic interactions on community assembly, focus on competitive interactions at the same trophic level, without considering multi-trophic, enemy-mediated interactions (Cavander-Bares et al. 2006, Pontarp & Petchey 2016, Wetzel &

Whitehead 2020, Henn et al. *in review*). Also, while there has been research on both the biotic and abiotic environments separately, only a limited number of studies have considered both together (Sedio et al. 2021, Henn et al. *in review*). Considering these constraints on knowledge, a study that considers the effects of both the abiotic as well as the biotic environments, while taking into account the potentially multi-trophic nature of biotic interactions would be extremely valuable in predicting how community assembly mechanisms change across gradients. We bridge this gap by using the latest innovations in large-scale ecological, chemical metabolomics, with a clear focus on two different but interacting forces/ processes the result of which are patterns of morphological and chemical traits.

In this study we ask how community assembly is affected by both the biotic and the abiotic environments, by examining the relationships between two suites of plant functional traits, chemical and morphological, and abiotic site characteristics. We used data from 16 1-ha forest plots distributed across an ~3000 m elevational-diversity gradient in the Bolivian Andes (17-137 tree species per plot). By using large-scale chemical-metabolomic analytical techniques (Wang et al. 2016, Sedio 2017, Sedio et al. 2018), we were able to compare patterns of both foliar leaf-defensive compounds, as well as 8 selected morphological traits that describe the resource acquisition strategies of the tree species, in relation to a battery of climatic and soil variables that thoroughly characterize the abiotic environment along the gradient. We tested 2 hypotheses: 1) Biotic interactions are more important determinants of community structure in warmer, wetter, less seasonal, nutrient-rich environments, with the prediction that overdispersion of chemical traits will increase towards warmer, wetter, less seasonal, nutrient-rich environments and 2) Abiotic filtering is a more important determinant of community structure colder, drier, more seasonal, nutrient poor environments, with the prediction of greater underdispersion of

morphological traits towards colder, drier, more seasonal, nutrient poor environments. This study takes a more nuanced look at community assembly by examining different, but potentially interacting forces across a tropical-elevational gradient. It also provides a timely and complementary view of recent, similar studies done in the temperate zone (Henn et al. *in review*).

4.2 Methods

1. Floristic Data & Study Area

Tree species data, both taxonomic as well as morphological and chemical functional trait data, were collected as part of the Madidi project, a 20 plus year collaboration to document the flora of the Madidi region in the Andes of Bolivia (Tello et al. 2015; Figure 1). The region ranges in elevation from lowland rainforests located at around 200 m to high mountains above 6,000 m, above the tree line (Fuentes 2005). The elevational gradient is covered by different forest types and encompasses a broad range of abiotic (climatic and environmental) conditions (Rafiqpoor & Ibish 2004). Species composition and abundance of woody plants were obtained from 50, spatially mapped, 1-hectare plots. Plot elevations rise from 212m to 3334m. Within each plot, all woody plants with a diameter at breast height of at least 10 cm were spatially mapped measured and identified to a valid species or morphospecies name, through further taxonomic work at the Herbario Nacional de Bolivia, in La Paz, Bolivia and at the Missouri Botanical Garden in St. Louis, USA. The majority of the permanent plots have been censused a minimum of two times. Of these 50 plots, 16 plots with available chemical trait data were

selected, ranging from the least species rich, Kañupa with only 17 species to Lomaka, the most species rich, with 137 species, to carry out this investigation.

Trait Data

Functional trait data were collected during multiple censuses. Five individuals of each species, in each plot, were censused, when abundances allowed. The field collection and laboratory methods used in the determination of functional trait values for the Madidi censuses were adapted from (Cornelissen et al. 2003) and the Madidi Project methods manual (Jørgensen et al. 2015) describes the methods further. Before testing the predictions related to both sets of functional traits, the trait data were cleaned of misnamed or unidentified species. The species included in each data set was matched with the species found in each of the other datasets.

2. Morphological Data

We selected 8 morphological functional traits that were specifically intended to encompass the whole of a forest tree's life history strategy (Table S2; Figure S1). Max tree height and max diameter at breast height (DBH), are associated with competitive vigor and overall growth and life history strategy. Both leaf area and specific leaf area (SLA) are associated with photosynthetic or growth rate, SLA additionally gives clues about the life history trade off between defense and growth rate, while relative growth rate (RGR) is an inclusive measure of growth rate. The traits leaf thickness, leaf thickness, and twig specific density are all associated with plant defense/ resistance to the environment (biotic and abiotic), while twig specific density further gives insight into the growth, survival tradeoff (Perez-Harguindeguy et

al. 2013). The morphological traits were centered and transformed, prior to being included in a principal components analysis, of which all 8 trait axes were used in the downstream analyses.

3. Chemical Data

Chemical trait data were derived according to Sedio et al. (2021) from leaf tissue samples collected as part of the Madidi project censuses and then preserved with silica gel. Only samples collected after 2011 were used for these analyses. Briefly, for each species in each plot, 10 mg of powdered leaf tissues were extracted overnight in a high percentage methanol solution. Extracts of five individuals per species per plot were pooled to create 906 extract pools representing each unique species-by-plot for the subsequent analysis. Ultra-high-performance liquid chromatography tandem mass spectrometry followed by heated electrospray ionization were used to analyze extracts (Sedio et al. 2017, 2018) and the Global Natural Products Social (GNPS) Molecular Networking tool to cluster the MS/MS spectra into consensus spectra that represent unique structures (Wang et al. 2016). NPClassifier was used to classify metabolites according to basic biosynthetic pathways, into classes and superclasses (Kim et al. 2022). The consensus spectra are referred to as compounds or metabolites throughout. These methods differentiated between both primary metabolites involved in resource-acquisitive metabolic functions, (which tend to be conserved across most plants, Sedio et al. 2018), and secondary metabolites, hypothesized to be involved in defense. Secondary metabolites display a greater diversity and much greater interspecific variability (Salminen and Karonen 2011) than primary metabolites. Due to their widespread defensive use, were used in the analyses. We calculated the chemical structural and compositional similarity (CSCS), which is a measure of pairwise

metabolite similarity among species, that takes into account similar, but non-identical chemical structures (Sedio 2017).

Environmental Data

The environmental data variables were scaled and centered, prior to a principal components analysis being run, of which the first two principal components axes were taken as independent variables to be used in the downstream analyses. The environmental data used was at the whole plot level only.

4. Climate Data

Climate variables were selected to represent the variation in plot temperature, precipitation, and seasonality over the elevational gradient. The temperature variables annual mean temperature and temperature annual range were used from WorldClim Version 2.1 (Fick & Hijmans 2017). The precipitation variables total annual precipitation and precipitation seasonality were used from the Tropical Rainfall Measuring Mission (TRMM) because of the greater accuracy compared to WorldClim data in the Madidi region.

5. Soil Data

A total of 13 soil variables were selected to characterize the variation in plot-soil nutrient richness, soil type, and local growth conditions (Appendix C: Table S2). Through the use of these variables, we sought to quantify the differences in the available soil resources between plots along the gradient. The soil data were additionally log transformed prior to the PCA.

6. Statistical Tests of Hypothesis

To quantify the differences in trait values, the abundance-weighted mean pairwise distance (MPD) of chemical and morphological functional traits within subplots were calculated using the `ses.mpd` function in the `picante` package in R. MPD is simply the mean of all of the pairwise functional distances of all of the species within a community or subplot. The mean of all the subplots was taken as the overall mean plot value. The morphological and chemical data were first transformed into dissimilarity-distance matrices of species' within-plot trait values. The plot MPD was calculated at the 5, 10 and 20 m subplot scales.

In order to account for within-plot variation not directly due to the environmental variables, we utilized a randomized null model. The null model randomizes the species trait values within a pot while preserving the total number of individuals, species abundances, and positions of individuals within the plots or subplots. We ran $n=999$ null model permutations. Any significant variation in MPD values from the randomized expectations of the null model were thus deemed due to the effects of the environmental climate and soil variables. To test statistically whether the observed values differed from random, the standard effect size (SES) of the MPD values was used, calculated by subtracting the randomized MPD values from the observed values and dividing by the standard deviation of the randomized values. To quantify the relationship of the traits MPD with the environmental variables, we linearly regressed the SES values against the first two PCA axes of the climate and soil variables. Positive deviations from random (greater than zero) indicate that trait dissimilarity values differ more than expected with regard to the independent variables, and negative deviations from random (less than zero) indicate that trait dissimilarity values differ less than expected with regard to the variables.

4.3 Results

Overview Results

Overall, we found that the majority of plots had less dispersed trait patterns than expected (negative SES values) for both chemical and morphological traits. The relationships among the three subplot scales were similar (Figures 1&2). The range of functional dispersal increases with spatial size of subplots. There was no relationship between trait distances of chemical and morphological traits across the entire data set; chemical and morphological traits are not orthogonal (Figure S2).

PCA

Climate PC1 explains 71% variation and describes a gradient in temperature, precipitation and seasonality, with annual mean temperature and annual precipitation loading strongly and as well as annual temperature range partially loading on PC1. Soil PC1 explains 28% of the variation and describes a gradient in soil interchangeable ions, with interchangeable calcium, interchangeable magnesium, interchangeable potassium, and clay loading negatively. Soil PC2 explains 26% variation and describes a gradient in soil organic content and total nitrogen, with organic material, organic carbon, total nitrogen loading strongly on PC2.

Prediction 1

The first prediction (that chemical traits will show greater overdispersion in warmer, wetter, less seasonal, or more nutrient rich environments) was rejected. We did not find evidence of overdispersion of secondary metabolites at any spatial grain size (Figure 2). The

majority of the plots had subplots with less than expected functional dispersion of chemical traits, at all subplot scales.

Prediction 2

There was partial support for the second prediction that there would be less functional dispersion, for morphological traits, in cooler, dryer, less seasonal, more nutrient poor environments. We found (significant relationships) evidence of greater underdispersion (filtering) of morphological traits at lower values of soil PC1, at all scales (5m: p-value: 0.024, R^2 : 0.27; 10m: p-value: 0.0077, R^2 : 0.37; 20m: p-value: 0.014, R^2 : 0.32; Figure 3).

4.4 Discussion

Community assembly is influenced by a combination of factors (Vellend 2010). The abiotic environment consists of those factors that result from the physical conditions of the site, such as temperature, precipitation, or soil content. Morphological traits, specifically those that are involved in resource acquisition and physical structure, are thought to best approximate the abiotic environment. The biotic environment describes the influence of species interactions within the communities and can be approximated by traits that reflect the interactions between species, whether at the same trophic level (competition) or between different trophic levels (enemy-mediated interactions). Chemical-defensive traits were used in this case, to approximate enemy-mediated interactions. We investigated how community assembly is affected by both the abiotic and the biotic environments by examining relationships within two groups of functional

traits, chemical and morphological. The relationships were not immediately apparent. However, there was some evidence for an increasing relationship between morphological traits and wetter, warmer, more seasonal, and nutrient-rich environments, suggesting that the importance abiotic filtering increases (changes systematically) along the gradient. Across the whole of the sampled plots, for both chemical and morphological traits, the majority of the plots were more undispersed than expected by chance suggesting at community subplot scales, that the abiotic environment affect trait patterns more than expected, for both groups of traits.

There was no evidence of overdispersion of chemical traits in relation to climate or soil variables across the elevational gradient. This could mean that the importance of biotic interactions in determining community structure does not vary across the gradient. Perhaps more likely, it could be that the chemical defensive compounds sampled do not closely exactly approximate the influence of the biotic environment.

It has become apparent to ecologists that plant-enemy interaction networks can be ecologically complex and that understanding these networks is important in elucidating the mechanisms at work within assemblies of communities, and in selecting the proper functional traits to do so (Sam et al. 2020). For example, Volf et al. (2020) in a study on the variance of several individual classes of tree foliar metabolites, found that the individual classes of metabolites varied according to different factors and that some compound classes were better described by abiotic factors such as temperature or humidity. They also found that different compound classes may be employed differentially (i.e. as resistors to abiotic stressors, as defenses against herbivores and/or mammals). Sam et al. (2020) found that herbivore damage along an elevational gradient varied significantly with elevation and season. They found that herbivore pressure is strongest in the most productive environments, that have the most abundant

available water. Volf et al. (2022) found that the abiotic environment, specifically temperature stress, had a greater effect on the individual classes of chemical metabolites produced in trees along a temperate elevational gradient. They also found specific classes of compounds to have differing probable ecological functions. Each of the previous studies found that the abiotic environment was found to have a greater than expected effect on the composition of chemical produced by plant species, further complicating the link between enemy-mediated interactions and chemical defenses (Fine et al. 2004, Defosse et al. 2021). A second important commonality among these studies is that they found specific classes of chemical metabolites to vary in function and abundance along a gradient. These findings suggest that the traits chosen to approximate the biotic landscape must be chosen with the probable function of the specific metabolite or class of metabolites in mind.

We found evidence of greater morphological dispersion in more nutrient rich environments which could suggest that the strength of abiotic filtering is reduced in more productive environments as either the more benign environment allows a greater chance of survival or that it allows for a greater extent of successful functional strategies to coexist. In nutrient poor locales, the best survival strategies may involve tradeoffs (Fine et al. 2006) as environmental conditions change along a gradient. Defosse et al. (2018) found covariation in a suite of defensive traits (morphological as well as some volatile compounds) along a temperate elevational gradient, in which plant species were clustered into separate defense ‘syndromes’ along the gradient, likely as a result of defensive tradeoffs necessitated by limited resources, supporting this supposition.

Our results underscore the importance and complexity of the abiotic environment on community trait patterns. The abiotic environment may have a greater than expected effect on

community trait patterns. The question of how well the traits selected describe the biotic and abiotic environments remains.

If the relationships between the biotic and abiotic environments are not clear, neither will their effects on morphological and chemical trait patterns. Several studies have found morphological and chemical traits to be largely orthogonal (Sedio et al. 2021), however this pattern is not always consistent (Labarrere et al. 2019, Henn et al. *in review*, Figure S2). The literature on how morphological traits are influenced by the abiotic environment is already quite extensive (Laughlin et al. 2020). As the biotic environment is complex and difficult to capture, choosing the right set of traits that reflect fitness or factors influencing fitness is a challenging task (Laughlin et al. 2020). Studies showing that different classes of chemical metabolites correspond to predictable adaptive or defensive purposes, such as flavonoids providing protection against abiotic stress (Volf et al. 2020), allows us to view the landscape of enemy-mediation interactions in a finer degree of detail. However, chemical traits are also influenced by the abiotic environment. The three previous studies (Sam et al. 2020, Volf et al. 2020, Volf et al. 2022) found the abiotic environment to have a greater than expected effect on the production of chemical compounds. Not only were specific classes of metabolites strongly related to certain environmental factors such as temperature or humidity, the chemical, but the chemical composition of species changed seasonally. This stronger than expected effect further complicates the selection of chemical traits. The selection of chemical defensive traits must be further refined.

There are several notes worth mentioning regarding the interpretation of these data. Within all of the plots analyzed, only individual stems greater than 10 cm DBH were sampled, possibly

excluding juvenile or developmental stages of some species, and possibly adult individuals of smaller species. Despite the wide variation in plot diversity and abiotic conditions along it, the whole of the study site is tropical, and along the gradient, there was less trait dispersion than expected overall. Another point to consider is that plant species' metabolite concentrations can change seasonally or that plants may invest more in plant defenses at different times of the year (perhaps especially in more seasonal forests; Sam et al. 2020). The investment of chemical metabolites is costly (Schoonhoven et al. 2005), such that research has found that species may invest in groups defense strategies that would cause individual defensive compounds to co-occur (defensive syndromes, Defosse et al. 2018). This is especially important when considering the influence of climate or soil conditions on chemical trait patterns. An illuminating future direction would analyze separate classes of defensive compound individually, formulating specific predictions based off their likely defensive or adaptive uses. It would also be helpful to test which chemical metabolites co-occur along gradients or within a study system.

We found the abiotic environment to have complex and intertwined effects with biotic environment and on trait patterns. Overall, our findings support the idea that there is systematic variation in the biotic and abiotic factors that structure tropical forest communities, but that the abiotic may influence the biotic, and that the interactions between species must be examined in yet finer detail. Undoubtedly, both the biotic and abiotic environments are important in community assembly, and that further understanding of their relative roles is of critical importance.

Acknowledgements

We thank J. Dwenger for assistance in the lab and M. Volf for valuable discussion. We thank the Dirección General de Biodiversidad, the Bolivian Park Service (SERNAP), the Madidi National Park and local communities for permits, access and collaboration in Bolivia, where fieldwork was supported by the National Science Foundation (DEB 0101775, DEB 0743457, DEB 1836353). Additional financial support to the Madidi Project was provided by the Missouri Botanical Garden, the National Geographic Society (NGS 7754-04 and NGS 8047-06), The German Embassy in Bolivia, International Center for Advanced Renewable Energy and Sustainability (I-CARES) at Washington University in St. Louis, the Comunidad de Madrid (Spain), Consejo Superior de Investigaciones Científicas (Spain), Centro de Estudios de América Latina (Banco Santander and Universidad Autónoma de Madrid, Spain), and the Taylor and Davidson families. Financial support for this study was provided by a grant from the Living Earth Collaborative at Washington University in St. Louis and the National Science Foundation (DEB Award 1557094 to JAM, DEB Award 2240431 to JAM, DEB Award 2240430 to BES). DH was supported by a Catharine M. Lieneman Botany Fellowship at Washington University in St. Louis.

4.5 References

- Bever, J. D., Mangan, S. A., & Alexander, H. M. (2015). Maintenance of plant species diversity by pathogens. *Annual review of ecology, evolution, and systematics*, 46, 305-325.
- Cadotte, M. W., Arnillas, C. A., Livingstone, S. W., & Yasui, S. L. E. (2015). Predicting communities from functional traits. *Trends in Ecology & Evolution*, 30(9), 510-511.
- Carmona, D., Lajeunesse, M. J., & Johnson, M. T. (2011). Plant traits that predict resistance to herbivores. *Functional Ecology*, 25(2), 358-367.

- Cavender-Bares, J., Keen, A., & Miles, B. (2006). Phylogenetic structure of Floridian plant communities depends on taxonomic and spatial scale. *Ecology*, 87(sp7), S109-S122.
- Coley, P. D., Endara, M. J., & Kursar, T. A. (2018). Consequences of interspecific variation in defenses and herbivore host choice for the ecology and evolution of *Inga*, a speciose rainforest tree. *Oecologia*, 187, 361-376.
- Connell, J. H. (1971). On the role of natural enemies in preventing competitive exclusion in some marine animals and in rain forest trees. *Dynamics of populations*, 298(312).
- Cornelissen, J. H. C., Lavorel, S., Garnier, E., Díaz, S., Buchmann, N., Gurvich, D. E., ... & Poorter, H. (2003). A handbook of protocols for standardised and easy measurement of plant functional traits worldwide. *Australian journal of Botany*, 51(4), 335-380.
- Defossez, E., Pellissier, L., & Rasmann, S. (2018). The unfolding of plant growth form-defence syndromes along elevation gradients. *Ecology Letters*, 21(5), 609-618.
- Defossez, E., Pitteloud, C., Descombes, P., Glauser, G., Allard, P. M., Walker, T. W., ... & Rasmann, S. (2021). Spatial and evolutionary predictability of phytochemical diversity. *Proceedings of the National Academy of Sciences*, 118(3), e2013344118.
- Dobzhansky, T. (1950). Evolution in the tropics. *American scientist*, 38(2), 209-221.
- Ehrlich, P. R., & Raven, P. H. (1964). Butterflies and plants: a study in coevolution. *Evolution*, 586-608.
- Endara, M. J., Soule, A. J., Forrister, D. L., Dexter, K. G., Pennington, R. T., Nicholls, J. A., ... & Coley, P. D. (2022). The role of plant secondary metabolites in shaping regional and local plant community assembly. *Journal of Ecology*, 110(1), 34-45.
- Fick, S. E., & Hijmans, R. J. (2017). WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. *International journal of climatology*, 37(12), 4302-4315.
- Fine, P. V., Mesones, I., & Coley, P. D. (2004). Herbivores promote habitat specialization by trees in Amazonian forests. *science*, 305(5684), 663-665.
- Fine, P. V., Miller, Z. J., Mesones, I., Irazuzta, S., Appel, H. M., Stevens, M. H. H., ... & Coley, P. D. (2006). The growth-defense trade-off and habitat specialization by plants in Amazonian forests. *Ecology*, 87(sp7), S150-S162.
- Fuentes, A. (2005). Una introducción a la vegetación de la región de Madidi. *Ecología en Bolivia*, 40(3), 1-31.
- HilleRisLambers, J., Adler, P. B., Harpole, W. S., Levine, J. M., & Mayfield, M. M. (2012). Rethinking community assembly through the lens of coexistence theory. *Annual review of ecology, evolution, and systematics*, 43, 227-248.
- Janzen, D. H. (1970). Herbivores and the number of tree species in tropical forests. *The American Naturalist*, 104(940), 501-528.
- Jørgensen, P. M., Fuentes, A. F., Miranda, T., & Cayola, L. (2015). Manual de Trabajo.
- Kim, S., Lim, S. W., & Choi, J. (2022). Drug discovery inspired by bioactive small molecules from nature. *Animal Cells and Systems*, 1-12.

- Kraft, N. J., Valencia, R., & Ackerly, D. D. (2008). Functional traits and niche-based tree community assembly in an Amazonian forest. *Science*, 322(5901), 580-582.
- Kraft, N. J., & Ackerly, D. D. (2010). Functional trait and phylogenetic tests of community assembly across spatial scales in an Amazonian forest. *Ecological monographs*, 80(3), 401-422.
- Kraft, N. J., Adler, P. B., Godoy, O., James, E. C., Fuller, S., & Levine, J. M. (2015). Community assembly, coexistence and the environmental filtering metaphor. *Functional ecology*, 29(5), 592-599.
- Kursar, T. A., Dexter, K. G., Lokvam, J., Pennington, R. T., Richardson, J. E., Weber, M. G., ... & Coley, P. D. (2009). The evolution of antiherbivore defenses and their contribution to species coexistence in the tropical tree genus *Inga*. *Proceedings of the National Academy of Sciences*, 106(43), 18073-18078.
- Labarrere, B., Prinzing, A., Dorey, T., Chesneau, E., & Hennion, F. (2019). Variations of secondary metabolites among natural populations of sub-Antarctic ranunculus species suggest functional redundancy and versatility. *Plants*, 8(7), 234.
- Laughlin, D. C., Gremer, J. R., Adler, P. B., Mitchell, R. M., & Moore, M. M. (2020). The net effect of functional traits on fitness. *Trends in Ecology & Evolution*, 35(11), 1037-1047.
- Lebrija-Trejos, E., Pérez-García, E. A., Meave, J. A., Bongers, F., & Poorter, L. (2010). Functional traits and environmental filtering drive community assembly in a species-rich tropical system. *Ecology*, 91(2), 386-398.
- MacArthur, R., & Levins, R. (1967). The limiting similarity, convergence, and divergence of coexisting species. *The american naturalist*, 101(921), 377-385.
- May, F., Giladi, I., Ristow, M., Ziv, Y., & Jeltsch, F. (2013). Plant functional traits and community assembly along interacting gradients of productivity and fragmentation. *Perspectives in Plant Ecology, Evolution and Systematics*, 15(6), 304-318.
- McGill, B. J., Enquist, B. J., Weiher, E., & Westoby, M. (2006). Rebuilding community ecology from functional traits. *Trends in ecology & evolution*, 21(4), 178-185.
- Mithöfer, A., & Boland, W. (2012). Plant defense against herbivores: chemical aspects. *Annual review of plant biology*, 63, 431-450.
- Mittelbach, G. G., & Schemske, D. W. (2015). Ecological and evolutionary perspectives on community assembly. *Trends in ecology & evolution*, 30(5), 241-247.
- Perez-Harguindeguy, N., Diaz, S., Garnier, E., Lavorel, S., Poorter, H., Jaureguiberry, P., ... & Cornelissen, J. H. C. (2016). Corrigendum to: New handbook for standardised measurement of plant functional traits worldwide. *Australian Journal of botany*, 64(8), 715-716.
- Pontarp, M., & Petchey, O. L. (2016). Community trait overdispersion due to trophic interactions: concerns for assembly process inference. *Proceedings of the Royal Society B: Biological Sciences*, 283(1840), 20161729.
- Rafiqpoor, D., & Ibsch, P. (2004). The physical environment of Bolivia. *Biodiversity: The richness of Bolivia. State of knowledge and conservation. Editorial Fundación Amigos de la Naturaleza, Santa Cruz de la Sierra*, 4-45.
- Salazar, D., Jaramillo, A., & Marquis, R. J. (2016). The impact of plant chemical diversity on plant–herbivore interactions at the community level. *Oecologia*, 181, 1199-1208.

- Salminen, J. P., & Karonen, M. (2011). Chemical ecology of tannins and other phenolics: we need a change in approach. *Functional ecology*, 25(2), 325-338.
- Sam, K., Koane, B., Sam, L., Mrazova, A., Segar, S., Volf, M., ... & Novotny, V. (2020). Insect herbivory and herbivores of *Ficus* species along a rain forest elevational gradient in Papua New Guinea. *Biotropica*, 52(2), 263-276.
- Schoonhoven, L. M., Van Loon, J. J., & Dicke, M. (2005). *Insect-plant biology*. Oxford University Press on Demand.
- Schemske, D. W., Mittelbach, G. G., Cornell, H. V., Sobel, J. M., & Roy, K. (2009). Is there a latitudinal gradient in the importance of biotic interactions?. *Annu. Rev. Ecol. Evol. Syst.*, 40, 245-269.
- Sedio, B. E. (2017). Recent breakthroughs in metabolomics promise to reveal the cryptic chemical traits that mediate plant community composition, character evolution and lineage diversification. *New Phytologist*, 214(3), 952-958.
- Sedio, B. E., Rojas Echeverri, J. C., Boya P, C. A., & Wright, S. J. (2017). Sources of variation in foliar secondary chemistry in a tropical forest tree community.
- Sedio, B. E., Parker, J. D., McMahon, S. M., & Wright, S. J. (2018). Comparative foliar metabolomics of a tropical and a temperate forest community.
- Sedio, B. E., Boya P, C. A., & Rojas Echeverri, J. C. (2018). A protocol for high-throughput, untargeted forest community metabolomics using mass spectrometry molecular networks. *Applications in Plant Sciences*, 6(3), e1033.
- Sedio, B. E., Spasojevic, M. J., Myers, J. A., Wright, S. J., Person, M. D., Chandrasekaran, H., ... & Vandermeer, J. H. (2021). Chemical similarity of co-occurring trees decreases with precipitation and temperature in North American forests. *Frontiers in Ecology and Evolution*, 9, 679638.
- Švamberková, E., Vítová, A., & Lepš, J. (2017). The role of biotic interactions in plant community assembly: What is the community species pool?. *Acta Oecologica*, 85, 150-156.
- Tello, J. S., Myers, J. A., Macía, M. J., Fuentes, A. F., Cayola, L., Arellano, G., ... & Jørgensen, P. M. (2015). Elevational gradients in β -diversity reflect variation in the strength of local community assembly mechanisms across spatial scales. *PLoS One*, 10(3), e0121458.
- Terborgh, J. (2012). Enemies maintain hyperdiverse tropical forests. *The American Naturalist*, 179(3), 303-314.
- Vellend, M. (2010). Conceptual synthesis in community ecology. *The Quarterly review of biology*, 85(2), 183-206.
- Volf, M., Laitila, J. E., Kim, J., Sam, L., Sam, K., Isua, B., ... & Segar, S. T. (2020). Compound specific trends of chemical defences in *Ficus* along an elevational gradient reflect a complex selective landscape. *Journal of chemical ecology*, 46(4), 442-454.
- Volf, M., Volfová, T., Hörandl, E., Wagner, N. D., Luntamo, N., Salminen, J. P., & Sedio, B. E. (2022). Abiotic stress rather than biotic interactions drives contrasting trends in chemical richness and variation in alpine willows. *Functional Ecology*, 36(11), 2701-2712.
- Wallace, A. R. (1878). *Tropical nature, and other essays*. Macmillan and Company.

Wang, M., Carver, J. J., Phelan, V. V., Sanchez, L. M., Garg, N., Peng, Y., ... & Bandeira, N. (2016). Sharing and community curation of mass spectrometry data with Global Natural Products Social Molecular Networking. *Nature biotechnology*, 34(8), 828-837.

Wetzel, W. C., & Whitehead, S. R. (2020). The many dimensions of phytochemical diversity: linking theory to practice. *Ecology letters*, 23(1), 16-32.

4.6 Figures

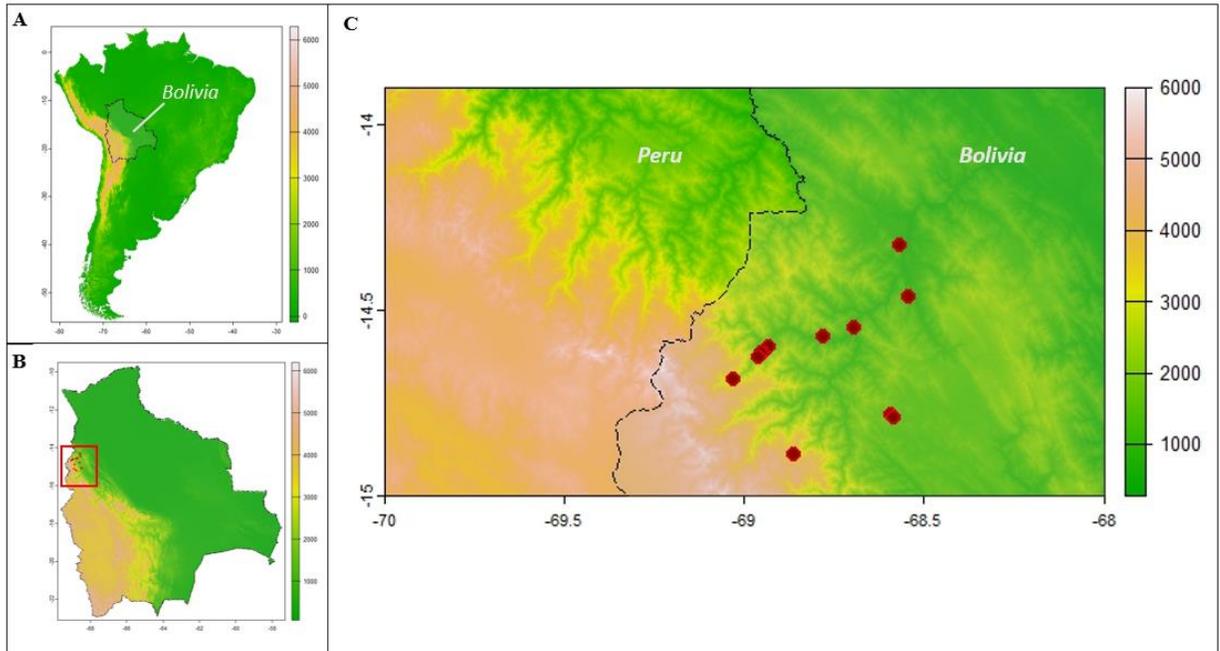


Figure 1 - Map of the 16 1-ha forest plots in the Bolivian Andes used in these analyses. A) study region location of Bolivia within the Cono del Sur, B) study region within Bolivia, and C) distribution of plots along the eastern slopes of the Andes (~356 to 3,328 masl) within Madidi National Park (part of the Madidi region) that form the elevational gradient. The x axes show longitude, and the y axes show latitude in decimal degrees.

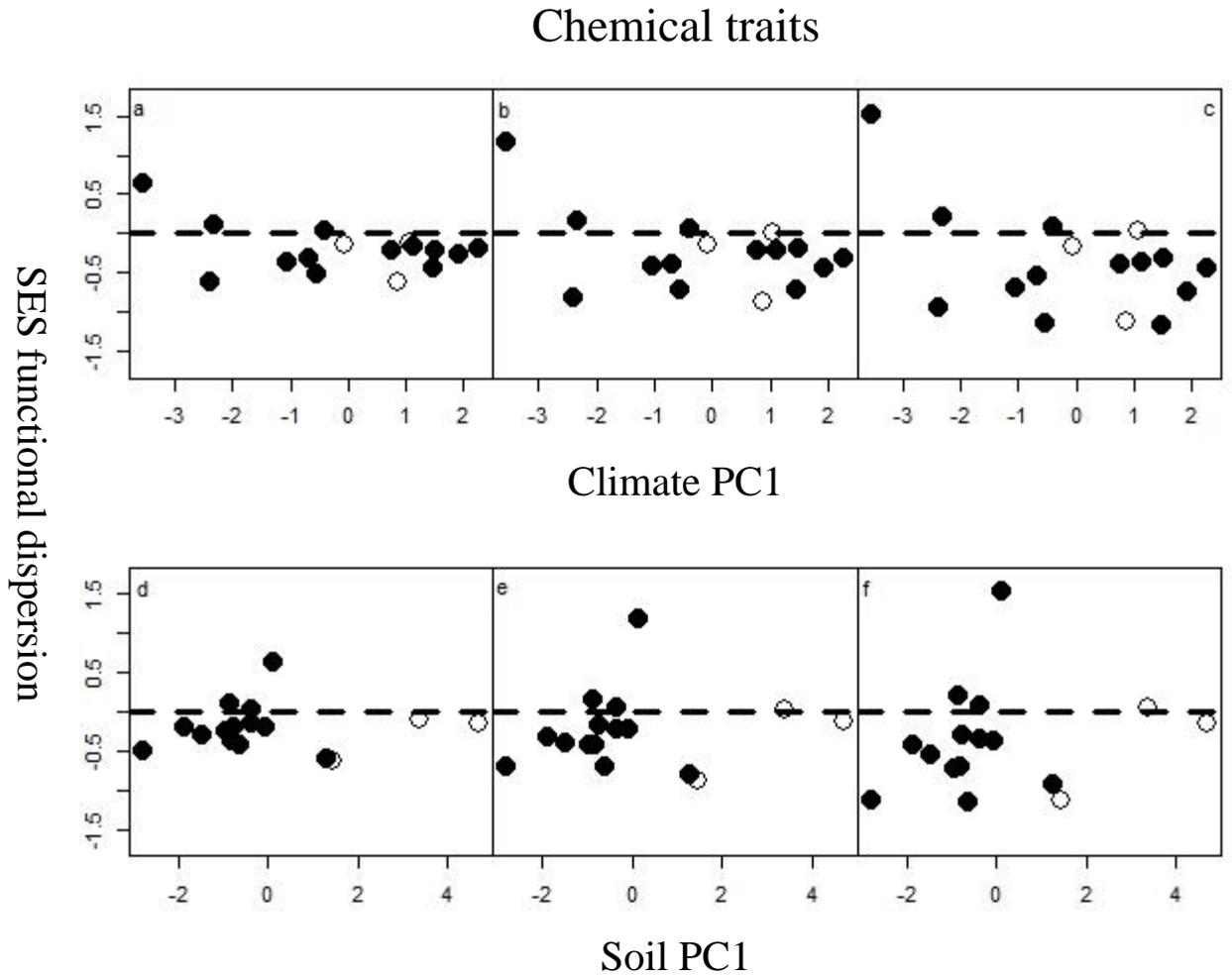


Figure 2 – SES of relationships of functional dispersion of chemical traits (secondary metabolites) across gradients: climate PC1 (top row) and soil PC1 (bottom row), at 3 different subplot sizes: 5m (left column), 10m (middle column), and 20m (right column). Filled points are moist montane forest plots. Open points are dry seasonal forest plots.

Morphological traits

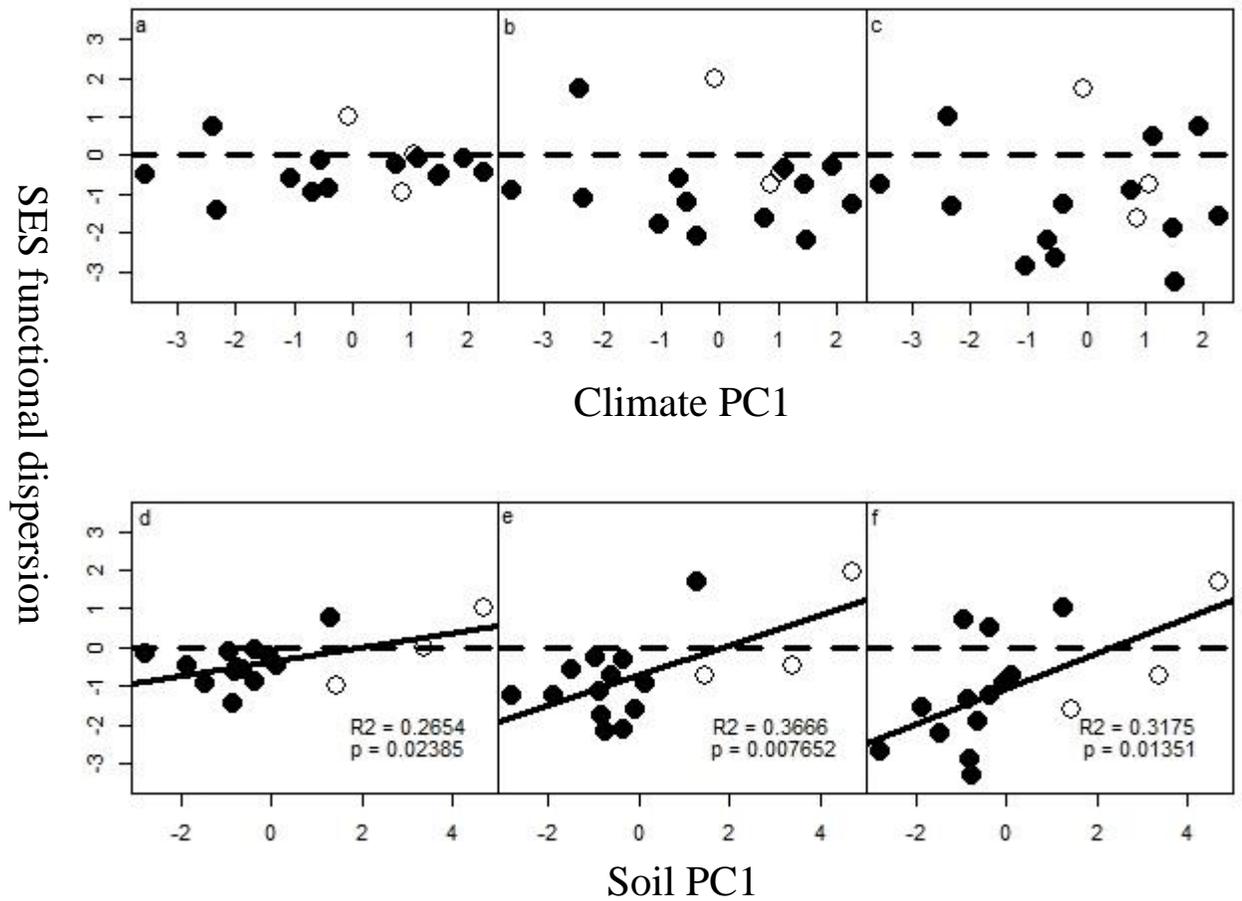


Figure 3 – SES of relationships of functional dispersion of morphological traits across gradients: climate PC1 (top row) and soil PC1 (bottom row), at 3 different subplot sizes: 5m (left column), 10m (middle column), and 20m (right column). Filled points are moist montane forest plots. Open points are dry seasonal forest plots.

Chapter 5: Conclusion

5.1 Conclusion

The three studies that comprise this dissertation investigated how biotic interactions structure diversity at large scales and small, and examined their role in maintaining diversity gradients. This dissertation contributes by addressing the conceptual gaps in our understanding of how local biotic interactions vary in their relative importance for community assembly across large-scale diversity gradients. It bridges this gap in three important ways: 1) by testing the relative importance of antagonistic species interactions among tree species across an elevational-diversity gradient in the tropical Andes (Chapters 2, 3), 2) by testing the relative importance of one particular type of antagonistic species interaction (apparent competition via chemically-mediated niche differences) across the same elevational-diversity gradient (Chapters 3, 4), and 3) by exploring the interaction between the biotic and abiotic environments using chemical and morphological traits (Chapter 4).

This research helps to reconcile the links between local-scale biotic interactions and large-scale diversity patterns. This work adds to recent works (Schemske et al. 2009) that have investigated the seminal hypotheses (Darwin 1859, Wallace 1878, Dobzhansky 1950, Erlich & Raven 1964, Janzen 1970, Connell 1971) that have long predicted the role that biotic interactions play in community structure, even across very large scales. The second chapter found that local biotic interactions become more stochastic w/ increasing diversity, suggesting that competitive interactions appear to be more unpredictable among neighborhoods of the same species in

higher-diversity tree communities. That patterns were not detected in the functional traits suggests that there are other forces at play than competitive biotic interactions.

Chapter 3 investigated how multi trophic, enemy mediated interactions varied over a gradient of climate, diversity, and elevation. The results showed that natural enemies impose a stronger selective pressure on plant chemical defenses in more diverse communities and in more productive climates, and that the defensive chemical compounds that hosts use are less phylogenetically conserved in higher diversity communities. The differential strength of biotic interactions and the greater evolutionary lability of defensive metabolites across communities implies that these enemy mediated interactions likely play a role in the maintenance and origin of biodiversity gradients.

While both chapters 2 and 3 both suggest that biotic interactions are a key factor in community structure, chapter 4 tested how community assembly is affected by both the biotic and the abiotic environments, by examining the relationships between two suites of plant functional traits, and abiotic site characteristics, finding evidence supporting systematic variation in the biotic and abiotic factors that structure communities. The finding that the abiotic environment might have a stronger than expected influence on chemical trait patterns suggests that a logical next step would be to further differentiate the chemical defensive compounds to quantify their individual influences on community assembly and dynamics.

Chapters 3 and 4 both utilized recent advances in chemical metabolomics, to become the first of few studies that have examined chemical metabolite patterns at a community scale. The use of these advanced tools allows for examination of community interaction patterns at a very fine scale. The results of these chapters as well as some recent studies (Salazar et al. 2016, Sam et al. 2020), show that this level of precision is necessary to accurately gauge the local

relationships working within communities, as individual classes of chemical-defensive metabolites vary according to different biotic and abiotic influences (Sam et al. 2020, Volf et al. 2020, 2022). Shedding light up how local interactions scale up and affect very large regions is one of the principal contributions of this work. Understanding the link between local scale interactions and large-scale patterns is key for finding ties between complimentary research and for protecting and conserving biodiversity by most accurately predicting future changes between species. The link between scales makes the importance of local patterns in predicting large-scale changes clear.

The implications of this research are potentially far-reaching. In an ever-changing world, an understanding of natural systems becomes more important than ever. The pace of change has continued to increase in the 250-plus years since the original hypotheses were recorded. In order to conserve biodiversity and to accurately predict species' shifts and relationships changes, accessible information is a must, and the scaling up of local data provides prodigious leverage. These studies have forged new ground in the direction toward a brighter future and provide a key step along the path to stewardship of biodiversity based on solid principles.

5.2 Conclusion References

Connell, J. H. (1971). On the role of natural enemies in preventing competitive exclusion in some marine animals and in rain forest trees. *Dynamics of populations*, 298(312).

Darwin, C. (1859). On the origin of species. *published on*, 24.

Ehrlich, P. R., & Raven, P. H. (1964). Butterflies and plants: a study in coevolution. *Evolution*, 586-608.

Hubbell, S. P. Foster R.B. (1986) Biology, chance and history and the structure of tropical rain forest tree communities. *Community ecology*"(ed. Diamond, J. & Case, TJ), 314-329.

Janzen, D. H. (1970). Herbivores and the number of tree species in tropical forests. *The American Naturalist*, 104(940), 501-528.

Salazar, D., Jaramillo, A., & Marquis, R. J. (2016). The impact of plant chemical diversity on plant–herbivore interactions at the community level. *Oecologia*, 181(4), 1199-1208.

Sam, K., Koane, B., Sam, L., Mrazova, A., Segar, S., Volf, M., ... & Novotny, V. (2020). Insect herbivory and herbivores of *Ficus* species along a rain forest elevational gradient in Papua New Guinea. *Biotropica*, 52(2), 263-276.

Schemske, D. W., Mittelbach, G. G., Cornell, H. V., Sobel, J. M., & Roy, K. (2009). Is there a latitudinal gradient in the importance of biotic interactions?. *Annu. Rev. Ecol. Evol. Syst.*, 40, 245-269.

Volf, M., Laitila, J. E., Kim, J., Sam, L., Sam, K., Isua, B., ... & Segar, S. T. (2020). Compound specific trends of chemical defences in *Ficus* along an elevational gradient reflect a complex selective landscape. *Journal of chemical ecology*, 46(4), 442-454.

Volf, M., Volfová, T., Hörandl, E., Wagner, N. D., Luntamo, N., Salminen, J. P., & Sedio, B. E. (2022). Abiotic stress rather than biotic interactions drives contrasting trends in chemical richness and variation in alpine willows. *Functional Ecology*, 36(11), 2701-2712.

Appendix A:

Supplementary table and figures to Chapter 2

Tables

Table S1 – Summary of eight plant functional traits measured on 701 tree species in 31 forest plots, including their ecological function (Pérez-Harguindeguy et al. 2013), unit of measurement, and the range of trait values from all plots.

Functional trait	Ecological function	Units	Range
Maximum height	Associated with growth form, position of the species in the vertical light gradient of the vegetation, competitive vigor, reproductive size, whole-plant fecundity, potential lifespan, and whether a species is able to establish and attain reproductive size between two disturbance events	meters	1.6 - 40
Maximum diameter at breast height (DBH)	An alternative, workable, proxy for height	centimeters	10 - 132
Relative growth rate (RGR)	Indicates plant strategy with respect to environmental productivity (soil nutrients, light moisture)	relative measure - unitless	- 0.124539788 - 0.17521802
Specific leaf area (SLA)	Associated with potential relative growth rate and abiotic environment. In general, species in resource-rich	g / m ²	0.8705576 – 242.23061

	environments have high SLA, while those in resource poor or harsh abiotic environments have low SLA. The leaf is also an organ with high investment in important secondary compounds		
Leaf thickness	Associated with abiotic environment, where sun leaves tend to be thicker than shade leaves as well as lower in N%, slower in CO ₂ diffusion and subject to more internal shading of chloroplasts	centimeters	0.004800000 - 0.1502
Leaf size	Represents a compromise between functional (growth) and resource use efficiency.	millimeters	88.320 - 491553.40
Twig specific density	Similar to wood density but can be measured for herbaceous species. Twig specific density is associated with stability, defense, architecture, hydraulics, C gain and growth potential of plants. Stem density partly underlies the growth-survival tradeoff; a low stem density (with large vessels) leads to a fast growth, because of cheap volumetric construction costs and a large hydraulic capacity, whereas a high stem density (with small	change in water volume of a 2.5cm stem cutting (g / g * cm ³)	0.01868481 - 1.996000

	vessels) leads to a high survival, because of biomechanical and hydraulic safety, resistance against pathogens, herbivores, or physical damage		
Relative twig bark thickness	Thicker bark insulates and protects buds from high temperatures. Thick bark may also provide protection of vital tissues against attack by pathogens, herbivores, frost or drought. In general, this trait has special relevance in trees or large shrubs subject to surface-fire regimes.	relative measure - unitless	0.01351351 – .8781250

Table S2 – Plot data of plots used in the analyses including plot type, plot richness, plot elevation, mean annual temperature, annual precipitation, dimensions, total number of stems (an individual may have multiple stems), average DBH and plot shape.

Plot Name	Forest Type	Plot Richness	Elevation (m)	Mean Annual Temperature	Annual Precipitation (mm)	Plot Dimension X (m)	Plot Dimension Y (m)	Total Stems	Plot mean DBH (cm)
Chaqui 31	moist, montane	38	2909	14.18333	1104	100	100	805	16.62085

Chaqui 32	moist, monta ne	34	3101	11.91667	975	100	100	780	17.608 7
Collpa 48	moist, monta ne	95	1259	20.59167	1449	100	100	848	20.507 45
Fuerte 26	moist, monta ne	73	2029	17.59583	1324	100	100	879	17.380 76
Fuerte 27	moist, monta ne	91	1853	17.975	1350	100	100	733	18.510 66
Kanup a 44	moist, monta ne	16	3334	10.72917	961	100	100	466	23.929 09
Lomak a 40	moist, monta ne	140	1268	21.00417	1332	100	100	791	19.480 73
Lomas a 39	moist, monta ne	102	1049	21.52917	1383	100	100	883	16.954 97
Resina 12	dry, season al	49	724	23.0875	1858	100	100	806	17.042 15
Resina 13	dry, season al	42	869	22.57083	1806	100	100	775	16.723 41
Resina 14	dry, season al	61	1099	21.70417	1696	100	100	876	16.156 55
Sanma r 21	moist, monta ne	89	1195	21.04583	1803	100	100	635	22.743 43
Sanma r 22	moist, monta	90	1328	20.4	1761	100	100	521	22.660 89

	ne								
Santaa 36	moist, monta ne	41	2490	14.56667	1114	100	100	527	21.251 03
Sumpu 133	moist, monta ne	71	1086	21.31667	1610	100	100	614	20.350 37
Sumpu 134	moist, monta ne	97	1339	20.36667	1500	100	100	745	19.741 93
Tanhua 37	moist, monta ne	94	1976	16.25	1206	100	100	703	19.387 07
Tapuri 45	moist, monta ne	43	2708	15.20417	1013	100	100	702	21.032 23
Tapuri 46	moist, monta ne	44	2362	16.44583	1062	100	100	1111	14.592 59
Terraz 41	moist, monta ne	32	2887	13.54583	934	100	100	894	17.693 02
Tintay 24	moist, monta ne	112	1478	19.65833	1358	100	100	722	17.073 13
Titiri 42	moist, monta ne	34	2854	13.06667	945	100	100	631	20.361 23
Tocoaq 28	moist, monta ne	77	2210	16.84583	1288	100	100	724	18.037 68
Tocoaq 29	moist, monta ne	83	2407	15.6	1202	100	100	734	17.362 28

Tocoaq 30	moist, monta ne	60	2508	15.69167	1221	100	100	683	17.746 16
Waturu 43	moist, monta ne	24	3088	11.69167	972	100	100	642	22.202 67
Yarimi 10	dry, season al	48	1006	21.63333	1628	100	100	610	18.656 27
Yarimi 11	dry, season al	55	1197	21.81667	1647	100	100	579	22.284 02
Yarimi 9	dry, season al	40	927	22.175	1698	100	100	433	23.563 42

Figures

Functional Trait Correlations

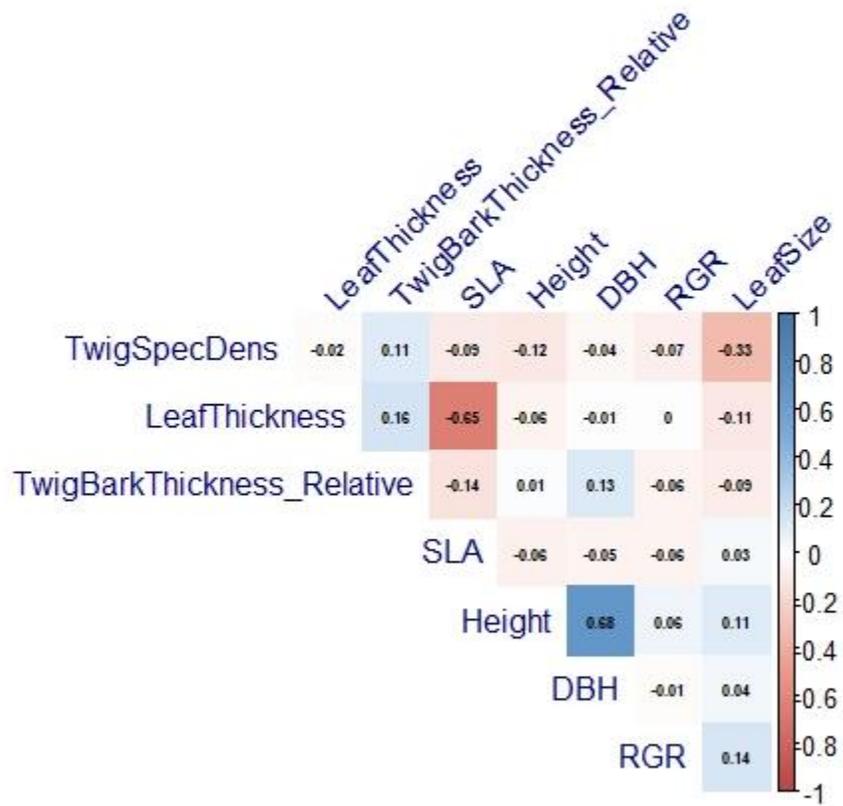


Figure S1 – Bivariate trait correlations (Pearson correlation coefficient) for the 8 selected traits.

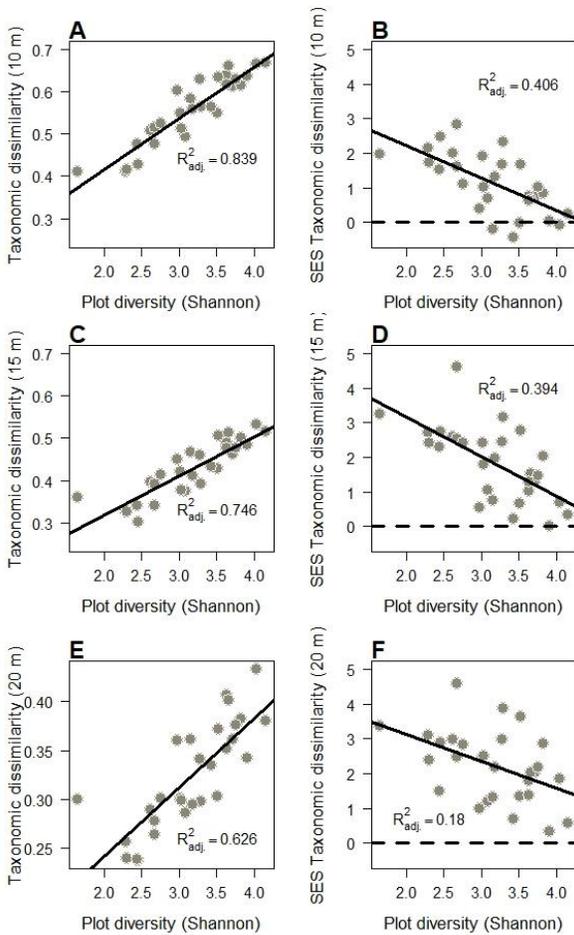


Figure S2 – Mean taxonomic dissimilarity of tree neighborhoods across the diversity gradient. The left panels show relationships between the mean observed taxonomic dissimilarity (Bray Curtis dissimilarity) of conspecific tree neighborhoods and species diversity (Shannon diversity Index; N = 29 forest plots) at A) 10-m, C) 15-m, and E) 20-m neighborhood radii. The right panels show the same relationships for null-model deviations (mean standardized effects sizes, SES). The dashed line shows the null expectation (SES = 0); positive and negative SES indicate higher and lower taxonomic dissimilarity than expected from random sampling of individuals from the plot species pool, respectively. Solid trend lines indicate significant linear relationships (P < 0.05).

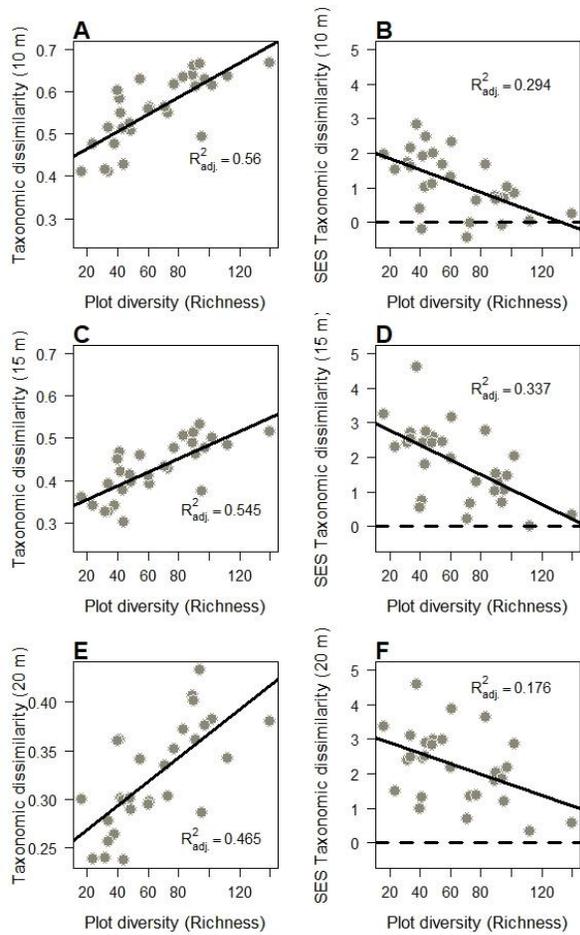


Figure S3 – Mean taxonomic dissimilarity of tree neighborhoods across the diversity gradient. The left panels show relationships between the mean observed taxonomic dissimilarity (Bray Curtis dissimilarity) of conspecific tree neighborhoods and species diversity (Plot species richness; $N = 29$ forest plots) at A) 10-m, C) 15-m, and E) 20-m neighborhood radii. The right panels show the same relationships for null-model deviations (mean standardized effects sizes, SES). The dashed line shows the null expectation ($SES = 0$); positive and negative SES indicate higher and lower taxonomic dissimilarity than expected from random sampling of individuals from the plot species pool, respectively. Solid trend lines indicate significant linear relationships ($P < 0.05$).

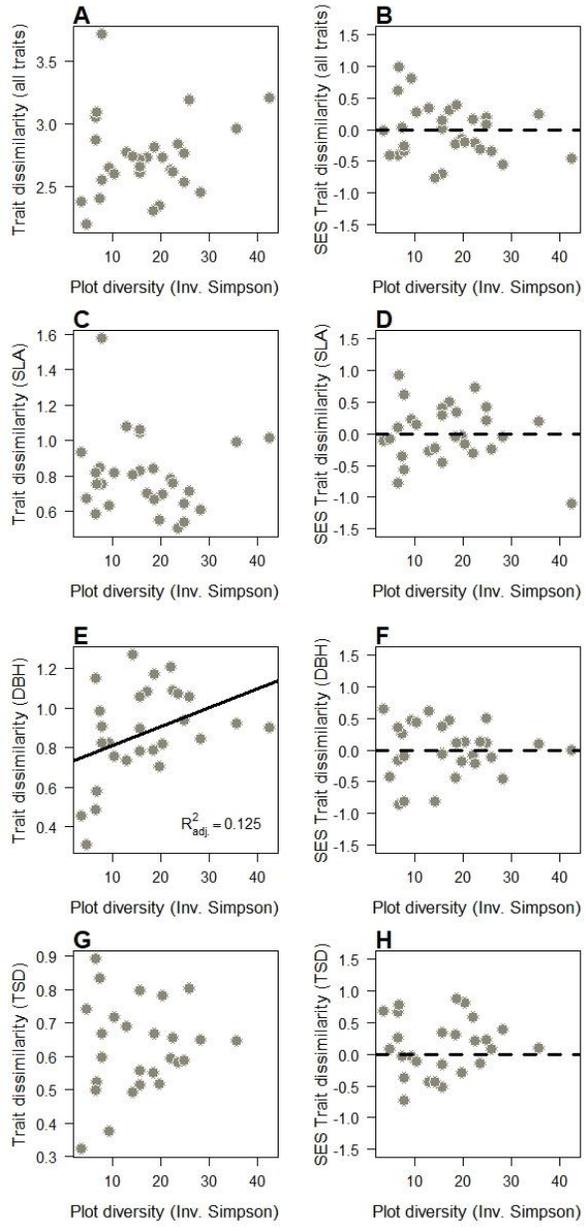


Figure S4 – Plot of functional traits at 10m scale corresponding to Figure 3.

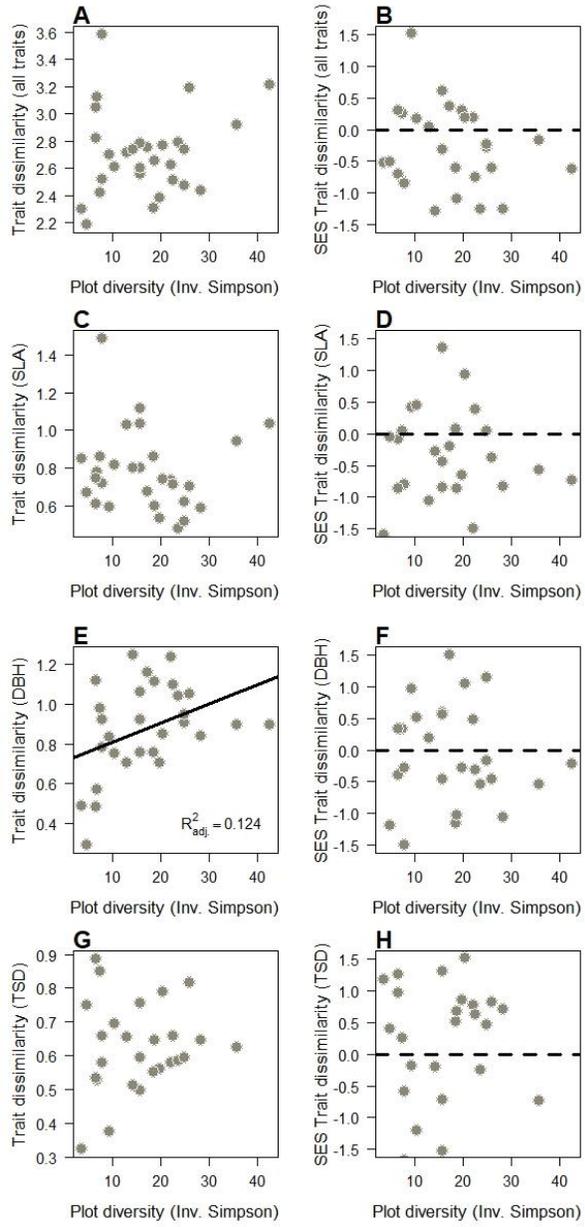


Figure S5 – Plot of functional traits at 20m scale corresponding to Figure 3.

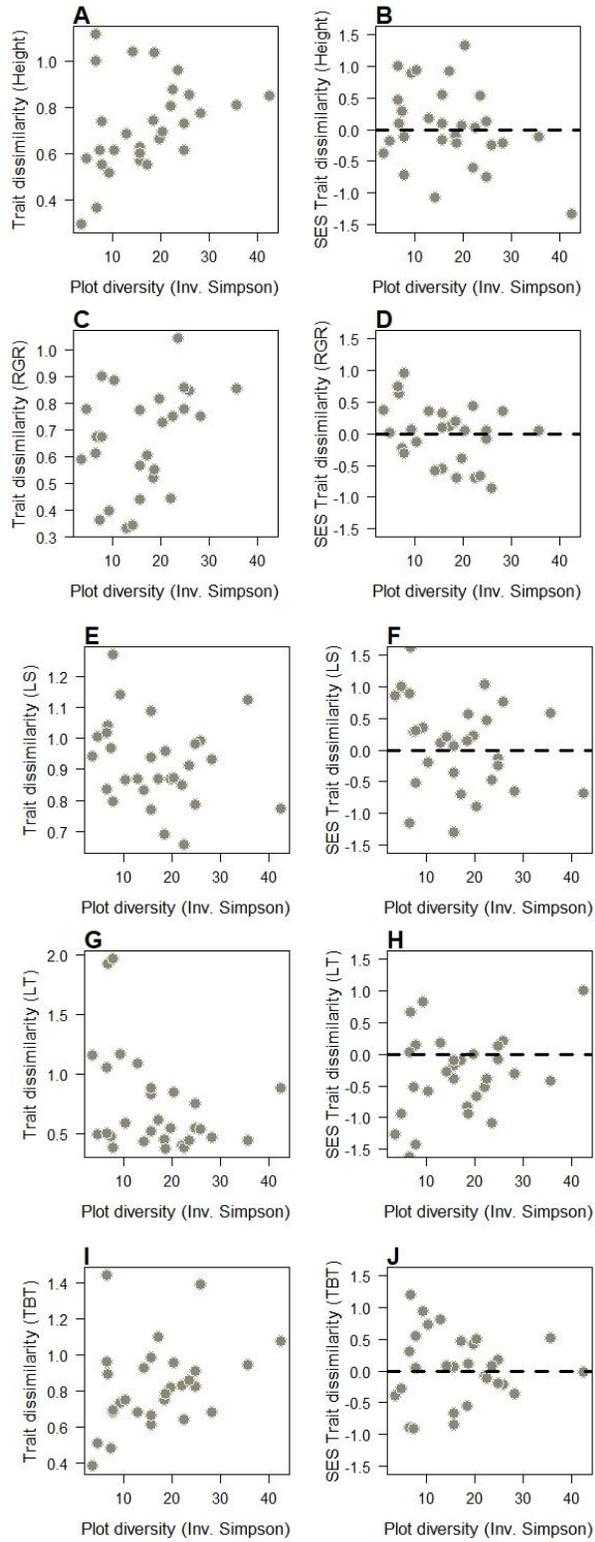


Figure S6 – Remaining functional traits at the 15m neighborhood radius scale.

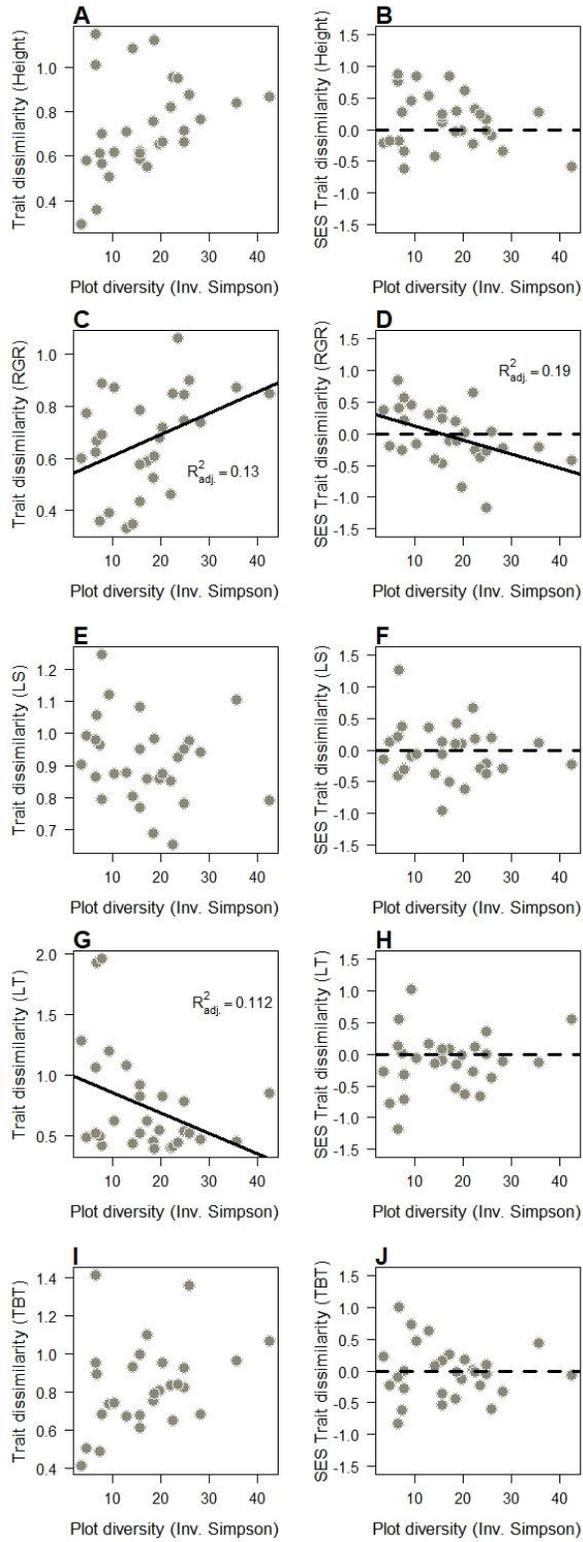


Figure S7 – Remaining functional traits at the 10m neighborhood radius scale.

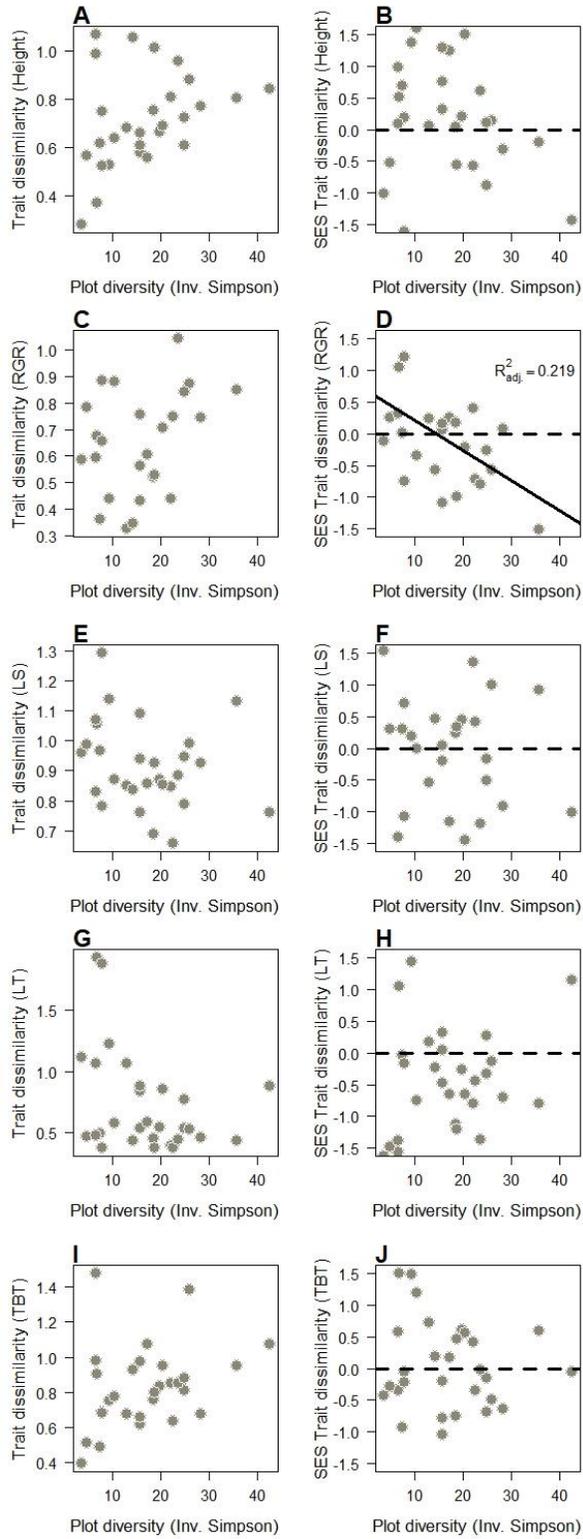


Figure S8 – Remaining functional traits at the 20m neighborhood radius scale.

Appendix B:

Supplementary figures to Chapter 3

Figures

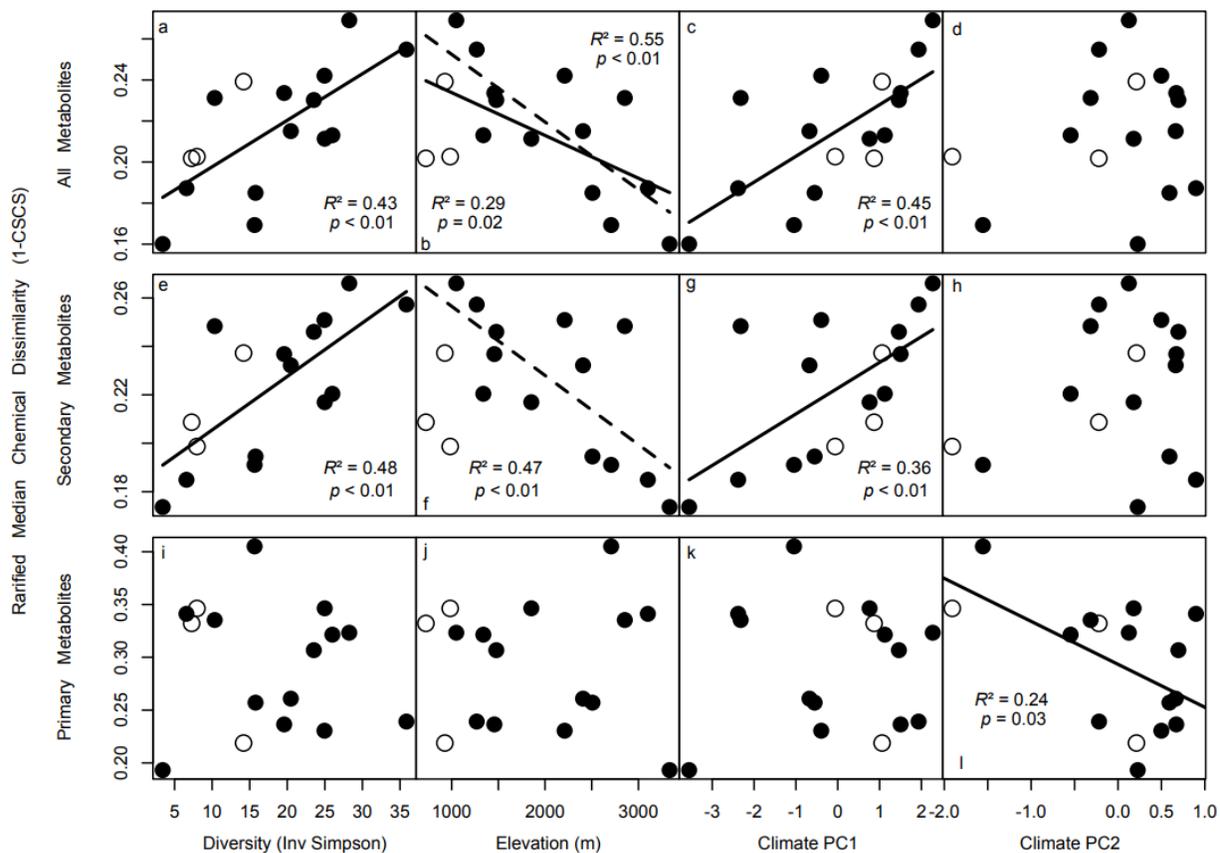


Figure S1 – Variation in rarified median chemical dissimilarity (1-CSCS) vs species diversity (inverse Simpson index), elevation (m), and climate among 16 forest plots in Madidi, Bolivia. Panels a-d represent linear regressions between rarified ($n = 12$) median chemical dissimilarity among co-occurring species with respect to the whole metabolite and (a) species diversity, (b) elevation, (c) Climate PC1, and (d) Climate PC2, respectively. Panels e-h represent linear regressions between chemical dissimilarity with respect to secondary metabolites and (e) species diversity, (f) elevation, (g) Climate PC1, and (h) Climate PC2, respectively. Panels i-l represent linear regressions between chemical dissimilarity with respect to primary metabolites and (i) species diversity, (j) elevation, (k) Climate PC1, and (l) Climate PC2, respectively. Secondary metabolites are defined as those derived from the Alkaloids, Amino acid and Peptides, Polyketides, Shikimates and Phenylpropanoids, and Terpenoids biosynthetic pathways using NPClassifier (Kim et al. 2022). Primary metabolites are defined as those derived from the Carbohydrates and Fatty acids pathways. Three seasonally dry forests are represented by open

circles. Regressions using all 16 forest plots are represented by solid lines; regressions excluding three seasonally dry forests are represented by dashed lines. Adjusted R^2 and p -values are presented for significant ($p < 0.05$) regressions only.

Appendix C:

Supplementary table and figures to Chapter 4

Tables

Table S1 - Overview of forest plots. Variation in tree species richness, elevation, total annual precipitation (MAP), and mean annual temperature (MAT) among 16 1-ha forest plots in the Madidi Project, Bolivia. Dry forest plots *italicized*.

Plot Name	Species richness	Elevation (m)	MAP (mm)	MAT (°C)	N trees sampled
Chaqui_32	35	3116	975	11.9	85
Fuerte_27	80	1900	1350	18	236
Kanupa_44	17	3324	961	10.7	52
Lomaka_40	137	1242	1332	21	248
Lomasa_39	95	1054	1383	21.5	250
<i>Pintat_5</i>	48	880	1684	22	164
<i>Resina_12</i>	50	662	1858	23.1	141
Sumpul_34	103	1223	1500	20.4	245
Tapuri_45	43	2697	1013	15.2	121
Tintay_24	117	1400	1358	19.7	283
Tintay_25	94	1468	1357	20	253
Titiri_42	34	2859	945	13.1	106
Tocoaq_29	83	2407	1202	15.6	170
Tocoaq_30	62	2510	1221	15.7	133
Tocoaq_28	81	2200	1288	16.8	209
<i>Yarimi_9</i>	40	850	1698	22.2	129

Table S2– Environmental climate (n=4) and soil (n=13) variables used in the analyses.

Climate Variable	Soil Variable
Annual Mean Temperature (bio01_WC2.1)	pH Dissolved
Temperature Annual Range (bio07_WC2.1)	Total Nitrogen
Annual Precipitation (bio12_TRMM)	Organic Carbon
Precipitation Seasonality (bio15_TRMM)	Organic Material
	Available Phosphorus
	Exchangeable Sodium
	Exchangeable Potassium
	Exchangeable Calcium
	Exchangeable Magnesium
	Cation Exchange Capacity (CEC)
	Percentage Sand
	Percentage Silt
	Percentage Clay

Figures

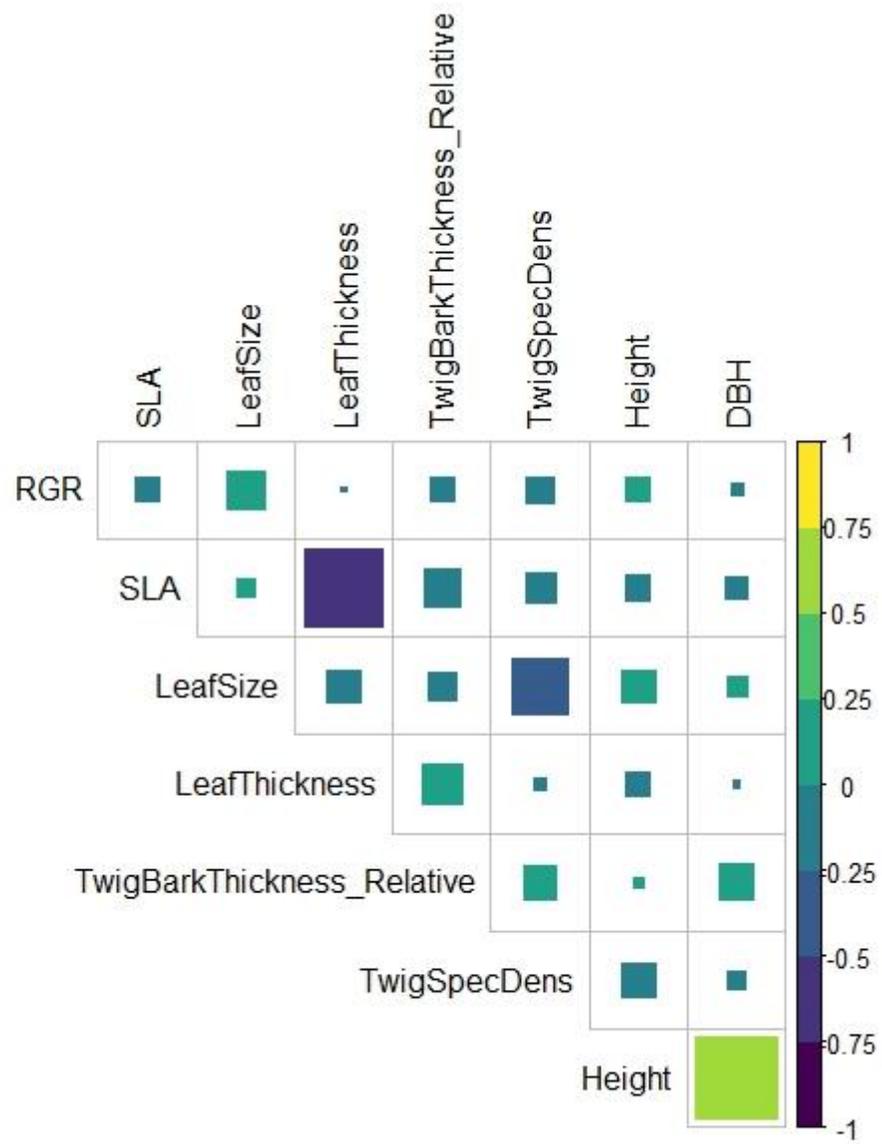


Figure S1 – Bivariate, Pearson correlations between each of the 8 selected morphological traits.

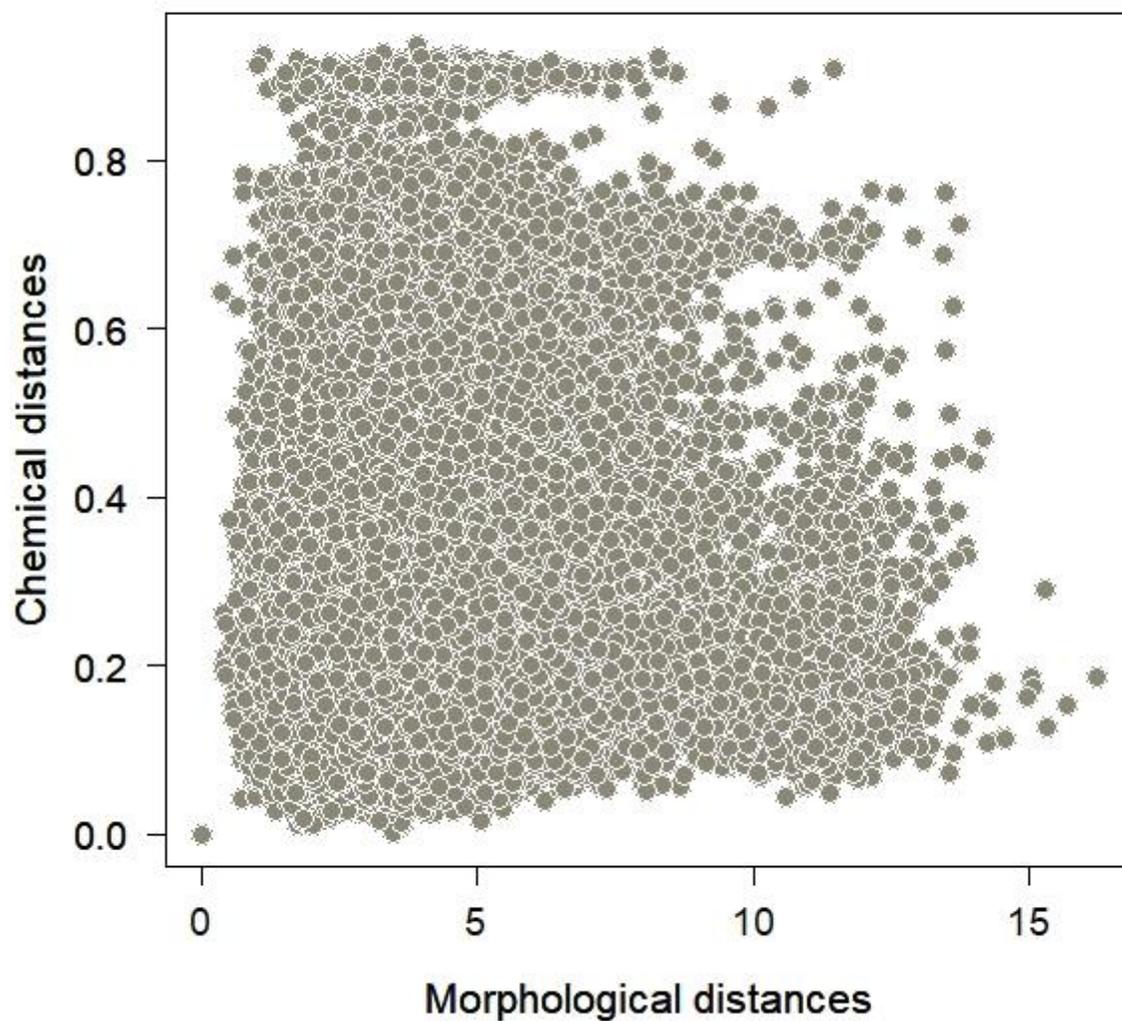


Figure S2: Relationship between chemical distances and morphological distances for the entire dataset (n=891).

Chemical traits

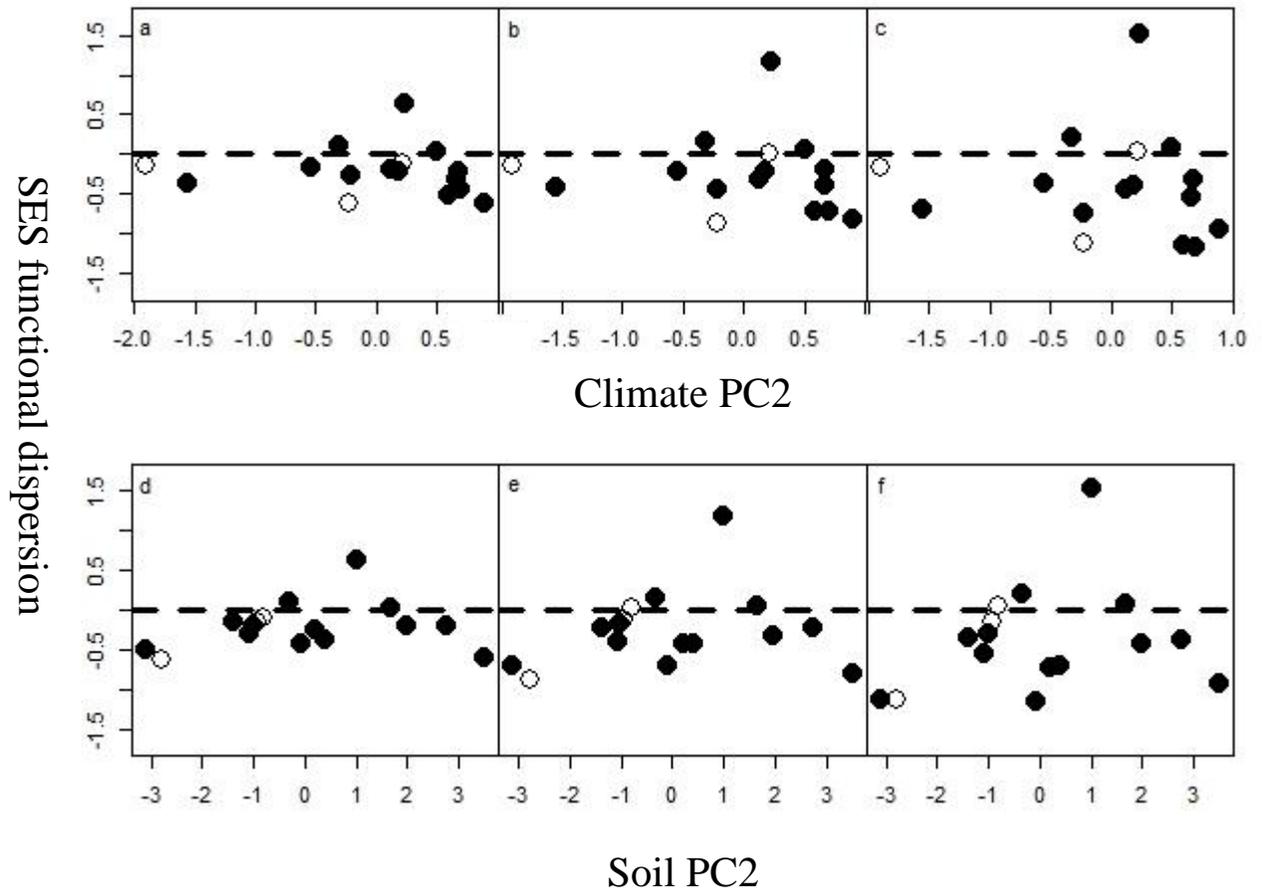


Figure S3 – SES relationships functional dispersion of chemical traits (secondary metabolites) across gradients: climate PC2 (top row) and soil PC2 (bottom row), at 3 different subplot sizes: 5m (left column), 10m (middle column), and 20m (right column). Filled points are moist montane forest plots. Open points are dry seasonal forest plots.

Morphological traits

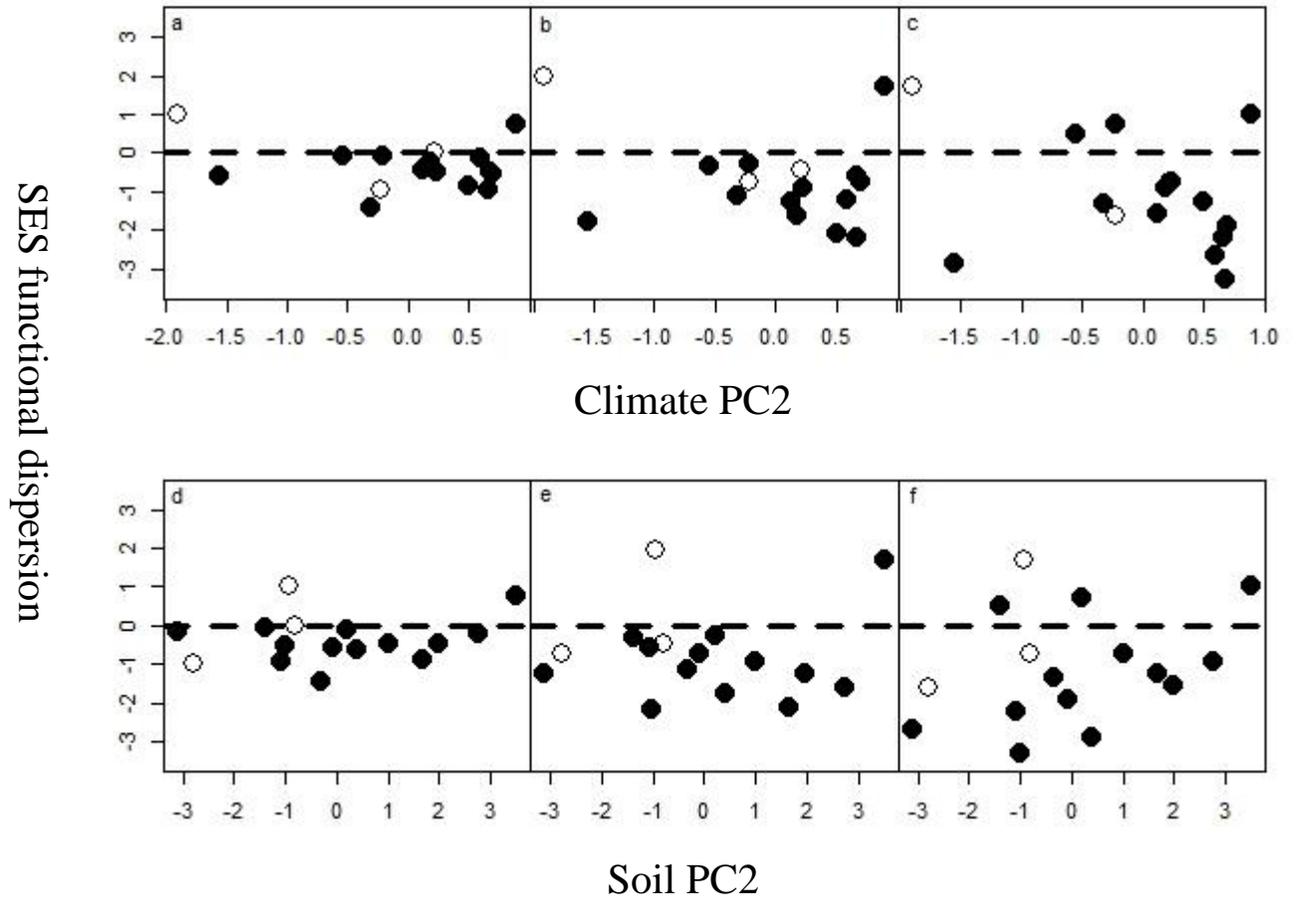


Figure S4 – SES relationships functional dispersion of morphological traits across gradients: climate PC2 (top row) and soil PC2 (bottom row), at 3 different subplot sizes: 5m (left column), 10m (middle column), and 20m (right column). Filled points are moist montane forest plots. Open points are dry seasonal forest plots.

Appendix D: R code

R code for Chapter 2

title: "Chapter 2 R code"

output: word document

Functions for analyses

``{r}

print("Yes!")

FUNCTION 1

neig.sim.sp <- function(sp.s.name, plot.p, trait.dists,

neig.limit, dists.p, which.close.edge, show.animation=FALSE)

{

Open packages necessary

require(vegan)

require(plotrix)

require(picante)

require(FD)

diag(dists.p) <- NA # Ensures that the focal individual is not selected

as part of its own neighborhood

which.sp.s <-

```
which(as.character(plot.p$BinomialSpeciesName)==sp.s.name) # which individuals  
_____ # in the plot belong  
_____ # to species  
_____ # "sp.s.name"
```

```
neig.n <- length(which.sp.s) # the number of individuals of species  
_____ # 'sp.s.name' in the plot
```

```
which.sp.s <- setdiff(which.sp.s, which.close.edge) # which individuals in the  
_____ # plot are usable because  
_____ # they are not too close  
_____ # to the edge
```

```
usable.neig.n <- length(which.sp.s) # number of trees for which neighborhoods  
_____ # can be estimated
```

```
if(show.animation==TRUE)  
{  
plot(plot.p$Y~plot.p$X, pch=21, bg="grey80", col="grey80", asp=1,  
ylim=range(plot.p$Y, na.rm=TRUE), xlim=range(plot.p$X, na.rm=TRUE),  
ylab="Y", xlab="X", main=sp.s.name)  
  
points(plot.p$Y[which.sp.s]~plot.p$X[which.sp.s], pch=21,  
bg="forestgreen", col="forestgreen", cex=1.5)
```

```

for(j in 1:length(which.sp.s))
  draw.circle(x=plot.p$X[which.sp.s[j]], y=plot.p$Y[which.sp.s[j]],
    radius=neig.limit, nv=100, border="black", col=NA, lty=1, lwd=1)
}

## Initiates the result objects as all NAs
ind.n <- ind.n.hetero <- ind.n.conspecific <-
rich <- rich.hetero <- NA

BC <- Jac <- BC.hetero <- Jac.hetero <-
trait.mpd <- trait.mpd.hetero <- NA

## Calculations are run only if there are more than 2 usable neighborhoods
if(usable.neig.n >= 2)
{

  dists.s.in.p <- dists.p[which.sp.s,] # distances for individuals that can
  # be used

  ind.names <- rownames(plot.p)[which.sp.s] # individuals that could be used
  if(identical(rownames(dists.s.in.p), ind.names)==FALSE)
  stop() # test that the distance matrix and the tree table match
}

```

```

## Creates empty tables for the community composition by neighborhood
spp.in.plot <- sort(unique(plot.p$BinomialSpeciesName))
compo <- matrix(0, nrow=usable.neig.n, ncol=length(spp.in.plot))
colnames(compo) <- spp.in.plot
rownames(compo) <- ind.names

## Loops through all usable neighborhoods
for(i in 1:usable.neig.n)
{

  dists.to.i <- dists.s.in.p[i,] # Isolate distances of all individuals to
  # individual 'i'.

  which.neig.i <- which(dists.to.i <= neig.limit) # Determines which
  # individuals are within the
  # neighborhood, meaning
  # 'neig.limit' away or
  # closer to individual 'i'.

  All.i <- plot.p[which.neig.i,] # Reduces the plot data to only those
  # rows corresponding to all individuals
  # within the neighborhood of
  # individual 'i'

```

```

## Creates a composition table for all individuals in the neighborhood of 'i'
compo.i <-
as.matrix(table(as.character(All.i$BinomialSpeciesName)))

match.names <-
match(rownames(compo.i), colnames(compo))

compo[i, match.names] <- compo.i
}

## For compo matrix with only heterospecifics, all conspecifics are 0
compo.hetero <- compo
compo.hetero[,sp.s.name] <- 0

## Eliminates columns (species) with zero individuals
compo <- compo[,colSums(compo)>0]
compo.hetero <- compo.hetero[,colSums(compo.hetero)>0]

## For each neighborhood, calculates numbers of individuals and richness
ind.n <- rowSums(compo)
ind.n.hetero <- rowSums(compo.hetero)
ind.n.conspecific <- ind.n - ind.n.hetero

rich <- rowSums(compo>0)

```

```

rich.hetero <- rowSums(compo.hetero>0)

## Calculates taxonomic beta-diversity across neighborhoods
BC <- vegdist(x=compo, method="bray", binary=FALSE)
Jac <- vegdist(x=compo, method="jaccard", binary=FALSE)

BC.hetero <- vegdist(x=compo.hetero, method="bray", binary=FALSE)
Jac.hetero <- vegdist(x=compo.hetero, method="jaccard", binary=FALSE)

## Calculates functional beta-diversity across neighborhoods

trait.dists <- as.matrix(trait.dists)

shared.spp.1 <- intersect(colnames(compo), colnames(trait.dists))

if(length(shared.spp.1) >= 2)
{
  trait.dists.1 <- as.dist(trait.dists[shared.spp.1, shared.spp.1])
  compo.1 <- compo[,shared.spp.1]

  trait.mpd <-
  comdist(comm = compo.1,

```

```

dis = trait.dists.1, abundance.weighted = TRUE)
}

shared.spp.2 <- intersect(colnames(compo.hetero), colnames(trait.dists))

if(length(shared.spp.2) >= 2)
{
trait.dists.2 <- as.dist(trait.dists[shared.spp.2, shared.spp.2])
compo.hetero.2 <- compo.hetero[,shared.spp.2]

trait.mpd.hetero <-
comdist(comm = compo.hetero.2,
dis = trait.dists.2, abundance.weighted = TRUE)
}
}

## Combines results for each neighborhood into a table
neigh.res.table <-
cbind(ind.n, ind.n.hetero, ind.n.conspecific,
rich, rich.hetero)
colnames(neigh.res.table) <- c("ind.n", "ind.n.hetero",
"ind.n.conspecific", "rich", "rich.hetero")

## Save results of neighborhood pairwise comparisons into table

```

```

pairwise.res.table <- cbind(as.numeric(BC),
  as.numeric(Jac), as.numeric(BC.hetero), as.numeric(Jac.hetero),
  as.numeric(trait.mpd), as.numeric(trait.mpd.hetero))

colnames(pairwise.res.table) <- c("BC",
  "jaccard", "BC.hetero", "jaccard.hetero",
  "trait.mpd", "trait.mpd.hetero")

res.summary <- c(neig.n, usable.neig.n,
  colMeans(neigh.res.table, na.rm=TRUE),
  colMeans(pairwise.res.table, na.rm=TRUE))

names(res.summary) <- c("neig.n", "usable.neig.n",
  paste("mean.", colnames(neigh.res.table), sep=""),
  paste("mean.", colnames(pairwise.res.table), sep=""))

output <- list(res.summary, pairwise.res.table, neigh.res.table)
names(output) <- c("res.summary", "pairwise.res.table", "neigh.res.table")

output
}

```

```

### FUNCTION 2 ###

neig.sim.plot <- function(plot.p, trait.dists, neig.limit,
  rand.n=10, SAD="fixed", show.animation=FALSE)
{
  require(R.utils)

  if(is.factor(plot.p$BinomialSpeciesName)==FALSE)
    stop("'BinomialSpeciesName' needs to be a factor")

  spp.list <- unique(as.character(plot.p$BinomialSpeciesName))
  spp.n <- length(spp.list)

  dists.p <- as.matrix(dist(plot.p[,c("X", "Y")]))
  diag(dists.p) <- NA

  ## NOTE: this assumes a rectangular plot
  plot.y.range <- range(plot.p$Y, na.rm=TRUE)
  plot.x.range <- range(plot.p$X, na.rm=TRUE)

  which.close.edge <- which(plot.p$Y<plot.y.range[1]+neig.limit |
  plot.p$Y>plot.y.range[2]-neig.limit |
  plot.p$X<plot.x.range[1]+neig.limit |
  plot.p$X>plot.x.range[2]-neig.limit)

  elements.to.randomize <-

```

```

c("mean.ind.n", "mean.rich",
"mean.BC", "mean.jaccard",
"mean.BC.hetero", "mean.jaccard.hetero",
"mean.trait.mpd", "mean.trait.mpd.hetero")

pb <- txtProgressBar(min=0, max=spp.n, style=3)

for(s in 1:spp.n)
{
#print(paste("Species", s, "of", spp.n))
setTxtProgressBar(pb, s)

sp.s.name <- spp.list[s]

which.sp.s <- which(as.character(plot.p$BinomialSpeciesName)==sp.s.name)
which.not.sp.s <- setdiff(c(1:nrow(plot.p)), which.sp.s)

emp.results.sp <- neig.sim.sp(sp.s.name=sp.s.name, plot.p=plot.p,
trait.dists=trait.dists, neig.limit=neig.limit, dists.p=dists.p,
which.close.edge=which.close.edge, show.animation=show.animation)

emp.results.sp <- emp.results.sp$res.summary

```

```

__ if(s==1)
__ {
__   spp.results.names <- c(names(emp.results.sp),
__   paste("rand.", elements.to.randomize, sep=""),
__   paste("SES.", elements.to.randomize, sep=""))

__   spp.results <- as.data.frame(
__   matrix(NA, ncol=length(spp.results.names), nrow=spp.n))

__   colnames(spp.results) <- spp.results.names
__   rownames(spp.results) <- spp.list
__ }

__ rand.results.table <- as.data.frame(
__   matrix(NA, nrow=rand.n, ncol=length(emp.results.sp)))
__   colnames(rand.results.table) <- names(emp.results.sp)

__ if(FALSE %in% (elements.to.randomize %in% colnames(rand.results.table)))
__   stop("'elements.to.randomize' are not all part of the results from fuction 'neig.sim.sp'")

__ #Creates fixed and randomized null models
__ for(r in 1:rand.n)
__ {
__   rand.plot.p <- plot.p

```

```

if(SAD=="fixed")
  rand.plot.p$BinomialSpeciesName[which.not.sp.s] <-
  sample(rand.plot.p$BinomialSpeciesName[which.not.sp.s])

if(SAD=="randomized")
  rand.plot.p$BinomialSpeciesName[which.not.sp.s] <-
  sample(unique(rand.plot.p$BinomialSpeciesName[which.not.sp.s]),
  size=length(which.not.sp.s), replace="TRUE")

rand.which.sp.s <-
  which(as.character(rand.plot.p$BinomialSpeciesName)==sp.s.name)

if(identical(rand.which.sp.s, which.sp.s)==FALSE)
  stop("something wrong with randomization 1")

rand.results.sp <-
  neig.sim.sp(sp.s.name=sp.s.name, plot.p=rand.plot.p,
  trait.dists=trait.dists, neig.limit=neig.limit,
  dists.p=dists.p, which.close.edge=which.close.edge,
  show.animation=show.animation)

rand.results.table[r,] <- rand.results.sp$res.summary
}

rand.results.table <-

```

```

rand.results.table[,which(colnames(rand.results.table) %in%
elements.to.randomize)]

if(identical(colnames(rand.results.table), elements.to.randomize)==FALSE)
{
print(colnames(rand.results.table))
print(elements.to.randomize)
stop("'colnames(rand.results.table)' and 'elements.to.randomize' are not the same or in the
same order")
}

rand.means <- colMeans(rand.results.table, na.rm=TRUE)
rand.sd <- apply(rand.results.table, 2, sd, na.rm=TRUE)

#Calculates standard effect size
ESs <- emp.results.sp[which(names(emp.results.sp) %in%
colnames(rand.results.table))]-rand.means
SESs <- ESs/rand.sd

spp.results[s,] <- c(emp.results.sp, rand.means, SESs)
}
close(pb)

spp.results
}

```

```\n

### Tree maps and data preparation

```\n{r}

Load plot and individual data

```
plot.data <- read.csv(paste0(path.to.files,\n  ""02 Data/01 PlotData Clean v4.1 2020-09-07.txt""),\n  header = TRUE, sep = "\\t", fileEncoding="UTF-8")\n  \ndim(plot.data)
```

```
tree.data <- read.csv(paste0(path.to.files,\n  ""02 Data/04 TreeData PP Clean v4.1 2020-09-07.txt""),\n  header = TRUE, sep = "\\t", fileEncoding="UTF-8")\n  \ndim(tree.data)
```

```
tree.data <- tree.data[which(tree.data$IsAlive=="True"), ] #Denote only trees alive in the first\n  census, excludes new recruits
```

```
dim(tree.data)
```

Need to remove a single uncertain data point

```
tree.data <- tree.data[-which(tree.data$Gx > 1977), ]\n  \ndim(tree.data)
```

Trait Data

trait.data <- read.csv(

paste0(path.to.files,

"02_Data/06_TraitData_Clean_v4.1_2020-09-07.txt"),

header = TRUE, sep = "\t", fileEncoding="UTF-8")

dim(trait.data)

SELECT ONLY PLOTS WITH TRAIT DATA

#####

plots.not.to.use<-c("PP Heatht 8", "PP Heathi 7", "PP Hondo 1", "PP Tuichi 4",
"PP Chiriu 2", "PP Mamaco 3", "PP ", "PP Pintat 5")

plots.with.traits <- unique(trait.data\$PlotName)

intersect(plots.with.traits, plots.not.to.use)

plots.with.traits <- setdiff(plots.with.traits, plots.not.to.use)

plot.data.traits <- plot.data[plot.data\$PlotName %in% plots.with.traits,]

dim(plot.data)

dim(plot.data.traits)

```

tree.data.traits <- tree.data[tree.data$PlotName %in% plots.with.traits,]

dim(tree.data)

dim(tree.data.traits)

tree.coords <- as.data.frame(cbind(tree.data.traits$Gy, tree.data.traits$Gx))

colnames(tree.coords)<-c("Y", "X")

tree.data.traits <- tree.data.traits[,c("PlotName", "BinomialSpeciesName", "DetRank", "Family",
"Diameter_t2")]

tree.data.traits <- data.frame(tree.data.traits, tree.coords)

dim(tree.data.traits)

##### CALCULATE MEAN TRAITS BY SPECIES AND PLOT
#####

traits.to.use.1 <- c("RGR", "SLA", "LeafSize", "LeafThickness",
"TwigBarkThickness_Relative", "TwigSpecDens")

traits.to.use.2 <- c("Height", "DBH")

mean.traits <- aggregate(trait.data[,traits.to.use.1],
by = list(trait.data$BinomialSpeciesName, trait.data$PlotName),
FUN = "mean", na.rm = TRUE)

```

```

max.traits <- aggregate(trait.data[,traits.to.use.2],
  by = list(trait.data$BinomialSpeciesName, trait.data$PlotName),
  FUN = "quantile", probs=0.90, na.rm = TRUE)

traits.summary <- data.frame(mean.traits[,1:2],
  scale(log1p(mean.traits[,traits.to.use.1])),
  scale(log1p(max.traits[,traits.to.use.2])))

colnames(traits.summary)[1:2] <- c("BinomialSpeciesName", "PlotName")

head(traits.summary)

# Correlation of variables
pairs(traits.summary[,c(-1,-2)]) # visually see the correlations on plots

traits.cor<-cor(traits.summary[,c(-1,-2)][complete.cases(traits.summary[,c(-1,-2)])],) # Calculate
PeARSON'S R between variables

traits.cor

library(corrplot)

p.mat<-cor.mtest(traits.summary[,c(-1,-2)][complete.cases(traits.summary[,c(-1,-2)])],)#
calculate the P values for correlations

# do the correlation plot
corrplot(traits.cor,method="color",
  type="upper", order="hclust",
  addCoef.col = "black", # Add coefficient of correlation

```

```
tl.col="black", tl.srt=45, #Text label color and rotation  
# Combine with significance  
p.mat = p.mat$p, sig.level = 0.05,  
# hide correlation coefficient on the principal diagonal  
diag=FALSE )
```

```
##### WRITE FILES #####
```

```
write.table(x=plot.data.traits,  
file=paste0(path.to.files, "02_Data/plot.data.traits.txt"),  
sep="\t", fileEncoding="UTF-8")
```

```
write.table(x=tree.data.traits,  
file=paste0(path.to.files, "02_Data/tree.data.traits.txt"),  
sep="\t", fileEncoding="UTF-8")
```

```
write.table(x=traits.summary,  
file=paste0(path.to.files, "02_Data/traits.summary.txt"),  
sep="\t", fileEncoding="UTF-8")
```

```
##### CREATING MAPS OF STEMS #####
```

```
tree.sizes <- sqrt(as.numeric(tree.data.traits$Diameter t2)) # had to add as.numeric/ no longer  
necessary b/c read.csv correction
```

```
size.to.plot.tree <- (tree.sizes-min(tree.sizes, na.rm=TRUE)) / (max(tree.sizes, na.rm=TRUE)-  
min(tree.sizes, na.rm=TRUE))
```

```
size.to.plot.tree[is.na(size.to.plot.tree)] <- 0.5
```

```
size.to.plot.tree <- (size.to.plot.tree)+0.75
```

```
par(mfrow=c(1,2))
```

```
plot(tree.sizes~size.to.plot.tree)
```

```
hist(size.to.plot.tree)
```

```
tree.data.traits <- data.frame(tree.data.traits, size.to.plot.tree)
```

```
tree.data.traits <- tree.data.traits[order(tree.data.traits$size.to.plot.tree, decreasing=TRUE), ]
```

```
plot.list<-unique(as.character(tree.data.traits$PlotName))
```

```
for(i in 1:length(plot.list))
```

```
{
```

```
  #i<-50 #To test a single iteration of the for loop
```

```
  PlotName.i<-plot.list[i]
```

```
  print(PlotName.i) #Why print the plotName here?
```

```
  tree.data.traits.i <- tree.data.traits[which(tree.data.traits$PlotName==PlotName.i),]
```

```
  tree.coords.i <- tree.data.traits.i[,c("Y", "X")]
```

```

which.plot.i<-which(plot.data$PlotName == PlotName.i)
plot.info<-plot.data[which.plot.i,]

spp.list<-unique(as.character(tree.data.traits.i$BinomialSpeciesName))
colors.spp<-rainbow(length(spp.list))
colors.spp<-colors.spp[sample(1:length(spp.list))]

png(filename = paste(path.to.files, "04 Results/Figures/PlotMaps/", i, " ", PlotName.i, ".png",
sep=""), width=480*7, height=480*7.9, res=800)

try(
{
  y.range <- range(c(tree.coords.i$Y), na.rm=TRUE)
  y.labels<-seq(y.range[1], y.range[2], length.out=6)
  x.range <- range(c(tree.coords.i$X), na.rm=TRUE)
  x.labels<-seq(x.range[1], x.range[2], length.out=6)

  plot.title <- paste(plot.info$PlotName, ": ", length(spp.list), " species", sep="")

  plot(tree.coords.i$Y~tree.coords.i$X, pch=21, bg="forestgreen", asp=1,
  ylim=range(tree.coords.i$Y, na.rm=TRUE), xlim=range(tree.coords.i$X, na.rm=TRUE),
  ylab="Y", xlab="X", main=plot.title, type="n", cex.axis=1, cex.lab=1, cex.main=1, las=1,
  axes=FALSE)

  axis(side=1, at=x.labels, labels = round(x.labels, 0), lwd = 1, lwd.ticks = 1, col = NULL,
  col.ticks = NULL, hadj = NA, padj = NA, cex.axis=1)

  axis(side=2, at=y.labels, labels = round(y.labels, 0), lwd = 1, lwd.ticks = 1, col = NULL,
  col.ticks = NULL, hadj = NA, padj = NA, cex.axis=1, las=1)

  for(j in 1:length(spp.list))

```

```
points(tree.coords.i$Y[tree.data.traits.i$BinomialSpeciesName==spp.list[j]]~tree.coords.i$X[tree  
.data.traits.i$BinomialSpeciesName==spp.list[j]], pch=21, bg=colors.spp[j], col="grey30",  
cex=tree.data.traits.i$size.to.plot.tree, lwd=1)
```

```
}, silent=TRUE)
```

```
dev.off() #Makes sure it closes files after creation when starting a new file
```

```
}
```

```
```
```

Neighborhood analyses

```
```{r}
```

```
rm(list=objects())
```

```
set.seed(1981)
```

```
library(FD)
```

```
library(vegan)
```

```
### OPEN FUNCTIONS NEEDED FOR ANALYSES
```

```
source(paste0(path.to.files,"03 Rcode/00 FunctionsForAnalyses INDEVELOPMENT 2021Oct  
19.R"))
```

OPEN BOTH LISTS, INDIVIDUALS AND COORDS + TRAITS BY INDS.

```
plot.data <- read.csv(paste0(path.to.files,  
"02_Data/plot.data.traits.txt"),  
header = TRUE, sep = "\t", fileEncoding="UTF-8")  
dim(plot.data)
```

```
tree.data <- read.csv(paste0(path.to.files,  
"02_Data/tree.data.traits.txt"),  
header = TRUE, sep = "\t", fileEncoding="UTF-8")  
dim(tree.data)
```

```
traits.summary <- read.csv(paste0(path.to.files,  
"02_Data/traits.summary.txt"),  
header = TRUE, sep = "\t", fileEncoding="UTF-8")  
dim(traits.summary)
```

```
plot.list <- plot.data$PlotName
```

```
SAD.Options <- c("fixed")
```

```
Neig.Limits <- c(20, 15, 10, 5)
```

```
#Neig.Limits <- c(15)
```

```
rand.n <- 50
```

```
save.res <- TRUE
```

```
#trait.option <- "SLA"
```

```
trait.option <- "all"
```

```
Animation <- FALSE
```

```
nl <- 1
```

```
SAD.o <- 1
```

```
p <- 1
```

```
### IMPUTATION OF MISSING VALUES IN THE FULL TRAIT MATRIX ###
```

```
library(missForest)
```

```
imputation <- missForest(traits.summary[,-c(1,2)])
```

```
imputation <- imputation$ximp
```

```
traits.summary <- cbind(traits.summary[,c(1,2)], imputation)
```

MODIFY THE TRAIT MATRIX DEPENDING ON THE TRAITS TO USE

If all traits are to be used, then the trait matrix is first

run through a PCA

if(trait.option=="all")

{

traits.pcs <- princomp(traits.summary[,-c(1,2)], cor=TRUE)\$scores

traits.summary <- cbind(traits.summary[,c(1,2)], traits.pcs)

}

CALCULATES PLOT-LEVEL FUNCTIONAL DIVERSITY FOR ALL PLOTS
SIMULTANEOUSLY ###

compo.by.plot <- table(tree.data\$PlotName, tree.data\$BinomialSpeciesName)

dim(compo.by.plot)

plot.richness <- rowSums(compo.by.plot>0)

plot.shannon <- diversity(x=compo.by.plot, index="shannon", MARGIN = 1, base = exp(1))

plot.invsimpson <- diversity(x=compo.by.plot, index = "invsimpson", MARGIN = 1, base =
exp(1))

plot.TDiv.table <- cbind(plot.richness=plot.richness, plot.shannon=plot.shannon,

plot.invsimpson=plot.invsimpson)

```

# Create a species composition matrix and trait table using species-by-site
# combinations.

plot.compo.temp <- table(tree.data$PlotName, paste(tree.data$PlotName,
tree.data$BinomialSpeciesName, sep=" "))

traits.summary.temp <- traits.summary
rownames(traits.summary.temp) <- paste(traits.summary$PlotName,
traits.summary$BinomialSpeciesName, sep=" ")

spp.with.traits.temp <- intersect(rownames(traits.summary.temp), colnames(plot.compo.temp))

traits.summary.temp <- traits.summary.temp[spp.with.traits.temp,-c(1,2)]
plot.compo.temp <- as.data.frame.matrix(plot.compo.temp[,spp.with.traits.temp])

plot.dbFD <- dbFD(x=traits.summary.temp, a=plot.compo.temp,
w.abun=TRUE, calc.FGR=FALSE, calc.CWM=FALSE)

MPDs <- mpd(samp=plot.compo.temp,
dis=as.matrix(dist(traits.summary.temp)), abundance.weighted=TRUE)

names(MPDs) <- rownames(plot.compo.temp)

plot.FDiv.table <- cbind(nbsp=plot.dbFD$nbsp, FRic=plot.dbFD$FRic, FDiv=plot.dbFD$FDiv,

```

FDis=plot.dbFD\$FDis, RaoQ=plot.dbFD\$RaoQ, MPD=MPDs)

Joins both diversity tables

identical(rownames(plot.TDiv.table), rownames(plot.FDiv.table))

plot.diversity.table <- cbind(plot.TDiv.table, plot.FDiv.table)

pairs(plot.diversity.table)

RUNS THE NEIGHBORHOOD CALCULATIONS FOR MULTIPLE OPTIONS IN LOOPS

###

for(nl in 1:length(Neig.Limits))

{

 neig.limit <- Neig.Limits[nl]

 for(SAD.o in 1:length(SAD.Options))

 {

 sad.option <- SAD.Options[SAD.o]

 plot.res.list <- sapply(rep(NA, length(plot.list)), list)

 names(plot.res.list) <- plot.list

```

for(p in 1:length(plot.list))
{
plot.name <- plot.list[p]
print(paste(plot.name, "; Plot", p, "of", length(plot.list)))

## individuals and their coordinates + traits by individuals
plot.p <- tree.data[tree.data$PlotName == plot.name,]
plot.p <- plot.p[order(plot.p$BinomialSpeciesName), ]
dim(plot.p)

# Functional distances only if there is functional data
traits.p <- traits.summary[traits.summary$PlotName == as.character(plot.name),]
rownames(traits.p) <- traits.p$BinomialSpeciesName
traits.p <- traits.p[order(traits.p$BinomialSpeciesName), ]
traits.p <- traits.p[,-c(1,2)]

spp.R.abund <- as.numeric(table(plot.p$BinomialSpeciesName))

which.0.abund <- which(spp.R.abund==0)
if(length(which.0.abund)>0)
spp.R.abund <- spp.R.abund[-which.0.abund]

```

```

spp.with.traits <-
intersect(plot.p$BinomialSpeciesName, rownames(traits.p))
plot.rich.with.traits <- length(spp.with.traits)

plot.p$BinomialSpeciesName <- as.factor(plot.p$BinomialSpeciesName)

if(trait.option=="all")
{
  trait.dists <- dist(traits.p, method = "euclidean")
}
else
{
  traits.p <- traits.p[trait.option]
  trait.dists <- dist(traits.p, method = "euclidean")
}

# give to the function one object with the tree and coordinates info (the original version),
# but add another argument that receives an object with the info on traits per individual.

plot.res <-
neig.sim.plot(plot.p=plot.p, trait.table=traits.p, trait.dists=trait.dists,

```

```

_____ neig.limit=neig.limit, rand.n=rand.n,
_____ SAD=sad.option, show.animation=Animation)

_____ plot.res.means.1 <- colMeans(plot.res[,c(1,2)], na.rm=TRUE)
_____ plot.res.means.2 <- apply(plot.res[, -c(1,2)], 2,
_____ weighted.mean, w=plot.res$usable.neig.n, na.rm=TRUE)

_____ plot.res.means <- c(plot.res.means.1, plot.res.means.2)

_____ if(p==1)
_____ {
_____ plot.res.table <- as.data.frame(
_____ matrix(NA, ncol=(1+length(plot.res.means)), nrow=length(plot.list)))

_____ rownames(plot.res.table) <- plot.list
_____ colnames(plot.res.table) <- c("plot.rich.with.traits", names(plot.res.means))
_____ }

_____ plot.res.table[p,] <- c(plot.rich.with.traits, plot.res.means)

_____ plot.res.list[[p]] <- plot.res
_____ }

_____ plot.res.table <- plot.res.table[order(rownames(plot.res.table)),]

```

```

plot.diversity.table <- plot.diversity.table[order(rownames(plot.diversity.table)),]

if(!identical(rownames(plot.res.table), rownames(plot.diversity.table)))
stop("Names of these tables are not the same")

plot.res.table <- data.frame(plot.diversity.table, plot.res.table)

if(save.res)
{
save(plot.res.list, file=paste(path.to.files, "04 Results/plot.res.list ", neig.limit, "m SAD.",
sad.option, sep=""))
write.table(plot.res.table, file=paste(path.to.files, "04 Results/plot.res.table ", neig.limit,
"m SAD.", sad.option, ".txt", sep=""), sep="\t")
}
}
}

...
==

```

R code for Chapter 3

title: "Chapter 3 R code"

output: word document

```
```{r}
```

```
Create phylogenetic tree
```

```
library(picante)
```

```
library(V.PhyloMaker)
```

```
library(phyloomatic)
```

```
library(phangorn)
```

```
library(phylotools)
```

```
library(ape)
```

```
library(dplyr)
```

```
library(phytools)
```

```
library(ggtree)
```

```
#Load tree plot data
```

```
treeplotdata = read.table(paste0(path.to.files,
_____ "02_Data/tree.plot.data.txt"))
```

```
unique(treeplotdata$PlotName)
```

```
Fix Kanupa
```

```
treeplotdata$PlotName[which(treeplotdata$PlotName=="PP_Kanupa_44")]<-
rep("PP_Kanupa_44",length(which(treeplotdata$PlotName=="PP_Kanupa_44")))
```

```
unique(treeplotdata$PlotName)
```

```
Make tree from treeplotdata
```

```
v.phylo.df <- as.data.frame(treeplotdata[,2])
```

```
v.phylo.df$Genus <- sapply(strsplit(treeplotdata$BinomialSpeciesName, " "),
function(x) x[1])
```

```
v.phylo.df$Family <- treeplotdata[,4]
```

```
colnames(v.phylo.df) = c("Species", "Genus", "Family")
```

```
Make tree
```

```
phy <- phylo.maker(v.phylo.df)
```

```
tree.3 <- phy$scenario.3
```

```
class(tree.3) #phylo
```

```
length(tree.3$tip.label) #1124
```

```
Root tree
```

```
tree.3.rooted <- multi2di(tree.3)
```

```
...
```

```
``{r}
```

```
Calculate abundance-weighted median & mean CSCS
```

```
Calculate mean and median CSCS for all plots
```

```
cscs.plot.list <- list(Chaqui32, Fuerte27, Kanupa44, Lomaka40, Lomasa39,
```

```
_____ Pintat5, Resina12, Sumpul34, Tapuri45, Tintay24,
```

```
_____ Tintay25, Titiri42, Tocoaq28, Tocoaq29, Tocoaq30, Yarimi9)
```

```
cscs.plot.mean.med.df <- data.frame(Plot=unique(cscs$Plot), plot.mean.cscs=rep(0, 16),
```

```
_____ plot.med.cscs=rep(0, 16))
```

```
Check to see that plots are in original order (matching cscs.plot.list)
```

```
cscs.plot.mean.med.df$original.plot <-
```

```
Loop species pairs each plot
```

```
for(i in 1:length(cscs.plot.list))
```

```
{
```

```
 # calculate the mean of dissimilarity matrix of
```

```
 i.mean <-
```

```
 _____ sum(cscs.plot.list[[i]])/(nrow(cscs.plot.list[[i]])*ncol(cscs.plot.list[[i]])-
 _____ nrow(cscs.plot.list[[i]]))
```

```
 _____ cscs.plot.mean.med.df[i, 2] <- i.mean
```

```

calculate the median of dissimilarity matrix of
i.med <-
median(as.vector(as.matrix(cscs.plot.list[[i]]))
which(as.vector(as.matrix(cscs.plot.list[[i]])!=0))

cscs.plot.mean.med.df[i, 3] <- i.med

}

add columns to dataframe
cscs.plot.mean.med.df$abundWtCSCS <- plotabmeanCSCS
cscs.plot.mean.med.df$elevation <- plot.elevations.2
cscs.plot.mean.med.df$invSimpson <- plot.invSimp

Save dataframe for later
write.csv()
```
  


```

```{r}
# Phylogenetic (Mantel) and Rarefaction analyses by plot
library(vegan)
library(ggplot2)
library(dplyr)
library(sfsmisc)

```


```

```
library(tidyr)
```

```
library(picante)
```

```
library(matrixStats)
```

```
library(V.PhyloMaker)
```

```
#path.to.files <- "~/Documents/"
```

```
setwd("~/Documents/")
```

```
cscs.plot.mean.med.df = read.csv("cscs.plot.mean.med.df_20220930.csv", header = TRUE, sep =
".")
```

```
cscs.plot.mean.med.df <- cscs.plot.mean.med.df[, -1]
```

```
colnames(cscs.plot.mean.med.df)
```

```
Opens phylogenetic tree
```

```
load("mad_tree_rooted1")
```

```
phylo.dist <- as.matrix(cophenetic(tree.3.rooted))
```

```
rownames(phylo.dist) <- gsub(" ", "", rownames(phylo.dist))
```

```
colnames(phylo.dist) <- gsub(" ", "", colnames(phylo.dist))
```

```
All Compounds
```

```

load(paste0(path.to.files,

"04_Results/Individual_CompoundClasses/Madidi_plotabundmeanCSCS_metabtot_20220608.R
data"))

load("Madidi_plotabundmeanCSCS_metabtot_20220608.Rdata")

cscs.all <- as.matrix(mad_cscs_all)

Defense Compounds

load(paste0(path.to.files,

"04_Results/Individual_CompoundClasses/MadAll_metabsim_npclass_defense_20220614_2022
0806.Rdata"))

load("MadAll_metabsim_npclass_defense_20220614_20220806.Rdata")

cscs.def <- as.matrix(1-cscs)

Primary Compounds

load(paste0(path.to.files,

"04_Results/Individual_CompoundClasses/CSCSsppBCI-prim-20220822.Rdata"))

load("CSCSsppBCI-prim-20220822.Rdata")

cscs.prim <- as.matrix(1-cscs)

dim(cscs.all)

dim(cscs.def)

```

dim(cscs.prim)

## \*\*\*NOTE\*\*\* ##

# Species in row and colum 384 is present for all and defense compounds,

# but not for the primary compounds matrix. To standardize matrices, we

# remove if from all matrices

#which(rownames(mad\_cscs\_all) == "PP Sumpul 34 Pleurothyrium trianae") #  
mad\_cscs\_all[384, 384]

#Not included in CSCSsppBCI-prim-20220822.Rdata

cscs.all <- cscs.all[-384, -384]

cscs.def <- cscs.def[-384, -384]

identical(rownames(cscs.def), rownames(cscs.prim))

# Change rownames, colnames

rownames(cscs.def) <- rownames(cscs.all)

colnames(cscs.def) <- colnames(cscs.all)

identical(rownames(cscs.all), rownames(cscs.def))

rownames(cscs.prim) <- rownames(cscs.all)

colnames(cscs.prim) <- colnames(cscs.all)

identical(rownames(cscs.all), rownames(cscs.prim))

```
NOTE
```

```
One species in the cscs matrix has no name "NA NA"
```

```
which(rownames(cscs.all) == "NA NA")
```

```
cscs.all <-
```

```
cscs.all[-which(rownames(cscs.all) == "NA NA"),
```

```
-which(colnames(cscs.all) == "NA NA")]
```

```
cscs.def <-
```

```
cscs.def[-which(rownames(cscs.def) == "NA NA"),
```

```
-which(colnames(cscs.def) == "NA NA")]
```

```
cscs.prim <-
```

```
cscs.prim[-which(rownames(cscs.prim) == "NA NA"),
```

```
-which(colnames(cscs.prim) == "NA NA")]
```

```
Add columns for species and plot
```

```
species.var <- sapply(strsplit(colnames(cscs.all), "[_]"), function(x) x[4])
```

```
length(unique(species.var))
```

```
setdiff(species.var, rownames(phylo.dist)) ## Some species are missing from the phylogeny
```

```
setdiff(rownames(phylo.dist), species.var)
```

```
plot.var.1 <- sapply(strsplit(rownames(cscs.all), "[_]"), function(x) x[[1]])
```

```
plot.var.2 <- sapply(strsplit(rownames(cscs.all), "[_]"), function(x) x[[2]])
```

```
plot.var.3 <- sapply(strsplit(rownames(cscs.all), "[_]"), function(x) x[3])
```

```
plot.var <- paste0(plot.var.1, " ", plot.var.2, " ", plot.var.3)
```

```
plot.list <- unique(plot.var)
```

```
Loop through each plot
```

```
for(i in 1:length(plot.list))
```

```
{
```

```
 plot.name.i <- plot.list[i]
```

```
 print(plot.name.i)
```

```
 which.plot.i <- which(plot.var==plot.name.i)
```

```
 species.var.i <- species.var[which.plot.i]
```

```
 cscs.all.i <- cscs.all[which.plot.i, which.plot.i]
```

```
 cscs.def.i <- cscs.def[which.plot.i, which.plot.i]
```

```
 cscs.prim.i <- cscs.prim[which.plot.i, which.plot.i]
```

```

rownames(cscs.all.i) <- colnames(cscs.all.i) <- species.var.i
rownames(cscs.def.i) <- colnames(cscs.def.i) <- species.var.i
rownames(cscs.prim.i) <- colnames(cscs.prim.i) <- species.var.i

Finds the phylogenetic distances for plot i
species.var.i <- intersect(species.var.i, rownames(phylo.dist))

phylo.dist.i <- phylo.dist[species.var.i, species.var.i]

phylo.cor.all <- mantel(xdis=as.dist(phylo.dist.i),
 ydis=as.dist(cscs.all.i[species.var.i, species.var.i]))$statistic

phylo.cor.def <- mantel(xdis=as.dist(phylo.dist.i),
 ydis=as.dist(cscs.def.i[species.var.i, species.var.i]))$statistic

phylo.cor.prim <- mantel(xdis=as.dist(phylo.dist.i),
 ydis=as.dist(cscs.prim.i[species.var.i, species.var.i]))$statistic

cscs.plot.mean.med.df$phylo.cor.all[
 cscs.plot.mean.med.df$Plot == plot.name.i] <- phylo.cor.all

cscs.plot.mean.med.df$phylo.cor.def[

```

```

cscs.plot.mean.med.df$Plot == plot.name.i] <- phylo.cor.def

cscs.plot.mean.med.df$phylo.cor.prim[
cscs.plot.mean.med.df$Plot == plot.name.i] <- phylo.cor.prim

calculate the mean of dissimilarity matrix of
mean.all.i <- mean(as.dist(cscs.all.i))
mean.def.i <- mean(as.dist(cscs.def.i))
mean.prim.i <- mean(as.dist(cscs.prim.i))

calculate the median of dissimilarity matrix of
median.all.i <- median(as.dist(cscs.all.i))
median.def.i <- median(as.dist(cscs.def.i))
median.prim.i <- median(as.dist(cscs.prim.i))

cscs.plot.mean.med.df$mean.cscs.all[
cscs.plot.mean.med.df$Plot == plot.name.i] <- mean.all.i

cscs.plot.mean.med.df$mean.cscs.def[
cscs.plot.mean.med.df$Plot == plot.name.i] <- mean.def.i

cscs.plot.mean.med.df$mean.cscs.prim[
cscs.plot.mean.med.df$Plot == plot.name.i] <- mean.prim.i

```

```

cscs.plot.mean.med.df$median.cscs.all[
cscs.plot.mean.med.df$Plot == plot.name.i] <- median.all.i

cscs.plot.mean.med.df$median.cscs.def[
cscs.plot.mean.med.df$Plot == plot.name.i] <- median.def.i

cscs.plot.mean.med.df$median.cscs.prim[
cscs.plot.mean.med.df$Plot == plot.name.i] <- median.prim.i

Rarified speices analyses
n.rand <- 1000
target.S <- 12

#Mean
mean.all.i.j <- rep(NA, times=n.rand)
mean.def.i.j <- mean.prim.i.j <- mean.all.i.j

for(j in 1:n.rand)
{
 if(nrow(cscs.all.i) <= target.S)
 {
 mean.all.i.j <- mean.all.i

```

```

mean.def.i.j <- mean.def.i
mean.prim.i.j <- mean.prim.i

warning("Not enough species")
break()
}

sample.j <- sample(1:nrow(cscs.all.i), size=target.S)

cscs.all.i.j <- cscs.all.i[sample.j, sample.j]
cscs.def.i.j <- cscs.def.i[sample.j, sample.j]
cscs.prim.i.j <- cscs.prim.i[sample.j, sample.j]

mean.all.i.j[j] <- mean(as.dist(cscs.all.i.j))
mean.def.i.j[j] <- mean(as.dist(cscs.def.i.j))
mean.prim.i.j[j] <- mean(as.dist(cscs.prim.i.j))

}

cscs.plot.mean.med.df$raref.mean.cscs.all[
cscs.plot.mean.med.df$Plot == plot.name.i] <- mean(mean.all.i.j)

cscs.plot.mean.med.df$raref.mean.cscs.def[
cscs.plot.mean.med.df$Plot == plot.name.i] <- mean(mean.def.i.j)

```

```

cscs.plot.mean.med.df$raref.mean.cscs.prim[
cscs.plot.mean.med.df$Plot == plot.name.i] <- mean(mean.prim.i.j)

Median
median.all.i.j <- rep(NA, times=n.rand)
median.def.i.j <- median.prim.i.j <- median.all.i.j

for(j in 1:n.rand)
{
 if(nrow(cscs.all.i) <= target.S)
 {
 median.all.i.j <- median.all.i
 median.def.i.j <- median.def.i
 median.prim.i.j <- median.prim.i

 warning("Not enough species")
 break()
 }

 sample.j <- sample(1:nrow(cscs.all.i), size=target.S)

 cscs.all.i.j <- cscs.all.i[sample.j, sample.j]
 cscs.def.i.j <- cscs.def.i[sample.j, sample.j]
 cscs.prim.i.j <- cscs.prim.i[sample.j, sample.j]

```

```

median.all.i.j[j] <- median(as.dist(cscs.all.i.j))
median.def.i.j[j] <- median(as.dist(cscs.def.i.j))
median.prim.i.j[j] <- median(as.dist(cscs.prim.i.j))

}

cscs.plot.mean.med.df$raref.median.cscs.all[
cscs.plot.mean.med.df$Plot == plot.name.i] <- median(median.all.i.j)

cscs.plot.mean.med.df$raref.median.cscs.def[
cscs.plot.mean.med.df$Plot == plot.name.i] <- median(median.def.i.j)

cscs.plot.mean.med.df$raref.median.cscs.prim[
cscs.plot.mean.med.df$Plot == plot.name.i] <- median(median.prim.i.j)

}

Save dataframe for later
write.csv()
```
  


```

```{r}
# Run KMULT phylogenetic analysis - more appropriate than Mantel test

```


```

#### 2023-02-21: Use Kmult from Adams 2014 Syst Biol (in R package 'phylocurve') to calculate phylogenetic signal

#### regress the Kmult statistic against diversity, elevation, and climatic PC1 and PC2

setwd("~/Documents/Madidi\_Project/Chapter2\_2023\_02\_21")

library(vegan)

library(ggplot2)

library(dplyr)

library(sfsmisc)

library(tidyr)

library(picante)

library(matrixStats)

library(V.PhyloMaker)

library(phylocurve)

plotdata =

read.csv("~/Documents/Madidi\_Project/Chapter2\_2023\_02\_21/cscs.plot.mean.med.df\_20230221.csv", header = T)[-1]

dim(plotdata)

# [1] 16 26

```
head(plotdata)
```

```
load phylogenetic tree generated usig v.phylomaker
```

```
load("mad tree rooted1")
```

```
phylo.dist <- as.matrix(cophenetic(tree.3.rooted))
```

```
rownames(phylo.dist) <- gsub(" ", "", rownames(phylo.dist))
```

```
colnames(phylo.dist) <- gsub(" ", "", colnames(phylo.dist))
```

```
setwd("~/Documents/Madidi Project")
```

```
load("MadidiAllMetab_20220205.RData")
```

```
load("Madidid_30k_all_Sirius_NPClassifier_20220607.RData")
```

```
heat.def = heat.real[which(siri.itol$custom %in% c("Alkaloids", "Amino acids and Peptides",
"Polyketides", "Shikimates and Phenylpropanoids", "Terpenoids", "Multiple")),]
```

```
dim(heat.def)
```

```
heat.prim = heat.real[which(siri.itol$custom %in% c("Amino acids and Peptides",
"Carbohydrates", "Fatty acids")),]
```

```
dim(heat.prim)
```

```
[1] 1190 906
```

```
names(heat.real) = gsub(".Peak.area", "", names(heat.real))
```

```
names(heat.def) = gsub(".Peak.area", "", names(heat.def))
```

```
names(heat.prim) = gsub(".Peak.area", "", names(heat.prim))
```

```
length(tree.3.rooted$tip.label)
```

```
[1] 1124
```

```
metabig = read.table("~/Documents/Madidi Project/Madidi metadata MSV000090549.txt",
header = T, sep = "\t")
```

```
length(which(tree.3.rooted$tip.label %in% metabig$ATTRIBUTE_Genus_species))
```

```
[1] 359
```

```
tree.3.rooted$tip.label[which(tree.3.rooted$tip.label %in%
metabig$ATTRIBUTE_Genus_species)]
```

```
tree.3.rooted$tip.label[1:100]
```

```
tree.3.rooted$tip.label[grep("Psychotria", tree.3.rooted$tip.label)]
```

```
[1] "Psychotria_trivialis" "Psychotria_AF14766" "Psychotria_LC6490"
"Psychotria_CMG2521" "Psychotria_CMG2789"
```

```
[6] "Psychotria_AEC305" "Psychotria_trichotoma" "Psychotria_AF16987"
"Psychotria_carthagenensis"
```

```
metabig$ATTRIBUTE_Genus_species[grep("Psychotria",
metabig$ATTRIBUTE_Genus_species)]
```

[1] "Psychotria AEC-305" "Psychotria AEC-305" "Psychotria AEC-305" "Psychotria AEC-305" "Psychotria AEC-305"

tree.3.rooted\$tip.label[grep("Ocotea", tree.3.rooted\$tip.label)]

[1] "Ocotea NPZ4964" "Ocotea AM3556" "Ocotea olivacea" "Ocotea LC3806"  
"Ocotea caesifolia" "Ocotea LC6779"

[7] "Ocotea mandonii" "Ocotea RSS25" "Ocotea ALM253" "Ocotea ALM63"  
"Ocotea ETP63" "Ocotea CMG2534"

[13] "Ocotea LC5997" "Ocotea LC6203" "Ocotea subrutilans" "Ocotea AEC335"  
"Ocotea AEC147" "Ocotea oblonga"

[19] "Ocotea PCL155" "Ocotea NCH209" "Ocotea cernua" "Ocotea AM3387"  
"Ocotea FZR18951" "Ocotea PCL181"

[25] "Ocotea floribunda" "Ocotea AF15396" "Ocotea AF12244" "Ocotea LC4365A"  
"Ocotea micrantha" "Ocotea albida"

[31] "Ocotea LC4144A" "Ocotea weberbaueri" "Ocotea comata" "Ocotea ECQ72"  
"Ocotea obovata" "Ocotea bofo"

[37] "Ocotea longifolia" "Ocotea puberula" "Ocotea aciphylla"

metabig\$ATTRIBUTE\_Genus\_species[grep("Ocotea", metabig\$ATTRIBUTE\_Genus\_species)]

levels(as.factor(meta\$ATTRIBUTE\_Plot))

[1] "PP Chaqui 32" "PP Fuerte 27" "PP Ka\x96upa 44" "PP Lomaka 40"  
"PP Lomasa 39" "PP Pintat 5" "PP Resina 12" "PP Sumpul 34"

[9] "PP Tapuri 45" "PP Tintay 24" "PP Tintay 25" "PP Titiri 42" "PP Toqoq 28"  
"PP Toqoq 29" "PP Toqoq 30" "PP Yamiri 9"

meta[which(meta\$ATTRIBUTE Plot == "PP Tintay 24"),]

meta[which(meta\$ATTRIBUTE Plot == "PP Yamiri 9"),]

meta[which(meta\$ATTRIBUTE Plot == "PP Pintat 5"),]

meta[which(meta\$ATTRIBUTE Plot == "PP Ka\x96upa 44"),]

"PP Yamiri 9"

?phylocurve

metabig\$ATTRIBUTE Genus species = gsub("-", "", metabig\$ATTRIBUTE Genus species)

metapool = meta

metapool\$Genus sp = NA

for(i in 1:nrow(metapool)){

\_\_\_\_\_metapool\$Genus sp[i] =

metabig\$ATTRIBUTE Genus species[which(metabig\$ATTRIBUTE SpeciesCode ==  
meta\$ATTRIBUTE SppCode[i])]

}

head(metapool)

metapool\$ATTRIBUTE\_Plot = gsub("\x96", "n", metapool\$ATTRIBUTE\_Plot)

metapool\$ATTRIBUTE\_Plot = gsub("Toqoaq", "Tocoaq", metapool\$ATTRIBUTE\_Plot)

metapool\$ATTRIBUTE\_Plot = gsub("Yamiri", "Yarimi", metapool\$ATTRIBUTE\_Plot)

levels(as.factor(metapool\$ATTRIBUTE\_Plot))

[1] "PP Chaqui 32" "PP Fuerte 27" "PP Kanupa 44" "PP Lomaka 40" "PP Lomasa 39"  
"PP Pintat 5" "PP Resina 12" "PP Sumpul 34" "PP Tapuri 45"

[10] "PP Tintay 24" "PP Tintay 25" "PP Titiri 42" "PP Toqoaq 28" "PP Toqoaq 29"  
"PP Toqoaq 30" "PP Yamiri 9"

plotdata\$Kmult.tot.K = NA

plotdata\$Kmult.tot.p = NA

plotdata\$Kmult.def.K = NA

plotdata\$Kmult.def.p = NA

plotdata\$Kmult.prim.K = NA

plotdata\$Kmult.prim.p = NA

for(i in 14:length(levels(as.factor(metapool\$ATTRIBUTE\_Plot)))){

    ploti = as.character(levels(as.factor(metapool\$ATTRIBUTE\_Plot))[i])

    cat("Analyzing Kmult for plot", ploti, sep = " ", "\n")

    metapooli = metapool[which(metapool\$ATTRIBUTE\_Plot == ploti),]

    # filesploti = metapool\$filename[which(metapool\$ATTRIBUTE\_Plot == ploti)]

    # speciesploti = metapool\$Genus\_sp[which(metapool\$ATTRIBUTE\_Plot == ploti)]

    dummy = rep(1,nrow(metapooli))

    names(dummy) = metapooli\$Genus\_sp

```

phyloploti = prune.missing(x = dummy, phylo = tree.3.rooted)$tree
phyloploti = prune.missing(x = dummy, phylo = testphylo)$tree
heat.tot.ploti = heat.real[,which(names(heat.real) %in% metapooli$filename)]
heat.tot.ploti = heat.tot.ploti[which(rowSums(heat.tot.ploti) > 0),]
heat.def.ploti = heat.def[,which(names(heat.def) %in% metapooli$filename)]
heat.def.ploti = heat.def.ploti[which(rowSums(heat.def.ploti) > 0),]
heat.prim.ploti = heat.prim[,which(names(heat.prim) %in% metapooli$filename)]
heat.prim.ploti = heat.prim.ploti[which(rowSums(heat.prim.ploti) > 0),]
for(j in 1:ncol(heat.tot.ploti)){
 filenamej = names(heat.tot.ploti)[j]
 names(heat.tot.ploti)[j] = names(heat.def.ploti)[j] = names(heat.prim.ploti)[j] =
metapooli$Genus_sp[which(metapooli$filename == filenamej)]
}
evomodel.tot = evo.model(tree = phyloploti, Y = t(heat.tot.ploti), method = "Pairwise
ML")
kmultresults.tot = K.mult(model = evomodel.tot, nsim = 1000, plot = F)
plotdata$Kmult.tot.K[which(plotdata$Plot == ploti)] = kmultresults.tot$K
plotdata$Kmult.tot.p[which(plotdata$Plot == ploti)] = kmultresults.tot$Pval
evomodel.def = evo.model(tree = phyloploti, Y = t(heat.def.ploti), method = "Pairwise
ML")
kmultresults.def = K.mult(model = evomodel.def, nsim = 1000, plot = F)
plotdata$Kmult.def.K[which(plotdata$Plot == ploti)] = kmultresults.def$K
plotdata$Kmult.def.p[which(plotdata$Plot == ploti)] = kmultresults.def$Pval
evomodel.prim = evo.model(tree = phyloploti, Y = t(heat.prim.ploti), method = "Pairwise
ML")
kmultresults.prim = K.mult(model = evomodel.prim, nsim = 1000, plot = F)

```

```
plotdata$Kmult.prim.K[which(plotdata$Plot == ploti)] = kmultresults.prim$K
plotdata$Kmult.prim.p[which(plotdata$Plot == ploti)] = kmultresults.prim$Pval
}
```

```
write.csv(plotdata, file = "Madidi_plotdata_Kmult_20230222.csv", quote = F)
```

```
mod.div = lm(plotdata$Kmult.tot.K ~ plotdata$invSimpson)
```

```
summary(mod.div)
```

```
mod.elev = lm(plotdata$Kmult.tot.K ~ plotdata$elevation)
```

```
summary(mod.elev)
```

```
mod.elev.nodry = lm(plotdata$Kmult.tot.K[which(plotdata$ForestType == "moist montane")] ~
plotdata$elevation[which(plotdata$ForestType == "moist montane")])
```

```
summary(mod.elev.nodry)
```

```
mod.pc1 = lm(plotdata$Kmult.tot.K ~ plotdata$PCA1)
```

```
summary(mod.pc1)
```

```
mod.pc2 = lm(plotdata$Kmult.tot.K ~ plotdata$PCA2)
```

```
summary(mod.pc2)
```

```
K.mult(model = , nsim = 1000, plot = T)
```

```
test = heat.tot.ploti/rowMeans(heat.tot.ploti)
```

```
evomodel.tot = evo.model(tree = phyloploti, Y = t(test), method = "Pairwise ML")
```

```
rand.data <- sim.traits()
```

```
rand.data
```

```
$trait_data
```

```
_____ # V1 V2 V3 V4
t2 1.3969525 0.47883417 4.2579360 4.5814450
t6 2.9829437 -0.89376273 -2.1986710 -1.3402672
t7 2.8468325 -0.42314268 0.7725386 0.8158939
t3 0.2883345 -0.27384562 1.3952555 4.3724692
```

# t4 -1.2785160 -0.71825975 0.1592125 6.0280045  
# t8 -1.1388404 -0.65469252 2.3485904 7.8514698  
# t9 -0.3107582 0.22277635 2.4436410 5.3318763  
# t12 1.7461535 0.43612073 3.7927976 3.8182190  
# t14 1.4809865 1.26835503 5.0226705 3.8463821  
# t15 1.4043228 1.28734259 4.8853338 3.7569117  
# t13 1.4908855 1.10941296 4.4181962 3.5220316  
# t5 3.2182798 0.92199269 5.3783353 1.8242808  
# t10 0.4621692 -0.54571569 3.8448105 6.6462391  
# t11 1.3677619 -0.08560965 2.7896276 3.8337457  
# t1 2.5311757 1.94589249 5.5230757 1.5326607

# \$tree

# Phylogenetic tree with 15 tips and 14 internal nodes.

# Tip labels:

# t2, t6, t7, t3, t4, t8, ...

# Rooted; includes branch lengths.

# \$sim tree

# Phylogenetic tree with 15 tips and 14 internal nodes.

# Tip labels:

# t2, t6, t7, t3, t4, t8, ...

# Rooted; includes branch lengths.

null.model <- evo.model(tree = rand.data\$tree, Y = rand.data\$trait data,method = "Pairwise ML")

# Evolutionary rate(s) (sigma2mult):

# [1] 2.642536

# Log-likelihood: -257.9414

# Method: Pairwise ML

# Evolutionary model: BM

K.mult(model = null.model,nsim = 100)

# Bootstrapping under null model.

# \*\*\*\*\*Simulation results\*\*\*\*\*

# Test statistic (K) 0.9756736

# Critical test statistic 0.3503957

# Estimated Power 1.0000000

# P-value: 0

```
testKmult = K.mult(model = null.model,nsim = 100)
```

```
str(testKmult)
```

```
List of 5
```

```
$ K : num 0.912
```

```
$ Pval : num 0
```

```
$ power : num 1
```

```
$ K.expectation: num 1
```

```
$ plot.tools :List of 5
```

```
..$ test_statistic : num 0.912
```

```
..$ critical_test_statistic: num 0.268
```

```
..$ test_statistic_name : chr "K"
```

```
..$ null_sim_test_statistic: num [1:100] 0.0645 0.0881 0.0691 0.0617 0.0722 ...
```

```
..$ alt_sim_test_statistic : num [1:100] 0.702 0.937 0.984 0.929 0.999 ...
```

```
- attr(*, "class")= chr "compare.model"
```

```
testKmult$K
```

```
[1] 0.9115932
```

```
names(heat.tot.ploti)[-which(names(heat.tot.ploti) %in% phyloploti$tip.label)]
```

```
[1] "Weinmannia LC4951"
```

```
tree.3.rooted$tip.label[grep("Weinmannia", tree.3.rooted$tip.label)]
```

```
[1] "Weinmannia LC3488B" "Weinmannia MCM1203" "Weinmannia AF19282"
"Weinmannia MCM1102" "Weinmannia sorbifolia"
```

```
[6] "Weinmannia fagaroides" "Weinmannia AF13212" "Weinmannia lechleriana"
"Weinmannia AF16205" "Weinmannia LC4909"
```

```
[11] "Weinmannia reticulata" "Weinmannia davidsonii" "Weinmannia JGM42"
"Weinmannia LC4791" "Weinmannia LC4887"
```

```
[16] "Weinmannia haenkeana" "Weinmannia pinnata" "Weinmannia crassifolia"
"Weinmannia nebulorum" "Weinmannia ovata"
```

```
testphylo = tree.3.rooted
```

```
testphylo$tip.label[which(testphylo$tip.label == "Weinmannia LC4791")] =
"Weinmannia LC4951"
```

```
^^^
—
```

## **R code for Chapter 4**

title: "Chapter 4 R code"

output: word document

``{r}

library(vegan)

library(ggplot2)

library(dplyr)

library(sfsmisc)

library(tidyr)

library(picante)

library(matrixStats)

library(V.PhyloMaker)

library(FD)

library(ggfortify)

library(missForest)

library(picante)

path.to.files <-

"C:/Users/marvi/Box Sync/DavidHenderson/03 Chapter1/"

plot list <- c("PP Chaqui 32", "PP Fuerte 27", "PP Kanupa 44", "PP Lomaka 40",  
"PP Lomasa 39", "PP Pintat 5", "PP Resina 12", "PP Sumpul 34", "PP Tapuri 45",

```
"PP Tintay 24", "PP Tintay 25", "PP Titiri 42", "PP Tocoaq 28", "PP Tocoaq 29",
"PP Tocoaq 30", "PP Yarimi 9")
```

```
Load plot data
```

```
plot.data <- read.csv(paste0(path.to.files,
```

```
"02_Data/MadidiRawAndCleanData_v4.1/03_CleanMadidiData/01_PlotData_Clean_v4.1_2020
-09-07.txt"),
```

```
header = TRUE, sep = "\t", fileEncoding="UTF-8")
```

```
dim(plot.data)
```

```
colnames(plot.data)
```

```
plot.data <- plot.data[which(plot.data$PlotType=="PP"),]
```

```
Change name of Kanupa
```

```
unique(plot.data$PlotName)
```

```
plot.data$PlotName[19] <- "PP Kanupa 44"
```

```
plot.data$PlotName[which(plot.data$PlotName=="PP Kañupa 44")]<-
```

```
rep("PP Kanupa 44",length(which(plot.data$PlotName=="PP Kañupa 44")))
```

```
plot.data <- plot.data[which(plot.data$PlotName %in% plot_list),]
```

```
unique(plot.data$PlotName)
```

```
colnames(plot.data)
```

```

Load tree data
tree.data <- read.csv(paste0(path.to.files,
"02_Data/MadidiRawAndCleanData_v4.1/03_CleanMadidiData/04_TreeData_PP_Clean_v4.1
2020-09-07.txt"),
header = TRUE, sep = "\t", fileEncoding="UTF-8")

dim(tree.data)

colnames(tree.data)

#Change kanupa
unique(tree.data$PlotName)
tree.data$PlotName[19] <- "PP_Kanupa_44"

tree.data$PlotName[which(tree.data$PlotName=="PP_Kañupa_44")]<-
rep("PP_Kanupa_44",length(which(tree.data$PlotName=="PP_Kañupa_44")))

2022 OCT 27: For now remove all rows w/ NoName for binomial species name
#tree.data[which(".NoName" %in% tree.data$BinomialSpeciesName),]
tree.data <- tree.data[-which(tree.data$BinomialSpeciesName == ".NoName"),]

2022 NOV 08:
tree.data <- tree.data[-which(tree.data$BinomialSpeciesName == ".NoName AF18382"),]

2022 OCT 27: Fix any hyphenated names
#which(treedata$BinomialSpeciesName == "Guatteria sanctae-crucis") #
treedata$BinomialSpeciesName[14877] treedata[14877,]

```

```
tree.data$BinomialSpeciesName <- gsub("\\-", " ",
tree.data$BinomialSpeciesName)
```

```
tree.data <- tree.data[which(tree.data$PlotName %in% plot_list),]
```

```
unique(tree.data$PlotName)
```

```
colnames(tree.data)
```

```
Add in numeric for subplots of subplots here
```

```
tree.data.1 <- tree.data
```

```
tree.data.1$Subplot_subplot <- 0
```

```
#tree.data.1$Subplot_subplot <- tree.data.1$Subplot
```

```
tree.data.1 <- tree.data.1[which((tree.data.1$Ly >= 0) & (tree.data.1$Ly <= 20) &
(tree.data.1$Lx >= 0) & (tree.data.1$Lx <= 20)),]
```

```
tree.data.1$Subplot_subplot[which((tree.data.1$Ly >= 0) & (tree.data.1$Ly < 10) &
(tree.data.1$Lx >= 0) & (tree.data.1$Lx < 10))] <- 1
```

```
tree.data.1$Subplot_subplot[which((tree.data.1$Ly >= 0) & (tree.data.1$Ly < 10) &
(tree.data.1$Lx >= 10) & (tree.data.1$Lx <= 20))] <- 2
```

```
tree.data.1$Subplot_subplot[which((tree.data.1$Ly >= 10) & (tree.data.1$Ly <= 20) &
(tree.data.1$Lx >= 0) & (tree.data.1$Lx < 10))] <- 3
```

```
tree.data.1$Subplot_subplot[which((tree.data.1$Ly >= 10) & (tree.data.1$Ly <= 20) &
(tree.data.1$Lx >= 10) & (tree.data.1$Lx <= 20))] <- 4
```

```
tree.data.1$Subplot_subplot
```

```
nrow(tree.data.1[which(tree.data.1$Subplot_subplot == 0),])
```

```
tree.data.1$Subplot_quadrat <- paste(tree.data.1$Subplot, tree.data.1$Subplot_subplot, sep=".")
```

```
tree.data.1$Subplot_quadrat
```

```
tree.data <- tree.data.1
```

```
Load plot - climatic/ env data - create clim & soli PCA
```

```
path.to.files <-
```

```
"C:/Users/marvi/Box Sync/DavidHenderson/04_Chapter2/"
```

```
Load environmental data
```

```
env.data <- read.delim(paste0(path.to.files,
```

```
_____ "02_Data/12_EnviroRasterData_Clean_v4.1_2020-09-07.txt"))
```

```
env.data <- env.data[which(env.data$PlotType=="PP"),]
```

```
colnames(env.data)
```

```

Change Kanupa
env.data$PlotName[19] <- "PP Kanupa 44"
env.data$PlotName[which(env.data$PlotName=="PP KaÃ±upa 44")]<-
 rep("PP Kanupa 44",length(which(env.data$PlotName=="PP KaÃ±upa 44")))

env.data <- env.data[which(env.data$PlotName %in% plot_list),]
unique(env.data$PlotName)
colnames(env.data)

#Aggregate environmental data
plot.env.data <- merge(env.data, plot.data, by="PlotName")

colnames(plot.env.data)

#Run Clim PCA
imputation <- missForest(plot.env.data[, c(10,16,29,32)])
imputation <- imputation$ximp

plot.clim.summary <- cbind(plot.env.data[,1], imputation)
colnames(plot.clim.summary)[1] <- c("PlotName")
colnames(plot.clim.summary)

```

```

pca.clim.test <- prcomp(plot.clim.summary[2:5],
_____ center=TRUE, scale.=TRUE)

summary(pca.clim.test)

pca.clim.test$x[,1]

pca.clim.test$x[,2]

Plot

autoplot(pca.clim.test, data = plot.clim.summary, colour = 'PlotName',
_____ loadings = TRUE, loadings.colour = 'blue', loadings.label = TRUE, loadings.label.size = 3)

#Run Soil PCA

imputation <- missForest(plot.env.data[, c(72, 74:81, 83:86)])

imputation <- imputation$ximp

plot.soil.summary <- cbind(plot.env.data[,1], imputation)

colnames(plot.soil.summary)[1] <- c("PlotName")

colnames(plot.soil.summary)

Scale and log transform soil variables: DEC 02 2022

plot.soil.summary.1 <- data.frame(plot.soil.summary[,1],
_____ scale(log(plot.soil.summary[,2:14])))

```

```
pca.soil.test <- prcomp(plot.soil.summary.1[2:14],
 center=TRUE, scale.=TRUE)
```

```
summary(pca.soil.test)
```

```
pca.soil.test$x[,1]
```

```
pca.soil.test$x[,2]
```

```
Chemical Trait Data
```

```
path.to.files <-
```

```
"C:/Users/marvi/Box Sync/DavidHenderson/04 Chapter2/"
```

```
All Compounds
```

```
load(paste0(path.to.files,
```

```
"04 Results/Individual CompoundClasses/Madidi plotabundmeanCSCS metabtot 20220608.R
data"))
```

```
cscs.all <- as.matrix(mad_cscs_all)
```

```
Defense Compounds
```

```
load(paste0(path.to.files,
```

```
"04 Results/Individual CompoundClasses/MadAll metabstim npclass defense 20220614 2022
0806.Rdata"))
```

```
cscs.def <- as.matrix(1-cscs)
```

# Primary Compounds

load(paste0(path.to.files,

"04\_Results/Individual\_CompoundClasses/CSCSsppBCI-prim-20220822.Rdata"))

cscs.prim <- as.matrix(1-cscs)

# Shikimate Compounds

load(paste0(path.to.files,

"04\_Results/Individual\_CompoundClasses/MadAll\_metabsim\_npclass\_shikimates\_20220614.Rdata"))

cscs.shik <- as.matrix(1-cscs)

# Terpenoid Compounds

load(paste0(path.to.files,

"04\_Results/Individual\_CompoundClasses/MadAll\_metabsim\_npclass\_terpenoids\_20220614.Rdata"))

cscs.terp <- as.matrix(1-cscs)

dim(cscs.all)

dim(cscs.def)

dim(cscs.prim)

dim(cscs.shik)

dim(cscs.terp)

```

NOTE

Species in row and colum 384 is present for all and defense compounds,
but not for the primary compounds matrix. To standardize matrices, we
remove if from all matrices

#which(rownames(mad_cscs_all) == "PP Sumpul 34 Pleurothyrium trianae") #
mad_cscs_all[384, 384]

#Not included in CSCSsppBCI-prim-20220822.Rdata

cscs.all <- cscs.all[-384, -384]
cscs.def <- cscs.def[-384, -384]
cscs.shik <- cscs.shik[-384, -384]
cscs.terp <- cscs.terp[-384, -384]

identical(rownames(cscs.def), rownames(cscs.prim))
identical(rownames(cscs.shik), rownames(cscs.terp))
identical(rownames(cscs.shik), rownames(cscs.prim))

Change rownames, colnames

rownames(cscs.def) <- rownames(cscs.all)
colnames(cscs.def) <- colnames(cscs.all)
identical(rownames(cscs.all), rownames(cscs.def))

```

```
rownames(cscs.prim) <- rownames(cscs.all)
colnames(cscs.prim) <- colnames(cscs.all)
identical(rownames(cscs.all), rownames(cscs.prim))
```

```
rownames(cscs.shik) <- rownames(cscs.all)
colnames(cscs.shik) <- colnames(cscs.all)
identical(rownames(cscs.all), rownames(cscs.shik))
```

```
rownames(cscs.terp) <- rownames(cscs.all)
colnames(cscs.terp) <- colnames(cscs.all)
identical(rownames(cscs.all), rownames(cscs.terp))
```

```
NOTE
```

```
One species in the cscs matrix has no name "NA NA"
```

```
which(rownames(cscs.all) == "NA NA")
```

```
cscs.all <-
cscs.all[-which(rownames(cscs.all) == "NA NA"),
-which(colnames(cscs.all) == "NA NA")]
```

```
cscs.def <-
cscs.def[-which(rownames(cscs.def) == "NA NA"),
-which(colnames(cscs.def) == "NA NA")]
```

```
cscs.prim <-
cscs.prim[-which(rownames(cscs.prim) == "NA NA"),
-which(colnames(cscs.prim) == "NA NA")]
```

```
cscs.shik <-
cscs.shik[-which(rownames(cscs.shik) == "NA NA"),
-which(colnames(cscs.shik) == "NA NA")]
```

```
cscs.terp <-
cscs.terp[-which(rownames(cscs.terp) == "NA NA"),
-which(colnames(cscs.terp) == "NA NA")]
```

### Morphological Trait Data ###

```
path.to.files <-
"C:/Users/marvi/Box Sync/DavidHenderson/03 Chapter1/"
```

```
trait.data <- read.csv(
paste0(path.to.files,
"02_Data/MadidiRawAndCleanData v4.1/03 CleanMadidiData/06 TraitData Clean v4.1 202
0-09-07.txt"),
header = TRUE, sep = "\t", fileEncoding="UTF-8")
dim(trait.data)
colnames(trait.data)
```

```

Change name of Kanupa
unique(trait.data$PlotName)
trait.data$PlotName[7] <- "PP Kanupa 44"

trait.data$PlotName[which(trait.data$PlotName=="PP Kañupa 44")]<-
 rep("PP Kanupa 44",length(which(trait.data$PlotName=="PP Kañupa 44")))

trait.data <- trait.data[which(trait.data$PlotName %in% plot_list),]
unique(trait.data$PlotName)
colnames(trait.data)

Calculate mean by species and plot
traits.to.use.1 <- c("RGR", "SLA", "LeafSize", "LeafThickness",
 "TwigBarkThickness Relative", "TwigSpecDens")

traits.to.use.2 <- c("Height", "DBH")

mean.traits <- aggregate(trait.data[,traits.to.use.1],
 by = list(trait.data$BinomialSpeciesName, trait.data$PlotName),
 FUN = "mean", na.rm = TRUE)

max.traits <- aggregate(trait.data[,traits.to.use.2],
 by = list(trait.data$BinomialSpeciesName, trait.data$PlotName),
 FUN = "quantile", probs=0.90, na.rm = TRUE)

```

```
traits.summary <- data.frame(mean.traits[,1:2],
 mean.traits[,traits.to.use.1],
 max.traits[,traits.to.use.2])
```

```
traits.summary <- data.frame(mean.traits[,1:2],
 scale(log1p(mean.traits[,traits.to.use.1])),
 scale(log1p(max.traits[,traits.to.use.2])))
```

```
colnames(traits.summary)[1:2] <- c("BinomialSpeciesName", "PlotName")
```

```
colnames(traits.summary)
```

```
Imputation of missing values in trait matrix
```

```
imputation <- missForest(traits.summary[,-c(1,2)])
```

```
imputation <- imputation$ximp
```

```
traits.summary <- cbind(traits.summary[,c(1,2)], imputation)
```

```
head(traits.summary)
```

```
unique(traits.summary$PlotName)
```

```
colnames(traits.summary)
```

```

Run PCA for morphological traits

pca.test.morpho <- prcomp(traits.summary[, 3:10],
 center=TRUE, scale.=TRUE)

summary(pca.test.morpho)

Plot

autoplot(pca.test.morpho, data = traits.summary, colour = 'PlotName',
 loadings = TRUE, loadings.colour = 'blue', loadings.label = TRUE, loadings.label.size = 3)

traits.summary$PCA1 <- pca.test.morpho$x[,1]
traits.summary$PCA2 <- pca.test.morpho$x[,2]
traits.summary$PCA3 <- pca.test.morpho$x[,3]
traits.summary$PCA4 <- pca.test.morpho$x[,4]
traits.summary$PCA5 <- pca.test.morpho$x[,5]
traits.summary$PCA6 <- pca.test.morpho$x[,6]
traits.summary$PCA7 <- pca.test.morpho$x[,7]
traits.summary$PCA8 <- pca.test.morpho$x[,8]

colnames(traits.summary)

rownames(pca.test.morpho$x) <- paste(traits.summary$PlotName,
 traits.summary$BinomialSpeciesName, sep=" ")

trait.dists <- as.matrix(dist(pca.test.morpho$x))

```

```
Need to change code to save obs.mpd as well
```

```
Null randomization SES.MPD
```

```
plot_list <- c("PP Chaqui 32", "PP Fuerte 27", "PP Kanupa 44", "PP Lomaka 40",
"PP Lomasa 39", "PP Pintat 5", "PP Resina 12", "PP Sumpul 34", "PP Tapuri 45",
"PP Tintay 24", "PP Tintay 25", "PP Titiri 42", "PP Tocoaq 28", "PP Tocoaq 29",
"PP Tocoaq 30", "PP Yarimi 9")
```

```
Functional Distances and SES.MPD Null
```

```
plot.MPD.df <- data.frame(PlotName=plot_list)
```

```
for(p in 1:length(plot_list))
```

```
{
```

```
 plot = plot_list[p]
```

```
 plotdata = droplevels(tree.data[which(tree.data$PlotName == plot),])
```

```
 plot.compo.temp <- table(plotdata$Subplot, paste(plotdata$PlotName,
 plotdata$BinomialSpeciesName, sep="_"))
```

```
 cscs.temp <- cscs.def
```

```
 spp.intersecting <- intersect(intersect(rownames(cscs.temp),
 colnames(plot.compo.temp)), rownames(trait.dists))
```

```

cscs.temp <- cscs.temp[spp.intersecting, spp.intersecting]

trait.dists.temp <- trait.dists[spp.intersecting, spp.intersecting]

plot.compo.temp <- as.data.frame.matrix(plot.compo.temp[, spp.intersecting])

#cscs.temp <- as.dist(cscs.temp)

plot.MPD.chem.subquad <- ses.mpd(samp=plot.compo.temp, dis=cscs.temp,
null.model="taxa.labels",
abundance.weighted=T, runs=999)

#trait.dists.temp <- as.dist(trait.dists.temp)

plot.MPD.morpho.subquad <- ses.mpd(samp=plot.compo.temp, dis=trait.dists.temp,
null.model="taxa.labels", abundance.weighted=T, runs=3)

mpd.obs.chem <- mean(plot.MPD.chem.subquad$mpd.obs, na.rm=T)

mpd.z.chem <- mean(plot.MPD.chem.subquad$mpd.obs.z, na.rm=T)

mpd.obs.morpho <- mean(plot.MPD.morpho.subquad$mpd.obs, na.rm=T)

mpd.z.morpho <- mean(plot.MPD.morpho.subquad$mpd.obs.z, na.rm=T)

Save results

```

```

Chem

plot.MPD.df[p, 2] <- mpd.obs.chem

plot.MPD.df[p, 3] <- mpd.z.chem

Morpho

plot.MPD.df[p, 4] <- mpd.obs.morpho

plot.MPD.df[p, 5] <- mpd.z.morpho

}

colnames(plot.MPD.df)[2] <- "mpd.obs.chem"

colnames(plot.MPD.df)[3] <- "mpd.z.chem"

colnames(plot.MPD.df)[4] <- "mpd.obs.morpho"

colnames(plot.MPD.df)[5] <- "mpd.z.morpho"

Save dataframe for later

write.csv(plot.MPD.df, paste0(path.to.files,
_____ "04 Results/"))

```