Determinism vs. Stochasticity in Community Assembly Processes: The Role of Species Phylogeny and Dominance

by

Carlos Alberto Arnillas Merino

A thesis submitted in conformity with the requirements for the degree of Doctor of Philosophy

Department of Physical and Environmental Sciences University of Toronto

© Copyright by Carlos Alberto Arnillas Merino 2019

Determinism vs. stochasticity in community assembly processes: The role of species phylogeny and dominance

Carlos Alberto Arnillas Merino

Doctor of Philosophy

Department of Physical and Environmental Sciences University of Toronto

2019

Abstract

The deterministic or stochastic nature of the rules that regulate which species co-exist in a community has been long debated in ecology. This thesis aims to answer (1) if understanding the degree of determinism in the co-existence of species is more or less informative than other sources of uncertainty affecting regional biodiversity, (2) if biodiversity is important to understand ecosystem processes, and (3) if dominance and determinism are correlated.

Using a landscape model I show that the degree of determinism or stochasticity (1) controls regional species richness, and (2) causes more uncertainty on species richness than land use change or climate change in the Tropical Andes.

In herbaceous communities around the world, diversity is important to improve predictions of biomass productivity and litter accumulation mostly when the diversity and biomass descriptors of the communities were partitioned into legumes, forbs and graminoids. Graminoids, grasses and *Carex*, and one forb lineage were more likely to be dominant species than non-dominant, while more than a dozen forb lineages were more likely to be non-dominant species.

The environment affected the dominant plants the most, increasing the role of habitat filtering, with a less common effect of limiting similarity; in contrast, non-dominant plants were mostly affected by limiting similarity among the non-dominants, but with no signal of habitat filtering. This pattern was observed in herbaceous systems using a local dominant removal experiment and a global observational dataset.

The results show that dominant species are a more deterministic subset of species that converge towards a predictable optimum constrained by environmental conditions. In contrast, non-dominant species are better described as a diverging group of species, potentially with multiple optimums, and therefore with a less predictable response. Understanding the differences between dominant and non-dominants can improve models of ecosystem services that rely either on biomass accumulation or on diversity.

Acknowledgments

The last six years have been a long journey for me in a myriad of ways. It has been a long and life-changing adventure for my family, Carmen, Santiago and Andres, to whom I am most thankful for their ongoing support and patience. Your daily presence, our hopes and dreams, motivated me to pursue this project.

My family was the start, the goal and the path of the journey. However, the journey would not have existed without the generous and constant support of Dr. Marc Cadotte. I am extremely grateful for the opportunity he gave me to learn and investigate about how plant communities organize and work, for his trust, but also, and probably as important, for the opportunity to learn about how to disseminate our discoveries in science. Being part of a vibrant academic community is a unique opportunity and I am happy I had the opportunity to participate in one.

The ideas that I began with Marc for this thesis would never been fully shaped and framed without the comments and feedback from Dr. Marney Isaac and Dr. Arthur Weis. Their suggestions along the journey helped me to be conscious of the bigger picture, and pushed me to better communicate my ideas. I also want to thank Adam Martin, who provided insightful comments in the last version of this manuscript.

Prof. Isaac also provided me with lab support for some of the analyses related to my field experiment, as did Prof. Roberta Fulthorpe. Their students, specially Serra Buchanan and Rosemary Saati, helped me to complete several analyses that will be part of subsequent studies and have been already included in some undergraduate theses. I would like to thank Stephan Schneider at the Koffler Scientific Reserve for keeping an eye on my experiment, his assistance with many aspects of the projects, and with a bike.

At University of Toronto - Scarborough, I owe gratitude to many staff members who were amazing in helping me along through the years. Special thanks to Elaine Pick, Shelley Eisner, Julie Quennville, Joanna Ying and many others at the Department of Physical and Environmental Sciences and at the Center for International Experience. I want to thank the DPES staff that trusted me and helped me to get the Connaught International Scholarship for Doctoral Students. That scholarship, and the funds provided by TD Bank Group and the Natural Science and Research Council of Canada to Marc Cadotte supported my research. I am also grateful to the dozens of undergraduate students who spent long hours with me in the field pulling grasses, one-by-one, until not a single thin grass leaf was left. James, Allanah, Katerina, Ivy, Diana, Kaushar, Shajeentha, Rhokini, Ronald, Bill, Mandy, Antonio, Shatabdy, Justin, Apsara, Tiffany, Michelle, Shajaat, Marley, Fat, Diana, Wei, Reema are just a few of them. Your patience and amazing, positive attitude made this project possible. Among them, I want to mention Nicholas, Michell, Garland, Shameek, Jissan, Marylouisse, Henil, Meera, Parsa and Pratyasha who trusted me as a mentor with their undergraduate thesis projects. I hope that the hours we spent together sharing ideas, experiences, philosophy, or just working or swimming under sun and rain may have helped you, somehow, in your own paths.

Along these years, many of those that worked alongside me became close friends. Many of them were part of the Cadotte Lab for a few days, months or years. Life would have been harder without Simone, Scott, Stuart, Kelly, Adriano, Carina, Marta, Xingfeng, Sebastian, Darwin, Penny, Shuai, Jeya, Reham, Alice, Lanna, Caroline, Nicholas, and many of the undergraduates I already mentioned. So many people in my own lab and in other labs around the campus and at the Koffler Reserve helped to make my dissertation possible.

Of course, the Cadotte Lab would not be the same vibrant and integrated workplace without Shirley's presence, her time, and her dedication that help make the lab a great space for everybody both personally and professionally.

I am also grateful to Marc and the scientists of the Nutrient Network for the opportunity to participate in their amazing research community. I want to express my gratitude to many of the NutNetters for their support, collaboration and their research, particularly to Jennifer, Kim, Eric and Elizabeth. Eric and Elizabeth, thanks for keeping this wonderful research initiative running, and for the opportunity to participate in it.

Thanks to Robin and Ken, and Mauricio and Mariana, for their friendship and endless support, not just for me, but also for my family. Any list will be incomplete, but I will always remember some of the best suggestions to strategize the research project, academic discussions, dinners together, and a big etcetera. But probably nothing more important than making us feel as among family in this new country. You are awesome. Our Peruvian family, colleagues and friends also accompanied me in this journey. Carolina and Wouter are friends, colleagues and co-authors. Fernando, Carmen, Jorge, Cecilia, Valeria, Alicia, Wendy, Ana María, Amaru are some of the family and friends who visited us, supported us, and gave us the opportunity to share this part of the world with them. Of course I am grateful to Fico, Liani and Ana Maria, my brother, aunt and grandmother: without your unconditional support to my family and I, we, would never manage to finish this journey, and be ready for the next one. Among my closest family are my parents, Fico and Maria Isabel: when I think of you, I realize that this journey began many years before I first contacted Marc. Thanks for all those years.

This list is incomplete, and I rather think of it as a sample of an amazing community of extended family, friends and colleagues that I am extremely happy and grateful that my family and I had the opportunity to know, and who have helped me complete this research. Thanks again.

Table of Contents

Acknowledgments	iv
Table of Contents	vii
List of Tables	xiii
List of Figures	xiv
Chapter 1 Introduction	1
1.1 Stochasticity-determinism and neutral-niche axes in community assembly theory	3
1.2 The three-filters analogy of community assembly theory, co-existence theory and phylogenetic dispersion	4
1.3 Distinguishing between stochastic and deterministic communities	5
1.4 Plants as engineers and the mass-ratio-hypothesis: is there a correlation between the stochasticity experienced by a species and its dominance?	6
1.5 Thesis statement	8
Chapter 2 From patches to richness: Assessing the potential impact of landscape transformation on biodiversity	11
2.1 Abstract	11
2.2 Introduction	12
2.3 The Conceptual Framework: Assessing the Richness of a Community in a Patchy Environment	14
2.3.1 Building the toolbox	18
2.3.1.1 Patch-size profile: Visualizing habitat availability and inferring general richness patterns	18
2.3.1.2 Quantifying the potential impact of landscape changes on biodiversity	21
2.3.2 Applying the conceptual framework	24
2.4 Methods	25
2.4.1 Spatial database	25
2.4.2 Assessing fragmentation impact on Andean biodiversity	25

	2.4.3	Uncerta	inty in habitat richness under present and future conditions	26
	2.4.4	Effect o	f spatial reconfiguration in extreme LEM-communities per habitat	27
2.5	Result	s		27
	2.5.1	How do biodiver	es the current habitat configuration of Andean biomes impact their rsity?	27
	2.5.2	Present	and future habitats' richness uncertainty	29
	2.5.3	Effect o	f spatial reconfiguration in extreme LEM-communities per habitat	31
	2.5.4	Sensitiv	ity analysis	34
2.6	Discus	sion		34
	2.6.1	Describ	ing Andean biomes richness	34
	2.6.2	General	framework	37
2.7	Conclu	usion		40
Chapte don	er 3 ninant p	Domina lants	nt plants in a meadow are more deterministically assorted than non-	41
3.1	Abstra	ct		41
3.2	Introdu	uction		41
3.3	Metho	ds		45
	3.3.1	Experin	nental design	45
	3.3.2	Phyloge	enetic information	46
	3.3.3	Abunda	nce and definition of new dominants	46
	3.3.4	Statistic	al analysis	47
		3.3.4.1	Treatment efficacy	47
		3.3.4.2	Deterministic vs. stochastic trends	47
		3.3.4.3	Different deterministic processes: limiting similarity and habitat filtering	48
		3.3.4.4	Testing neutrality using species trends	49
3.4	Result	s		50

	3.4.1	Treatment efficacy	50
	3.4.2	Deterministic vs. stochastic trends	50
	3.4.3	Different deterministic mechanisms: Limiting similarity and habitat filtering	51
	3.4.4	Testing neutrality using species trends	53
3.5	Discus	sion	55
	3.5.1	The dominant species	56
	3.5.2	The non-dominant species	57
	3.5.3	Long-term trends	58
	3.5.4	A note on traits and phylogenies	60
3.6	Conclu	usions	60
Chapte asse	er 4 embly r	Dominant species converge, non-dominant species diverge: dominance, ules and phylogenies interact in herbaceous ecosystems around the world	61
4.1	Abstra	ct	61
4.2	Introd	uction	61
4.2 4.3	Introdu Metho	uction	61 64
4.2 4.3	Introdu Metho 4.3.1	ds Data sources	61 64 64
4.2 4.3	Introdu Metho 4.3.1	ds Data sources 4.3.1.1 Phylogenetic information	61 64 64 64
4.2 4.3	Introdu Metho 4.3.1	ds Data sources 4.3.1.1 Phylogenetic information 4.3.1.2 Site level data	61 64 64 65
4.2 4.3	Introdu Metho 4.3.1 4.3.2	ds Data sources 4.3.1.1 Phylogenetic information 4.3.1.2 Site level data Are dominant and non-dominant species similarly assorted? Estimating the relatedness of dominant, non-dominants and their disparity	61 64 64 65 65
4.2	Introdu Metho 4.3.1 4.3.2 4.3.3	ds Data sources 4.3.1.1 Phylogenetic information 4.3.1.2 Site level data Are dominant and non-dominant species similarly assorted? Estimating the relatedness of dominant, non-dominants and their disparity Are certain lineages more likely to be either dominant or non-dominant?	61 64 64 65 65 68
4.2	Introdu Metho 4.3.1 4.3.2 4.3.3 4.3.4	ds Data sources	61 64 64 65 66 68
4.2 4.3 4.4	Introdu Metho 4.3.1 4.3.2 4.3.3 4.3.4 Result	ds Data sources	61 64 64 65 66 68 69 70
4.24.34.4	Introdu Metho 4.3.1 4.3.2 4.3.3 4.3.4 Result 4.4.1	ds Data sources	61 64 64 65 66 68 69 70

	4.4.3	Are there environmental conditions, topological characteristics of the phylogeny or biogeographic aspects that drive the relatedness disparity?	.73
4.5	Discus	sion	.75
	4.5.1	Dominant and non-dominant species follow different ecological mechanisms	.75
	4.5.2	Environmental drivers	76
	4.5.3	Non-dominance as a strategy	.77
	4.5.4	Limitations and caveats	.78
	4.5.5	Concluding remarks	79
Chapte cau	er 5 sality so	Can partitioning plant communities by life history and scale dependent olve the productivity-biodiversity conundrum?	.80
5.1	Abstra	ct	.80
5.2	Introdu	action	.80
5.3	Metho	ds	.83
	5.3.1	Community information	.83
	5.3.2	Phylogenetic information	.84
	5.3.3	Statistical analysis	.85
		5.3.3.1 Potential biodiversity-biomass-litter correlation	.85
		5.3.3.2 Realized biodiversity-biomass-litter correlation	.88
5.4	Result	5	.90
	5.4.1	Biodiversity-productivity relationships among sites	.90
	5.4.2	Local analysis	94
5.5	Discus	sion	.95
Chapte	er 6	Conclusions	.98
6.1	Key re	sults	.99
6.2	Movin	g forward	101
	6.2.1	Ideas coming from a landscape perspective	101
	6.2.2	Partitioning the community using a dominance criterion	102

	6.2.3	From diversity to ecosystem processes and services	103
6.3	Final	remarks	104
Refere	nces		106
Appen	dices.		121
А	Appe	ndices of chapter 2	121
	A.1	Supplementary table	121
	A.2	Supplementary figures	122
	A.3	ISAR curves in nature and models: the prevalence of power-law	129
		A.3.1 Neutral theory	130
		A.3.2 Species fragmented area relationship: An ISAR for highly fragmented habitats	131
		A.3.3 References	132
	A.4	Proof of maximum richness when a patch is symmetrically divided	133
		A.4.1 General ISAR shape: Two even patches can foster more species than a single large one	133
		A.4.2 Arrhenius ISAR: Several even patches can foster more species than uneven ones	133
В	Appe	ndices of chapter 3	136
	B.1	Supplementary tables	136
	B.2	Supplementary figures	139
С	Appe	ndices of chapter 4	142
	C.1	Supplementary tables	142
	C.2	Supplementary figures	144
	C.3	Alternative definitions of disparity in phylogenetic dispersion	146
		C.3.1 Detecting Partition methods	146
		C.3.2 Comparison among partition methods	149
	C.4	Testing other legumes' proportion, biomass and their interaction with climatic variables	150

		C.4.1	Legumes as a response variable	151
		C.4.2	Legumes as a predictor variable	151
		C.4.3	Conclusions	152
	C.5	Other as	spects of the structural equation model	153
		C.5.1	Results	153
		C.5.2	Discussion	153
	C.6	Are C3/	C4 pathways relevant to characterize graminoids prevalence?	154
D	Appe	endices of	chapter 5	156
	D.1	Supplen	nentary tables	156
	D.2	Supplen	nentary figures	159
	D.3	Detailed biomas	description of structural equation models: potential biodiversity-	161
		D.3.1	Final model	161
		D.3.2	Building a full community model	167
		D.3.3	Single step-backwards regression	168
		D.3.4	Other climatic variables	169
		D.3.5	Including phylogenetic descriptors and species richness together	170
		D.3.6	Using PD but not MPD or MNTD	170
		D.3.7	Alternative litter models	172
		D.3.8	Building the step-backwards models without diversity info	173
		D.3.9	Productivity as a driver of diversity	173
	D.4	Details	of the realized biomass-biodiversity models	174
		D.4.1	Final Faith's PD model	175
		D.4.2	Final species richness model	179
		D.4.3	Using the same set of sites as for the global model	183

List of Tables

Table 1 R-values for current potential and remnant Andean biomes. 28
Table 2 Main sources of uncertainty in the species richness calculations of the Andean biomes
during the period 2040-2069
Table 3 Changes in species ranking in the dominant removal treatments compared with the
expectations for a neutral model
Table 4 Effect of treatments in new dominant species in comparison with the control plots at the
final visit of the experiment
Table 5: Expected and observed global phylogenetic dispersion patterns of the dominant (D _{ses})
and non-dominant (ND _{ses}) partitions and the relatedness disparity ($\Delta_{ses} = D_{ses} - ND_{ses}$)
Table 6 Effect of richness and phylogenetic diversity metrics and of partitioning the community
into functional groups on litter at a global and local scale
Table 7 Effect of the direction of the relationship between diversity and biomass descriptors with
basic structural equation models of diversity, biomass and litter globally and locally
Table 8 Structural equation model representing the effect of (1) location on climate; (2) climate,
management and phylogenetic relatedness on functional group biomass production; and (3)
management, phylogenetic relatedness and functional group biomass production on litter
accumulation

List of Figures

Figure 1 Key characteristics that influence the number of species that can survive in a
fragmented habitat
Figure 2 Effect of the maximum individual dispersal in the habitat configuration, represented by
the number of patches and the proportion of the total habitat in the largest patch
Figure 3 Construction and interpretation of a patch size profile
Figure 4 Connectivity profile using R-values and their relationship with patch metrics often used
in fragmentation analysis
Figure 5 Patch size profile for current potential (black lines) and remnant biomes (gray lines), assuming zero individual dispersal
Figure 6 Connectivity profiles for evergreen forest, montane shrubland and xeric puna at present
time for different maximum individual dispersal capabilities (x-axis) and for both tolerant (top)
and intolerant (bottom) communities
Figure 7 Patch-size profile for projected future biomes under different climate models (gray
lines) for potential and remnant areas, assuming zero individual dispersal
Figure 8 Expected dynamic of several plots after dominant plants are removed (vertical dark gray
bar) and some non-dominant species occupy the released space
Figure 9 Dominant species interact more with the environment and with other dominant species
(often intraspecific interaction) than with non-dominant species, while non-dominant species
follow the opposite pattern
Figure 10 Effect of the treatments on community dissimilarity measured as the proportion of
times the measured Bray-Curtis dissimilarity was lower than the dissimilarity before the
treatments were applied
Figure 11 Partial residuals (dots), confidence band (gray box), and prediction line (blue solid
line) of the standardized effect size (z-scores) of the mean phylogenetic distance (SES-MPD) and

mean nearest taxon distance (SES-MNTD) for each treatment and for all the community (All,
panels a,b), dominant species as defined at the end of the experiment (panels c,d) and non-
dominant species (panels e,f)
Figure 12: Initial model representing the basic relationships (solid black arrows) of legume mass
(LM), total biomass (BM), proportion of graminoids (pGr) and difference in the dispersion of the
dominants and non-dominant plants (Δ_{ses})
Figure 13: Phylogenetic dissimilarities among sites when each site is partitioned into dominant,
intermediate dominance and non-dominant species, each partition with a third of the species72
Figure 14: Phylogenetic tree of species observed in the experiment showing the probability of a
lineage to be dominant, intermediate or non-dominant
Figure 15: Causal paths between site location (latitude, Lat; elevation, Elev; hemisphere, Hem),
climate (precipitation, P; temperature range, TR; mean temperature, T), tree topology (Faith's
PD, PD), management (anthropogenic, An; grazing, Gz), productivity (legume and total
aboveground biomass, LM and BM, respectively) their effect on the proportion of graminoids
(pGr) and phylogenetic dissimilarity between dominants and non-dominants ($\Delta_{ses.mntd}$)
Figure 16 Basic models representing a situation where (a,b) diversity influences biomass and
(c,d) biomass influences diversity
Figure 17 Directed acyclic graph representing the effect of (1) location on climate; (2) climate,
management and phylogenetic relatedness on functional group biomass production; and (3)
management, phylogenetic relatedness and functional group biomass production on litter
accumulation
Figure 18 Explanatory power of the different methods used in this study at the global scale 93

Chapter 1 Introduction

Living systems are fascinating and complex. The recurrent interaction among species creates a forever-changing arena where the species compete and facilitate each other directly and indirectly. In this process, the species transform and organize matter and energy in their surroundings in ways that can currently be coarsely described (Holdridge 1967, Rosenzweig 1995), but hardly detailed (Laughlin 2014, Storch et al. 2018). Several mechanisms have been proposed to predict which species will eventually co-exist and form a community, ranging from stochastic to deterministic mechanisms (MacArthur and Levins 1964, MacArthur and Wilson 1967, Tilman 1980, Hanski 1982, Weiher and Keddy 1995, Chesson 2000, Hubbell 2001a).

One of the most exciting academic puzzles of our time is the study of the rules that explain species co-existence, a field known as community assembly (Weiher et al. 2011). This theoretical question may be particularly relevant in the present context because global environmental changes are impacting ecosystems worldwide and endangering our society's means of survival (Millennium Ecosystem Assessment 2005). A large amount of effort is continuously devoted to quantify ecosystem services such as carbon storage, water regulation and pollination (e.g. Guo et al. 2000, Fischlin et al. 2007, De Deyn et al. 2008). The quantification often accounts for how much benefits an ecosystem actually provides to a society and for ways to economically assess those benefits for the society.

The connections between ecosystem changes and the associated impacts to society are, however, not commonly modeled in large-scale analysis of climate change or land use impacts. For instance, the Stern report on the global impact of climate change (2007) projected the impact of climate change on the global economy without the most basic causal link: changes in climate causes changes in ecosystems, which in turn causes changes in ecosystem services provided by those ecosystems, changing the benefits we obtain from the environment. One of the efforts to fill that gap was the creation of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services (IPBES) (https://www.ipbes.net/) in 2012.

The fact that this gap exists in current models that link global change and national and local impacts to ecosystem services is understandable given the little amount of information available,

compared to the amount of information they may require. On the one hand, despite the many advances in ecosystem descriptions that allow some generalizations (Lavorel et al. 2010, Cadotte 2013, 2017), the models currently used to predict potential biodiversity changes require large amounts of information, which is often unavailable. These models are also theoretically and computationally complicated and often have little explanatory power (Thuiller et al. 2008, Steinmann et al. 2009, Kissling et al. 2011, Wisz et al. 2013, Ramirez-Villegas et al. 2014). On the other hand, some evidence indicates that so much detail can be of limited value: some simple empirical models can predict basic trends of biomass production without any explicit reference to the biodiversity in an area (Holdridge 1967), while other theoretical models suggest that diversity details may be of little importance because of strong physical constraints (Brown et al. 2004). If we add to this list the fact that there is little capability to generalize common patterns in ecosystems (Lawton 1999), then the ecosystem gap in regional and global models of ecosystem services is totally justifiable.

However, accepting that biodiversity is uninformative is counterintuitive for most ecologists, including myself. As do many, I think that there is a moral value in preserving and caring for the other species living on this planet. Besides, I also think that there is logical ground and enough evidence to assert that, as living organisms shape ecosystem services from carbon sequestration to pollination, biodiversity should be informative of these services (Tilman et al. 1997, Lavorel et al. 2010, Cadotte 2013, Garibaldi et al. 2016, Duffy et al. 2017, Manning et al. 2018). For instance, as species differ in their shape, chemical characteristics and requirements, the species will often differ in how they alter the environment (Lavorel and Garnier 2002). For instance, trees capture more carbon than forbs or grasses, legumes can help to fix nitrogen, and pollination is more benefitted by the diversity of bees than by an average single species (Garibaldi et al. 2016). Therefore, biodiversity should be informative, but it is not clear how much. In other words, if we were interested in predicting the services that a grassland can provide how important will be to know if a legume or a grass will survive in that grassland? And given the prevalence of noise and stochasticity in natural systems, should we be more interested in local drivers of species performance or in processes occurring outside that location?

2

1.1 Stochasticity-determinism and neutral-niche axes in community assembly theory

Stochasticity and determinism in the laws of nature is a common debate in philosophy and science that has existed for a long time, even before it became part of the community assembly theory discussion. Currently, stochastic processes are those that are better described by probability functions (Tucker et al. 2016) while determinism indicates a mostly predictable system. These probability functions can either represent truly random processes or just the combined effect of unknown deterministic processes (Clark et al. 2007). Ecology has long recognized the role of stochasticity in community assembly processes (Fisher et al. 1943, MacArthur 1955, Al-Mufti et al. 1977, Tilman 1980, Hanski 1982), but neutral models (MacArthur and Wilson 1967, Hubbell 2001a) are different because they assume that mechanisms are independent of species characteristics (Vellend et al. 2014).

Niche models are based on the Hutchisonian definition of the niche (Hutchinson 1978) to describe different performances of species in an environment. In contrast to neutral models, local species abundance in niche models is controlled by both the correlation between the species fundamental niche and the environmental conditions of the site, and by the interactions with other species, especially competition and trophic interactions (MacArthur and Levins 1967, Hutchinson 1978, e.g. Hanski 1982, Fukami 2015, Tucker et al. 2016). Niche models were traditionally validated because the patterns they can create were often qualitatively similar to those observed in natural conditions.

Neutrality is the assumption that all species are equivalent. Hubbell and others have shown that neutral models also generate realistic patterns like the ones originated by niche models (Hubbell 2001a, McGill et al. 2007). In Hubbell's model, regional species abundances are totally controlled by stochastic processes, while local species abundances result as a balance between (1) random mortality identical for all the individuals and (2) species migration probability (which is proportional to the abundances of the species in the region surrounding the local area). Therefore, at the local scale, the main difference between niche and neutral theories is if the uneven species abundances are caused by internal mechanisms (niche theories) or external constraints (neutral theories).

1.2 The three-filters analogy of community assembly theory, co-existence theory and phylogenetic dispersion

Community assembly theory often accounts for three filters or mechanisms to explain some patterns in the species co-existing in a location (Weiher et al. 2011). The first filter is dispersal limitation, which defines the group of species in the region, i.e. the species pool, that can migrate to the location. This filter can be neutral as mentioned previously (oscillating by large scale stochastic mechanisms, Hubbell 2001a), niche-based (e.g. controlled by species dispersal traits, Lowe and McPeek 2014) or fixed (and therefore left outside of the model, particularly useful when models focus in the next two filters). The second filter, *habitat filtering*, is non-neutral and is caused by the correspondence between the fundamental niche of the species and the abiotic conditions of the habitat that limit the species that can maintain positive fitness in the area (Hutchinson 1978, Weiher and Keddy 1995). The third filter, often called *limiting similarity*, is also non-neutral and is caused by the competition among species that make dissimilar species more likely to co-exist (Gause 1932, MacArthur and Levins 1967, Weiher and Keddy 1995).

Habitat filtering and limiting similarity filters are both based on species characteristics, which are not randomly distributed, but constrained by their evolutionary history (Gould 1990). Because relatedness is positively correlated with similarity, we expect to see similar species co-existing when habitat filtering is strong, creating an "under-dispersed" pattern in the phylogeny, i.e. co-existing species belong to fewer lineages than expected by random (Webb 2000). Conversely, when limiting similarity is strong co-existing species should tend to belong to different lineages in the phylogeny, creating an over-dispersed phylogenetic pattern. With the widespread access to phylogenetic information, these relatively simple predictions have proven useful to test the generality of the community assembly model worldwide (Cadotte and Davies 2016, cf. Gerhold et al. 2015).

A different point of view was proposed by Chesson's coexistence theory (2000). In his theory, Chesson proposes that co-existence of species can be promoted by stabilizing and equalizing mechanisms. In his frame, niche differences could promote co-existence (stabilizing effects). But when niches are similar, similarity in the performance of the species (i.e. fitness) will promote co-existence by delaying the extinction of competing species (equalizing effects), mostly because

4

the extinction odds of both species will be too similar. Only in that scenario, when the same niche is occupied by similar species with similar fitness differences, species interactions became neutral, and the role of stochastic mechanisms increases. Mayfield and Levine (2010) extended the model to include phylogenetic patterns, and pointed out that under some conditions, competition could also lead to underdispersion or clustering, particularly when the main environmental constraint in the habitat has a single optimum strategy to deal with it (e.g. plants competing for light are required to be tall, therefore similar species that grow tall are more likely to co-exist if they have small fitness differences or other niche differences allowing coexistence). Further, if relatedness controls niche and fitness similarities, we should expect closely related species to co-exist via equalizing effects (similar niche, similar fitness) and distantly related species via stabilizing effects (dissimilar niche, fitness not relevant).

1.3 Distinguishing between stochastic and deterministic communities

The long debate about the role of stochasticity in nature is not yet solved, with contrasting results from real systems. Experimental and observational evidence started to accumulate showing that no pure model was able to properly describe every ecosystem pattern (Chase 2007, Li et al. 2015). The contrasts arose not only between studies reporting different mechanisms, but even those studies designed to tease out the mechanisms reported that several mechanisms may be valid simultaneously (Gilbert et al. 2009).

Several reasons explain the challenges in identifying the underlying nature of the community assembly process. One of them is that neutral and niche theories predict similar patterns (McGill et al. 2007). Despite the large contrast in the assumptions underlying both types of theories, without clear distinctions in the predictions, there is no way to tease them apart.

Another problem arises because the two main deterministic mechanisms of the community assembly process (i.e. habitat filtering and limiting similarity) theoretically produce opposing signals. As previously mentioned, habitat filtering could produce underdispersion in the phylogeny, while limiting similarity could produce overdispersion. For instance, if both mechanisms are strong (or weak), which pattern should prevail (Chalmandrier et al. 2013)?

Yet another problem with the community assembly framework is that it focuses on negative interactions, with filters always reducing the species diversity, despite the evidence that positive interactions are common and widespread (Brooker and Callaway 2009). This is unsurprising, as the largest theoretical developments related to niche theory are strongly linked to the Hutchinsonian formalization of the niche (Soberón 2007), which highlights the *requirements* of the species, more than the transformation of the environment that the species create.

A final problem arises from the different ways stochasticity is interpreted and measured (Clark et al. 2007, Vellend et al. 2014). One of the oldest and most basic definitions of determinism (and by opposition, stochasticity) is a process, that when repeated several times under similar conditions, yields always the same result. Unfortunately, controlling similar conditions is practically impossible in natural conditions; historical contingency (the deterministic or stochastic dependency of a state on the history that originated that state, Fukami 2015) becomes indistinguishable from stochasticity; and an apparent deterministic pattern, such as some species more likely to occupy a location, can occur by purely stochastic mechanisms, as in some neutral models (Hubbell 2001a, Tucker et al. 2016). Other approaches to measure the degree of determinism or stochasticity in a system rely on the definition of a reference system (e.g., abundances in the species pool), and a measure of the discrepancy between an observed sample to the reference system: when the discrepancy is large enough it can be interpreted as proof of a deterministic process (Vellend et al. 2014). A potential problem in this approach can occur when abundances in the reference system are affected by deterministic mechanisms that create abundances similar to those observed in the sample. The problem arises because the two deterministic patterns can not be distinguished and the comparison between will suggest that stochastic mechanisms prevail.

1.4 Plants as engineers and the mass-ratio-hypothesis: is there a correlation between the stochasticity experienced by a species and its dominance?

The lack of congruence in the evidence related to the stochasticity or deterministic nature of community assembly is concordant with Lawton's (1999) assertion that there are no common patterns in ecology. However, I argue that for plants, combining two widely observed patterns can shed some light on the stochasticity-determinism discussion: the wide differences in

abundance among species, and that species, particularly plants, do not only compete but also transform the environment.

One of the oldest patterns observed in nature is that few species are very abundant and most are not (Fisher et al. 1943, MacArthur 1960, McGill et al. 2007). Several terms have been used to describe this difference, such as dominant-subordinate-subdominant, or core-satellite framework (Hanski 1982) and the implications of that difference have been largely discussed (Brown 1984, Magurran and Henderson 2003, Cornwell and Ackerly 2010, Mariotte 2014).

Similarly, it has largely been recognized that species transform their environment (Scott-Phillips et al. 2013), and that some species have a larger impact than others (Paine 1969). In a broad sense, every species transforms their environment, but some of them, noticeably plants, have been widely recognized in their role of habitat modification (Huston 1994, Jones et al. 1997, Odling-Smee et al. 2013). That generalization has been widely discussed in the case of plants and it is called the mass-ratio-hypothesis, which states that each species contributes to any ecosystem process or habitat transformation proportionally to its biomass (Grime 1998), and therefore to its abundance.

I present here three differences between dominant and non-dominant plants and how those differences can cause differences in the degree of determinism of each of those groups of plants. These differences are not exhaustive and not all of them may apply to every system, but exemplify why we should expect an asymmetry between dominant and non-dominant plants. First, less abundant species may be subject to more frequent stochastic extinctions than dominant species (MacArthur 1955). The rationale is that given a mean population size and a fluctuation in time described as a standard deviation of the population size, smaller populations will be more likely to cross the zero abundance line and become locally extinct. Second, because all species alter their environment (Jones et al. 1997) and that transformation of the environment is often proportional to the species biomass (Grime 1998), the non-dominant species should be more affected by the dominant species than by the environment. In contrast, dominant species should be relatively more affected by environmental conditions than by the interaction with other species. Because habitat filtering causes phylogenetic underdispersion and limiting similarity overdispersion, this hypothetical asymmetry between dominant and non-dominant species is testable in natural conditions (Webb 2000, Mayfield and Levine 2010, Gerhold et al. 2015). And

7

third, if by occupying space or by modifying the environment, dominant species reduce niche differences as well as fitness differences among non-dominant species, the dominant species will be increasing the role of stochastic mechanisms for non-dominant species (Chesson 2000).

The first and third differences between dominant and non-dominant plants previously mentioned indicates that stochastic mechanisms can be more important for non-dominant species than for dominant ones (Chase 2007, Tucker et al. 2016), but provide no predictions about the type of the deterministic mechanisms. The second difference distinguishes between stochastic mechanisms and the two opposing deterministic mechanisms previously mentioned, habitat filtering and limiting similarity (Webb 2000, Cadotte et al. 2013, Gerhold et al. 2015).

1.5 Thesis statement

Whether communities are deterministically organized or not is an apparently simple question, but answering it has been plagued with difficulties and complexities. In the wake of large global changes, is it worthwhile to invest in a major effort to answer this question? I think two conditions have to be true to positively answer that question: (I) biodiversity is meaningful to predict ecosystem services and (II) our lack of understanding on the degree of determinism in natural communities creates similar uncertainty on how much biodiversity could survive than drivers of global change. Further, I hypothesize that the dominance of a species and the degree of determinism it faces in the community are correlated. Because ecosystem services are also expected to be related to the biomass of the species, the degree of determinism in the community will also impact the predictability of the provision of ecosystem services.

The second chapter of my thesis aims to answer condition II: In order to assess the total number of species in a fragmented habitat, how much uncertainty is caused by our lack of knowledge of the degree of stochasticity-determinism in the predictability of species survival, compared to uncertainty caused by dispersal capability of the species, the effect of land use change and our ignorance about future climatic conditions? I developed a conceptual model base on common biogeographic patterns and used it to answer that question at a landscape level in the Tropical Andes.

The third and fourth chapters explore the dominance-determinism correlation. The third chapter tests the assumption that stochasticity and determinism are different for dominant and non-

dominant plants. This was done removing dominant plants in a meadow and studying the response of the non-dominant plants that became the new dominant species in the experimental plots. In particular, this chapter asks (1) if dominant species (in comparison to non-dominant species) are deterministic or stochastically organized, (2) if limiting similarity or habitat filtering are important drivers to explain phylogenetic dispersion of dominant and non-dominant plants, and (3) whether neutral mechanisms can explain the observed changes in species abundances.

The fourth chapter aims to scale-up some of the results observed in the third chapter to meadows globally distributed. In particular, I aimed to answer (1) if phylogenetic dispersion of dominant plants is lower than of non-dominant plants, as a sign of stronger habitat filtering acting on dominant species and stronger limiting similarity acting on the non-dominant species; and (2) what can drive that difference.

The fifth chapter builds on the findings of the previous chapters to explore condition I: In meadows around the world, will diversity help to predict biomass productivity and litter accumulation? In this chapter I hypothesized that graminoids (grasses and sedges), forbs and legumes relate differently to productivity and litter accumulation because, as shown in the previous chapter, they may follow different assembly rules. If this is the case, modeling the biodiversity-productivity relationships independently for each group should improve our predictions.

The sixth chapter summarizes the findings of the previous four chapters, and present potential future studies that can help either to confirm or expand those findings.

Chapters 2 to 5 have been developed as independent studies to be published in peer-reviewed journals, and one of them is already published. The independent studies are:

- Arnillas, C. A., C. Tovar, M. W. Cadotte, and W. Buytaert. 2017. From patches to richness: assessing the potential impact of landscape transformation on biodiversity. Ecosphere 8(11). DOI: 10.1002/ecs2.2004
- Arnillas, C. A., and M. W. Cadotte. (in preparation for *Ecology*). Dominant plants in a meadow are more deterministically assorted than non-dominant plants.

- Arnillas, C. A., M. W. Cadotte, and The Nutrient Network[†]. (in preparation for *Ecology Letters*). Globally, grassland species are deterministically phylogenetically over and under dispersed, and species' phylogeny and dominance controls the difference.
- Arnillas, C. A., M. W. Cadotte, J. Firn, K. J. La Pierre, and The Nutrient Network[†] (in preparation for *Ecology*). Can functional groups and scale dependent causality solve the productivity-biodiversity conundrum?

[†] For these studies, I will follow co-authorship guidelines of The Nutrient Network (<u>http://www.nutnet.org</u>).

Chapter 2 From patches to richness: Assessing the potential impact of landscape transformation on biodiversity

2.1 Abstract

Natural patchiness and human fragmentation results in habitats that are not continuously distributed. How spatial configuration of patches in fragmented habitats influence biodiversity remains largely controversial. Here, we propose a framework to extend the species-area relationship approach to analyze how changes in habitat configuration affect species richness in fragmented habitats. We use hypothetical communities that are characterized by (1) their tolerance to human activities, (2) the dispersal capability of the individuals of any species, (3) the species-area relationship and (4) the species turnover among patches. Further, the species turnover is a function of (4a) the predictability of species survival and (4b) the species recolonization odds.

In our framework, we identify three extreme communities that encompass the richness of all potential different communities, and thus encapsulate the richness of real communities. We propose a graph to visualize the effect of different patch sizes on species richness, an index to quantify those changes and a second graph using the index to visualize the effect of distance between patches on species richness.

After applying our framework and tools to the Tropical Andes we found strong differences in the impact of natural vs. human-driven fragmentation on richness between biomes. When projecting future richness values under climate change scenarios, the largest source of uncertainty in our richness calculation (>90%) were species turnover among patches and species dispersal for most of the biomes rather than future climate or species tolerance to human activities. Habitat loss consistently decreased the species richness, however, fragmentation per se often increased it. The increment was mostly linked to the species turnover rate among patches.

Our framework is a new theoretical tool to study the main patterns that underlie regional richness and, therefore, can provide new insights to face spatial habitat reconfiguration caused by human activities. **Keywords**: biodiversity; climate change; fragmentation; land use change; species-area relationship; richness; Tropical Andes.

2.2 Introduction

The impact of habitat loss and fragmentation on biodiversity is a topic that has garnered much attention the last couple of decades (e.g. Tilman et al. 1994, Loehle and Li 1996, Fahrig 2003, Rybicki and Hanski 2013). Habitat loss is known to be one of the most important causes of species extinction (Diamond et al. 1989), but the effect of fragmentation *per se* (the disruption of previously contiguous habitats without major habitat loss, hereafter fragmentation) and increasing habitat isolation (e.g. Fahrig 2003) on richness is still a topic of debate (Simberloff 2000, Fahrig 2013, Rybicki and Hanski 2013). For instance, although most researchers expect richness to decline with increasing fragmentation, Fahrig (2017) found that a positive or negligible impact was actually a more common pattern.

Concerns about the effect of habitat loss and fragmentation on biodiversity is largely related to the disproportionate impact of human activities on nature, and how land use and global warming are affecting the size and spatial configuration of patches within habitats (Parmesan 2006, Thuiller et al. 2008, Tylianakis et al. 2008). Often unrecognized is the fact that topography, climate and geographic barriers produce naturally patchy habitats with potentially strong effects on species diversity patterns (Bertuzzo et al. 2016). Regardless of the origin of the patches, the large number of interacting processes that links changes in spatial configuration with changes in species diversity limits our ability to adequately predict diversity responses to fragmentation. Perhaps more importantly, we currently lack the necessary information needed to identify which of these processes is more important for improving our ability to predict changes in diversity.

Several fragmentation metrics such as total habitat area, edge density or fractal related metrics have been used to measure the impact of habitat loss and fragmentation on biodiversity (McGarigal and Marks 1995), and experiments, often using microcosm systems, have been used to distinguish the relative role of each in natural conditions (Gonzalez 2000). But despite the intuitive relevance of these metrics for biodiversity and their statistical links to species richness, most of them lack a causal underpinning that can provide strong quantitative predictions. For example, most efforts to link fragmentation and habitat loss to biodiversity have focused on focal species (Saura and Pascual-Hortal 2007, Watts et al. 2010) without incorporating important

biological processes such as species interactions and dispersal restrictions (Kissling et al. 2011, Wisz et al. 2013, but see Leroux et al. 2017).

The species-area relationship (SAR) is one of the few widely accepted correlations between biodiversity and landscape patterns (Rosenzweig 1995, Lewis 2006, Whittaker and Triantis 2012), with several models proposing mechanisms underlying it (Hubbell 2001a, Rybicki and Hanski 2013). In particular, the SAR can convert a spatial measure (i.e. area) into a measure of biodiversity for a habitat, i.e. species richness (e.g. Simberloff and Abele 1976, Lomolino 2000, Triantis et al. 2012).

Whittaker and Triantis (2012) distinguished two main types of SAR: The species accumulation curve (SAC), also known as continental-SAR, is the result of expanding a sample area in an otherwise continuous habitat where the increase in richness is the result of sampling a larger proportion of the population. The SAC curves are tri-phasic, with steeper slopes at small and large spatial scales (Rosenzweig 1995, Hubbell 2001a). The second type is the island-SAR (ISAR), in which the number of species is counted in fragments of a discontinuous habitat. In this case the SAR is defined by the species distribution in the landscape. In the ISAR model, more species can be observed in larger islands because the latter have larger areas that 1) can accommodate more individuals or 2) have more spatial heterogeneity (MacArthur and Wilson 1967, Hubbell 2001a). Rybicki and Hanski (2013) identified a third type of SAR, the one-fragment-SAR (OF-SAR), which is similar to an ISAR but assumes that a fragment is isolated of the others because migration is not possible (see Appendix A2 for a more detailed discussion about ISAR). The most recent review on ISARs supports a convex upward curve without an asymptote, which can be described with a power-law function, is the most commonly observed ISAR in nature (Matthews et al. 2016).

We develop a framework to assess regional species changes caused by natural and artificial fragmentation, habitat loss, and other landscape changes. To do that, we used the ISAR because it represents a basic biological intuition tested in the field (Simberloff and Abele 1976, Matthews et al. 2016) and supported by some simple ecological models (Hubbell 2001a, Rybicki and Hanski 2013): we expect more species in a larger isolated or semi-isolated space for a given community (see Appendix A2). As usually defined in SAR models, we assume that the species of a taxonomic group, guild or trophic level living in a habitat constitute a community, as usually

13

defined in SAR models, and share two population level properties: their dispersal capability and their tolerance to human activities, and that two community level properties are fixed: the ISAR and species turnover among patches. The turnover is the degree at which species found on smaller fragments are not found on larger fragments (Baselga 2010), and is in turn affected by the impact of dispersal constraints and the predictability in extinctions and survivals caused by species interactions (Lomolino 1996). We define a Levins-McArthur community (hereafter LEM-community) as a hypothetical community with a particular combination of these four properties. The species in a particular LEM-community will perceive the fragments as one or more patches according to the individual organism dispersal capability. With these concepts, we (1) develop a conceptual framework to explore how LEM-communities and habitat configuration affect total habitat species richness and (2) develop two graphical approaches and an index to describe the impact of changes in habitat configuration on the richness of some particularly extreme LEM-communities. Finally, we apply these tools to describe the impact of fragmentation and habitat loss on species richness in the Tropical Andes and to discuss the main sources of uncertainty in our richness calculation for this biodiversity hotspot.

As other montane areas, the Tropical Andes are a highly biodiverse area important for human well-being (Körner and Ohsawa 2005) with a rough topography that makes their habitats naturally patchy. These characteristics (Dirnböck et al. 2003, 2011) make the Andes and other mountain areas particularly vulnerable to the synergistic interactions of habitat loss, fragmentation and climate change (Travis 2003). Their biological and social relevance and fragility make this area a natural area to apply this research.

2.3 The Conceptual Framework: Assessing the Richness of a Community in a Patchy Environment

We designed the framework to evaluate how landscape configuration influences (long-term) equilibrium species richness of a community within a habitat, accounting for its patchiness. To that end, we defined a LEM-community as a set of species that occupy only one habitat and respond similarly to human-induced habitat transformation, disperse similarly (individual dispersal and recolonization capability) and have consistent interactions, with either predictable or unpredictable survival across the patches (Figure 8). With this information, we estimated (1)

the available habitat and (2) the patches as perceived by a species of a LEM-community. Together, (1) and (2) describe habitat configuration; then, we (3) assessed the number of species per patch and (4) describe the species turnover rate among patches. In the next sections, we will use these results to develop a graphic representation and an equation to assess species richness changes caused by spatial configuration changes.



Figure 1 Key characteristics that influence the number of species that can survive in a

fragmented habitat.

The landscape is characterized by the natural distribution of the habitats and the human activities on it. Species of a habitat will use the habitat according to their tolerance to human activities and their dispersal capability, thus the patches represent the way the species perceive the habitat. Each patch can be composed of non-contiguous fragments, as far as the movement of individuals among them is so frequent that keeps their population synchronized preventing a meta-population dynamic (see detailed explanation in the conceptual framework in the main text). If all the species in that habitat share these characteristics, the species-area relationship (f) can estimate the richness in each patch base on its area (A_i). The total number of species in the habitat (S, total richness) will depend on the patch richness and on the species turnover among patches, represented as the proportion of species present in a patch and absent in any larger patch (p_i). The species survival caused by the interaction among species.

(1) Available habitat: The habitat is a set of fragments of the landscape that share similar environmental conditions, where the LEM-community can exist (i.e. biomes, sensu Olson et al. 2001). Each fragment is a continuous area of the habitat with a discrete boundary (e.g. Figure S1). If the community is composed of species intolerant to human-induced habitat loss, only "remnant habitats" are available for the community. Conversely, species tolerant to anthropogenic transformation of the habitat will occupy remnant and altered areas ("potential").

habitats"). Once the available habitat is defined, we can calculate fragment sizes and interfragment distances for a given habitat.

(2) *Patch definition and habitat configuration*: If the individual dispersal capability of species is high enough, species will perceive two (or more) fragments of the habitat as a single patch. This will happen when individuals can easily cross from one fragment to another, allowing individuals to use resources from different fragments. If dispersal is fully unrestricted, the whole habitat operates as a single patch, and the community responds as a "*unitary community*". A unitary community is the first of three cases of extreme LEM-communities. Also, by decreasing the dispersal capability of species, the species perception of the number, size and distance between patches will change (Figure 9).



Figure 2 Effect of the maximum individual dispersal in the habitat configuration,

represented by the number of patches and the proportion of the total habitat in the largest

patch.

Several fragments can act as a single patch if individual organisms can disperse more than the distance between those fragments. In (a) the two largest fragments are close to each other, causing a sharp increase in the proportion of the largest patch with the increase in the maximum individual dispersal distance. However, the small fragments are spread around, causing a smooth decrease in the number of patches. In (b) the smallest fragments coalescence quickly in two large patches creating a sharp decrease in the number of patches, while the proportion of the largest patch increases once the maximum individual dispersal equals the distance between these two groups of fragments.

(3) *Number of species in a patch*: The number of species in a patch is ultimately controlled by species interactions as well as resources availability and variability. We use the island-species-area relationship (ISAR) to estimate the expected number of species in each patch of the habitat. The ISAR is a monotonically increasing function, with a decreasing or constant slope. Once fragments below the dispersal threshold have been grouped into patches, we assumed that the ISAR of those patches become a good approximation of the one-fragment-species-area-relationship, which is defined as the ISAR for 'completely isolated populations' (Rybicki and Hanski 2013).

(4) Species turnover among patches: Turnover refers to the degree of species dissimilarity among different patches (Baselga 2010). High species turnover indicates that patches tend to have different sets of species. Species turnover is controlled by (a) the consistency of a species survival when interacting with other species (predictability of species survival), and (b) population dispersal, which in turns controls the probability of re-colonization. The population dispersal can be larger than the individual dispersal, since a population can spread through the habitat patches in different generations. As previously discussed, the extreme "unitary *community*" has species with unrestricted individual dispersal which implies that only one patch exists, and therefore turnover is meaningless (Figure S2c). In any other case, more than one patch exist, and meta-populations dynamics may appear (sensu Levins 1969). Metapopulation dynamics imply that species can become locally extinct in some patches, and is strongly linked to turnover rate: When recolonization is impossible (no population dispersal) each patch could have a different set of species in the long term, creating a total turnover and a "disjoint community". However, if recolonization is possible but limited (restricted population dispersal), turnover will decrease when (a) species become extinct in a consistent order (species survival is predictable, i.e. species A will be the first to disappear, then B, then C, etc.) and (b) each species can recolonize every patch (Wright et al. 1997, Ewers and Didham 2006). Conversely, if turnover is negligible then the largest patch has every species of the habitat and smaller patches simply contain predictable species subsets, this scenario corresponds to a "nested community" (Patterson and Atmar 1986). In contrast, turnover increases when each patch can contribute some species to the species pool of the habitat. This can happen when: (a) species survival is stochastic (Hubbell 2001a), (b) because patch composition depends on the order species arrive to the patch

(priority effect sensu Drake 1991), or (c) some patches are more isolated (Aranda et al. 2013) and stochastic extinctions are not always rescued (Brown and Kodric-Brown 1977).

In this way, for a given dispersal level, nested and disjoint communities define, respectively, the minimum and maximum richness that can exist in a patchy habitat. As individual dispersal capability increases from zero to unrestricted (Figure 9), the number of patches is reduced and the largest patch size increases, causing both communities to converge into the unitary community. This implies that nested communities with zero dispersal will have the smallest largest patch possible (the largest fragment) and hence will have the minimum species richness for any nested community. Conversely, a disjoint community with zero dispersal rate would have the maximum number of species for a given spatial configuration of the habitat because the number of patches is maximal and each patch would contain a different set of species (Figure S2). These two community types, nested and disjoint with zero individual dispersal, along with the unitary community, define three extreme community types that provide the possible boundaries of richness values for the full range of individual dispersal capability.

2.3.1 Building the toolbox

Here we present two novel graphical approaches and a numerical index to study the richness of the nested, disjoint and unitary communities, and their potential change due to changes in habitat configuration. The first graph (patch-size profile) provides qualitative information about potential richness change with relatively few assumptions. Most notably, no assumption is made about the specific shape of the ISAR. However, the index assumes a particular ISAR function and provides numerical estimates of richness change. The index is then used to explore the effect of distances between patches with the second graph, the dispersal profile.

2.3.1.1 Patch-size profile: Visualizing habitat availability and inferring general richness patterns

To visualize the effects of habitat patchiness on richness we plot the cumulative patch area against the order of the patch for a single habitat type (sorted from the largest to the smallest) (Figure 10a-b). This curve, hereafter referred to as the 'patch-size profile', shows the largest patch size (LPS, y-axis value of left side of the curve), the total habitat area (THA, y-axis value of right side of the curve), the contribution of the patches of intermediate size (slope of the curve), and the number of patches (maximal x-axis value). We log-transformed the x-axis to

facilitate the visualization of changes in the number of patches. Profiles closer to the bottom and right side of the graph indicate less habitat area or a more fragmented habitat.



Figure 3 Construction and interpretation of a patch size profile.

(a) For a particular habitat of a landscape, patches of a potential habitat (red) that are used by the tolerant LEMcommunities, are identified and sorted from left to right from the largest (order 1) to the smallest one. The same process is applied to patches of the remnant habitat (blue), used by intolerant communities. (b) Then, the cumulative sum of the patches area is plotted (solid lines). Dashed lines define the two habitat configuration extremes for a defined number of patches and total habitat area: First, every patch has the same area (bottom line) and second, one very large patch and all the others as small as possible (above line). These configurations define the extreme richness scenarios. If species turnover is null, the above line will indicate the highest richness and the bottom one the lowest richness. If turnover is complete, the values switch, and now the above line will indicate the lowest richness and the bottom line the largest one. Differences in the largest patch size (d_1) , the total habitat area (d_2) and the number of patches (d_3) between profiles are shown. (c) Three different profiles as a result of different underlying mechanisms and how to interpret them in terms of the expected impact on richness for the different communities. In the left panel, habitat loss of the largest patch caused the near parallel profiles. In the middle panel, 10 largest patches were split in two (fragmentation of largest patches without habitat loss) caused the convergent profiles. And in the right panel, halving by two the second to tenth largest patches (habitat loss of medium-sized patches) caused the divergent profiles. The logarithmic scale in the x-axis helps to visualize the profiles of the largest patches, but reduces the visual effect of divergence caused by fragmentation without habitat loss.

In addition, assuming only a general form of the ISAR (where largest patches have more species than smaller ones and that the number of new species per unit of area is smaller when area increases), one can observe that: (1) the right side of the profile (THA) defines the number of species that can subsist in a unitary community; (2) the left side (LPS) defines the richness in a nested community; and further, (3) the more evenly split the habitat is into patches (patch-size profile closer to the bottom dashed line in Figure 3b) the higher the richness in a disjoint community with the same total habitat area and number of patches (Appendix 3).

Although it is not possible to calculate richness using the patch-size profile directly, we can visualize relative changes by comparing multiple profiles plotted together. For example, multiple profiles representing the effect of climate change, land use activities, or another landscape transformation can be contrasted. If remnant and potential patches are plotted together, the distance between the profiles shows the impact of land use change on species richness (d2 in Figure 3b). Profiles that are parallel, convergent or divergent are caused by different changes in the habitat configuration and imply different impacts on intolerant LEM-communities (Figure 3c).

When comparing current and projected future profiles, one possibility is that the future profiles can be lower than the current ones. Lower future profiles indicate lower habitat availability and a higher risk of species loss. On the other hand, future profiles could be higher than current ones, implying that there is an increase in available habitat. We discuss the implications of this scenario later on.

Additionally, the uncertainty of future climatic conditions can be taken into account by plotting a set of profiles for different scenarios but with the same climatic models (i.e. uncertainty in the concentration of gasses in the atmosphere), or using different climatic models for a same scenario (i.e. uncertainty caused by our level of understanding of the climatic system). Thus, a lack of congruence of the profiles implies uncertainty for projecting species richness. Further, multiple profiles can differ between the right, middle or left side of the graphs, representing stronger impacts on unitary, disjoint and nested communities, respectively.

20

2.3.1.2 Quantifying the potential impact of landscape changes on biodiversity

For a given LEM-community described by its habitat, tolerance to human activities level and maximum individual dispersal capability, the habitat configuration can be described by the number of patches (k), and the area of each patch A_i , with i = 1 being the largest and i = k the smallest patch. Then, habitat richness (S) can be described by

$$S = \sum_{i=1}^{k} p_i \cdot S_i = \sum_{i=1}^{k} p_i \cdot cA_i^z$$
 eq. 1

where S_i is the number of species in patch *i*, and it is linked to the patch area by the Arrhenius SAR (cA_i^z , where the variables *c* and *z* of the Arrhenius expression remain constant (Ewers and Didham 2006)); and p_i represents the species turnover as the proportion of species in the patch *i* that are absent in any larger patch.

The species turnover rate can be easily described for the extreme communities previously presented. In a nested community, every species is present in the largest patch, so $p_1=1$ and $p_{2..k}=0$, that implies that its richness is $S_n=cA_1^z$ where A_1 is the largest patch size (LPS). In a disjoint community $p_i=1$ for every patch because each species in each patch is unique, and its richness $S_d=\sum S_i=c\sum A_i^z$. Finally, there is only one patch in a unitary community, such that k=1 and $p_1=1$, A_1 is the total area of the habitat (THA) and $S_u=c\sum A_1^z=cTHA^z$.

To quantify the impact of different habitat configurations on the richness of our LEMcommunities we define

$$R = \frac{\sum_{i=1}^{k} p_i \cdot cA_i^z}{c\left(\sum_{i=1}^{k} A_i\right)^z} = \frac{\sum_{i=1}^{k} p_i \cdot A_i^z}{THA_T^z}$$
 eq. 2

which is the ratio of the richness of any LEM-community to the richness of the unitary tolerant LEM-community at a given time. The selection of the unitary community as a reference is rather arbitrary, but by using it the index will show that (1) land use change reduces habitat richness of nested and unitary communities (which is intuitive) and (2) species turnover is meaningless when there is only one patch. When the reference community is the unitary one, R-values will range from that for nested community with null dispersal (which is always < 1 when more than
one fragment exist because z < 1) to the R-value of the disjoint community with null dispersal too (which is always > 1 when more than one fragment exist also because z < 1), and unitary communities will always have an R-value equal to 1. Any other community that can support the same number of species as the unitary community will also has an R-value of 1.

R-values can either be used directly or combined to assess the impact of different aspects of the landscape configuration (Figure 4). For instance, using the remnant habitat to define the patches in the numerator and assuming again unrestricted dispersal, the R-values will assess the impact of human activities on the richness of intolerant unitary communities. Further, by plotting the Rvalues for nested and disjoint communities with different maximum individual dispersal we can explore the effect of connectivity (distances between patches) in the richness of the habitat. We call this graph the connectivity profile (Figure 4a-b). The equation used to calculate the connectivity profile (eq.2) allow for a direct interpretation of the proportion of species of different LEM-communities (Figure 4c), such as the proportion of species with unrestricted dispersal and tolerant to human activities able to survive due to current land use change. Since the R-values are on a normalized scale, the values can be compared among habitats as the proportion of species of a unitary tolerant community able to survive in a particular LEMcommunity. Therefore, R-values can be subtracted between them to obtain Δ R-values that can isolate the impact or uncertainty related to a particular component (Figure 4d). For instance, the uncertainty related to natural fragmentation is caused by different species turnover rates, and can be measured as the difference in the R-values of the extreme disjoint and nested tolerant community.



Maximum individual dispersal distance

b) Variables definition

- z Exponent of the Arrhenius SAR expression
- n(ITI) Number of patches as perceived by a tolerant/non-tolerant community with a given dispersal capability.
- $LPS_{(T/l)}$ Largest patch size as perceived by a tolerant/non-tolerant community with a given dispersal capability. $LPS_{(T/l)}=A_{(T/l)}$,
- $THA_{(T/I)}$ Total habitat area as perceived by a tolerant/non-tolerant community. $THA_{(T/I)} = \sum_{i=1}^{n_{T}} A_{(T/I)i}$

c) Basic expresions represented in the connectivity profile

Symbole and	-				
formulas	Tolerant to humans	Tolerant to Dispersal humans capability		Interpretation	
$R_{Tu} = \frac{THA_T^z}{THA_T^z} = 1$	Yes	Unrestricted	N/A	Reference level	
$R_{Iu} = \frac{THA_I^z}{THA_T^z}$	No	Unrestricted	N/A	Number of species	
$R_{Tn} = \frac{LPS_T^z}{THA_T^z}$	Yes	Restricted	Null	able to survive	
$R_{In} = \frac{LPS_I^z}{THA_T^z}$	No	Restricted	Null	community expresed as	
$R_{Td} = \frac{\sum_{i=1}^{n_T} A_{Ti}^z}{THA_T^z}$	Yes	Restricted/ Null	Total	a proportion of species in a unitary	
$R_{Id} = \frac{\sum_{i=1}^{n_I} A_{Ii}^z}{THA_T^z}$	No	Restricted/ Null	Total	community.	

d) Derived equations represented in the profile

Symbols and formulas	Interpretation
$\Delta R_u = R_{Iu} - R_{Tu}$	Expected effect of human activities on species with unrestricted individual dispersal (unitary communities).
$\Delta R_d = R_{Id} - R_{Td}$	Expected effect of human activities on species with total turnover (disjoint communities).
$\Delta R_n = R_{In} - R_{Tn}$	Expected effect of human activities on species with no turnover (nested communities).
$\Delta R_T = R_{Td} - R_{Tn}$	Uncertainty in species richness caused by natural fragmentation.
$\Delta R_I = R_{Id} - R_{In}$	Uncertainty in species richness caused by natural and human-led fragmentation and habitat loss

Figure 4 Connectivity profile using R-values and their relationship with patch metrics often

used in fragmentation analysis.

a) Expected connectivity profile for disjoint and nested communities of tolerant and intolerant communities. Every basic equation used to estimate an R-value profile represented in the figure is a particular case of $R = \sum p^{i^{-1}} I^{A_i^z}/THA_T^z$, symbols as in (b), and equations detailed in (c). In particular, nested communities (R_n) are represented by p=0 (null species turnover rates among patches), while disjoint communities (R_d) by p=1 (total species turnover rates), intermediate values of p generate curves intermediate to the ones presented in (a). As maximum individual dispersal capability increases, the habitat configuration changes (e.g. number of patches decreases and proportion of largest patch increases) and both R-values for nested and disjoint communities converge in the unitary community (R_u). Human activities (dashed lines) can only reduce the area occupied by the habitat, reducing the species richness for nested and unitary communities (lower curves for R_{Iu} and R_{In} than for R_{Tu} and R_{Tn} , respectively). On the other

hand, when compared with the richness of natural fragmentation (R_{Td}), human-induced fragmentation can increase richness of a disjoint community (R_{Id}), while habitat loss will decrease it. Some equations of R-values (c) and Δ Rvalues (d) provide examples of how to link the figure, LEM-communities, common patch metrics (i.e. largest patch size, LPS, and total habitat area, THA) and their interpretation.

To study the effect of intermediate species turnover rates, we assumed that the proportion of new species will decrease with any new patch, so that $p_i=p^{i-1}$. Here, *p* represents a nested-turnover gradient, with *p*=0 being a fully nested community, and *p*=1 a disjoint community (full species turnover). It is important to highlight that even when *p*=0.9, it does not imply each patch will add 90% of new species. For instance, if *p*=0.9 the patch *i*=10 will have less than 40% of new species. Other functions can represent the proportion of new species in each patch, but we believe it is more likely to expect a decreasing function when new patches are observed (regardless the order in which the patches are added) because the probability of finding a new species should decrease as more samples of the population are observed.

Importantly, the R- and Δ R-values should be interpreted as potential impact assessments rather than probabilities of extinction. This is important because they ignore important ecological processes as adaptation, and, particularly, the time needed to achieve the new equilibrium.

2.3.2 Applying the conceptual framework

We applied the framework to the biomes of the Tropical Andes, a biodiversity hotspot (Myers et al. 2000) where human land use has been prevalent for centuries (Ellenberg 1979), and has impacted different habitats with different intensities. Apart from human transformation of the landscape, climate change will likely cause an upward displacement of biomes, further reducing habitat area and increasing fragmentation (Malcolm et al. 2006). We use our framework to analyze: 1) how current spatial configuration shapes the species richness of the Andean biomes, 2) the main sources of uncertainty in calculating habitat species richness when habitat loss and climate change are accounted for and 3) how changes in habitat configuration in the future would affect different extreme LEM-communities.

We used the patch-size profile to visualize and qualitatively estimate the impact of land use and climate change on richness. Numeric estimates of these impacts per biome and measures of the uncertainty on calculating changes in richness (R), were obtained using R-values for several LEM-communities assuming different maximum individual dispersal levels and different species

turnover rates. Dispersal levels were selected based on the minimum and maximum distance among fragments in the habitat.

2.4 Methods

2.4.1 Spatial database

The study area (Figure S3) covers the tropical Andean region from Venezuela to Bolivia above approximately 600 and up to 6000 m a.s.l. This region encompasses 8 biomes (sensu Tovar et al. 2013): glaciers and cryoturbated areas (hereafter cryoturbated areas), two forested biomes that are the evergreen montane forest (hereafter evergreen forest) and the seasonally dry tropical montane forest (hereafter seasonal forest), three tropical alpine grassland biomes that are the paramo, humid puna and xeric puna, and finally two lower altitude biomes that are the prepuna and montane shrublands. Here we used the biome maps produced by Tovar et al. (2013) at 1km² resolution for the present and the future climate change scenarios. The map for the present represents a baseline from the year 2000 (current scenario) and we assume that their distribution is representative of the 1960-2000 climate conditions. Future biome maps represent projections using future projected climate as described by WCRP CMIP3 multi-model dataset (Meehl et al. 2007), for two future periods (2010-2039 and 2040-2069), 8 global climate models (GCMs) for emission scenarios A2 and 10 for A1B (Table S1). Biomes surrounding the Andes were also modelled to capture the change in the lower boundary of the Andean biomes, but they are out of the scope of this study.

We defined potential biomes as the biomes projected by the model under present or future climatic conditions. Remnant biomes are defined as potential biomes excluding current human transformed areas (Tovar et al. 2013). Since we used only a map of current human transformed areas, the projected future patch sizes of remnants biomes are overestimated. Two pixels of the same habitat were assigned to the same patch if they shared an edge or a corner (*i.e.* the Queen contiguity rule). Patches of 1 pixel were discarded from the indices to reduce noise from misclassified pixels.

2.4.2 Assessing fragmentation impact on Andean biodiversity

We used potential and remnant patch-size profiles of current biomes to assess the impact of natural and human-induced fragmentation, respectively. To quantify the effect of natural and

human induced fragmentation and habitat loss on extreme community types we estimated Rvalues for the extreme LEM-communities (nested and disjoint communities with null dispersal, as well as unitary communities) for each biome as observed at the present (see formulas in Figure 4).

2.4.3 Uncertainty in habitat richness under present and future conditions

We used the R-values to get an initial estimate of the role of different sources of uncertainty (species turnover, dispersal capability, human impact and climate change). First, for present conditions we estimated R-values for different levels of species turnover (ranging p from 0 to 1), and for different dispersal capabilities (from no dispersal to unrestricted dispersal) for both, potential and remnant habitats and plotted connectivity profiles with them. Second, we estimated the extra uncertainty related to climate change with future R-values for each available combination of climatic scenario and model. With this future dataset, we fitted linear models were the R-values were the dependent variable, using as explanatory variables the climatic scenario, climatic model, tolerance to human activities, maximum individual dispersal and turnover rate. We also included the interaction between maximum individual dispersal and turnover rate because the predicted strong interaction between these two terms. We estimated the relative weight of the uncertainty related to each explanatory variable using the percentage of the ANOVA type I sum of squares of the R-values for each model. Because fully disjoint communities (p=1.0) created extremely high R-values, we discarded their values to prevent a biased result. Also, to control for different number of levels in each factor, we used ten maximum dispersal distances evenly distributed for each habitat. Then, we compared the mean sum of squares fitted by the model.

To calculate the ISAR, we used several *z* values between 0.1 and 1, but focused on z = 0.25 to describe our results. This value of the *z* exponent is frequently used and is theoretically justified (e.g. Thomas et al. 2004, Rybicki and Hanski 2013). Further empirical *z*-values converge towards that value when increasing the range in the patch size (Figure S4). The other values were used to test the sensitivity of the conclusions to that parameter. We also analyzed an extreme case when each new patch has the same probability of new species ($p_1 = 1$ and $p_{i>1} = p$).

2.4.4 Effect of spatial reconfiguration in extreme LEMcommunities per habitat

To analyze the effect of future spatial reconfiguration on species richness of extreme LEMcommunities we plotted patch-size profiles for future climatic conditions. We used scenario A1B, and included one profile per climatic model for future remnant and potential biomes.

All the analyzes were performed with R 3.3.1 (R Core Team 2016), and the packages raster (Hijmans et al. 2015), igraph (Csardi and Nepusz 2006) and ggplot2 (Wickham 2009).

2.5 Results

2.5.1 How does the current habitat configuration of Andean biomes impact their biodiversity?

Strong differences in natural fragmentation between biomes can be observed (potential patchsize profiles in Figure 5 and Figure S5, and Table 1). The high elevational grasslands xeric and humid puna have a large patch that covers most of their potential area (Figure 5c and Figure S5b, respectively), which imply large R-values for nested tolerant communities. In both habitats, the relatively low R-values of disjoint communities and the low difference between R-values of disjoint and nested (ΔR_T) extreme communities suggests low natural fragmentation and a low uncertainty in the species richness caused by natural fragmentation. Similarly, the patch-size profiles indicate that despite the large number of fragments, the small size of most of them may contribute very little to the richness value. Conversely, Glacier and cryoturbated areas and paramos have a small largest patch, several medium-sized patches and not as many small patches (Figure S5c and a, respectively). The R-values for nested tolerant communities in these two biomes is thus lower than for other biomes (0.61 and 0.85 respectively), while disjoint communities get intermediate values, caused by a balance between important medium-sized patches, but not as many small patches, as in other biomes.



Figure 5 Patch size profile for current potential (black lines) and remnant biomes (gray

lines), assuming zero individual dispersal.

a) Evergreen forest shows the highest human impact. Human impact considerably reduced the area of the largest patches (parallel and slightly divergent profiles, see Figure 3c) and increased fragmentation (more patches). b) Montane shrubland profile shows that the largest patches remained with similar size but medium size patches have reduced their size due to human impact (divergent profiles, see Figure 3c) and many small patches got lost. c) Xeric puna does not show a significant human impact. Triangles and dots show the threshold between patches smaller and larger than 10 km² and 1 km², respectively.

Table 1 R-values for current potential and remnant Andean biomes.

For each habitat, the table shows the R-values for nested, disjoint and unitary communities, either tolerant or intolerant to human activities. Intolerant communities occupy remnant patches only, while tolerant ones can occupy the entire potential habitat. Unitary communities use the whole habitat as a single patch. The area of the largest patch defines the R-values for nested communities. Disjoint communities occur when every patch has a different species subset. The R-values represent the proportion of species able to survive in each community type divided by the number of species able to survive in a unitary tolerant community. Patches of 1 km² were excluded from the analysis.

	Un	itary	Nes	sted	Disjoint		
Habitat	Tolerant	Intolerant	Tolerant	Intolerant	Tolerant	Intolerant	
Evergreen forest	1	0.874	0.902	0.547	100.6	221.1	
Glacial and cryoturbated areas	1	1.000	0.609	0.609	60.8	60.8	
Humid puna	1	0.979	0.985	0.946	41.9	58.5	
Montane shrublands	1	0.966	0.937	0.929	140.4	118.2	
Paramo	1	0.916	0.855	0.599	36.9	54.3	
Prepuna	1	0.992	0.898	0.892	22.3	24.9	
Seasonal forest	1	0.967	0.863	0.856	169.5	175.9	
Xeric puna	1	0.997	0.985	0.983	50.4	48.1	

The effect of current land use can be perceived in the switch to the bottom and right of the remnant patch size profiles in comparison to the potential patch size profile, in many habitats, but particularly in the evergreen forest (Figure 5a). The second most affected biome due to current land use is the paramo (Figure S5a). These two biomes showed a large difference in their largest patches between remnant and potential habitat (almost parallel profiles), suggesting a large impact of land use on the richness of unitary and nested communities. In fact, for evergreen forest the R-values (Table 1) suggest a decrease of ~12.6% of species in a unitary community (ΔR_u) , and ~35% for a fully nested community (ΔR_n) , once equilibrium condition of the new spatial configuration is achieved. On the other hand, the increase in the number of patches and the massive reduction in area of evergreen forest could also affect the disjoint communities (large ΔR_d). In the case of shrublands, human impact mostly affected medium size patches (divergent profiles) without a change in the largest patch size (Figure 5b). However, the observed reduction in the number of patches, leads to negligible, small and larger impacts on the richness of nested, unitary and disjoint communities, respectively ($\Delta R_n \approx 0$, medium ΔR_u , and high ΔR_d). Finally, habitat loss had little or no impact on the communities of some biomes such as xeric puna (Figure 5c), prepuna or cryoturbated areas (Figure S5e and c, $\Delta R_u \approx \Delta R_n \approx \Delta R_d \approx 0$).

2.5.2 Present and future habitats' richness uncertainty

For present conditions, the connectivity profiles using various levels of species turnover (Figure 6) suggest that the largest source of uncertainty in calculating R-values is related to the species turnover rate and to the individual dispersal capability. Given that there is not much difference between the curves of tolerant and intolerant communities, the effect of habitat loss is not as significant. R-values for species turnover rate (p) above 0.9 were extremely high, but even lower values caused high uncertainty on R-values on any maximum individual dispersal, as can be seen in each figure in the vertical distance between curves with different turnover rate.



Figure 6 Connectivity profiles for evergreen forest, montane shrubland and xeric puna at present time for different maximum individual dispersal capabilities (x-axis) and for both tolerant (top) and intolerant (bottom) communities.

Lines represent a gradient of species turnover rates (*p* values ranging from 0 to 1 in colours). The plots show a large source of uncertainty in the R-values related to species turnover rate, followed by maximum individual dispersal distance, while values representing species tolerant and intolerant to human activities (top and bottom respectively) are more similar among them when controlling for the other variables.

The uncertainty of future richness estimates, measured as variability on R-values, was also related mostly to species turnover (Table 2a). Although we discarded the most extreme turnover rate (p=1.0) to reduce biasing the results, we still found that the largest source of uncertainty, was related, first, to species turnover rate among patches and, second, to maximum individual dispersal. Those two components and their interaction term explained more than 90% of the uncertainty of the estimated richness of each habitat. When accounting for different degrees of freedom in each explanatory variable (Table 2b), species turnover still represented the largest source of uncertainty. In contrast to the present conditions, the second largest source of uncertainty differed among habitats: tolerance to human activities for evergreen forest, shrubland and seasonal forest; maximum individual dispersal for prepuna, paramo, humid and xeric puna and glacier and cryoturbated areas. Future climate, as a combination of uncertainty of climate scenario and model is a very large source of uncertainty for cryoturbated areas too. These results suggest different ranking in the main threats faced by each of those biomes.

Table 2 Main sources of uncertainty in the species richness calculations of the Andean biomes

during the period 2040-2069.

After fitting a fixed effects linear model with the R-values, uncertainty was estimated as (a) the percentage of the total variance (represented by the sum of squares) and (b) the mean sum of squares. The model described climate uncertainty by two greenhouse gas emission scenarios (A1B and A2) and several climate models (10 for A1B, 8 for A2), and the LEM-communities as tolerance to land use change (tolerant or intolerant), ten levels of maximum individual dispersal capability (MID, actual distances varied according to each biomes' spatial configuration) and ten different species turnover rates (p=0, 0.1, ..., 0.9). In (a) turnover rate explains by itself the largest proportion of the variance. Further, together, MID and turnover rate explained more than 90% of the total variance. In (b) meansum of squares figures suggest that predictability is the single largest source of uncertainty, and that the second most important source of variability is strongly related to habitat stressors, either land use or climate change.

Source	Evergreen Forest	Glacier and cryoturbated areas	Humid puna	Montane shrublands	Paramo	Prepuna	Seasonal forest	Xeric puna
Scenario	0.00	0.25	0.00	0.00	0.00	0.00	0.00	0.01
Model	0.10	3.60	1.45	0.83	0.30	0.97	0.34	1.73
Tolerance	1.85	0.00	0.02	1.26	0.37	0.02	0.79	0.01
MID	12.89	3.82	22.61	1.53	11.60	12.52	4.08	14.92
Turnover rate	46.20	69.49	33.55	87.96	59.30	56.25	81.09	51.35
MID * Turnover rate	35.29	21.80	39.26	6.04	26.91	26.03	12.37	28.91
Residuals	3.67	1.04	3.11	2.38	1.51	4.21	1.32	3.08

a) Proportion of the sum of squares by uncertainty sour

b) Mean sum of squares by uncertainty source

Source	Evergreen montane forest	Glacier and cryoturbated areas	Humid puna	Montane shrublands	Paramo	Prepuna	Seasonal montane forest	Xeric puna
Scenario	0.02	1.81	0.00	0.02	0.03	0.00	0.00	0.03
Model	0.06	2.85	0.33	1.30	0.29	0.35	0.48	0.58
Tolerance	10.64	0.00	0.05	17.91	3.25	0.06	9.97	0.02
MID	8.24	3.02	5.09	2.41	11.28	4.54	5.71	5.03
Turnover rate	29.53	54.96	7.55	138.82	57.69	20.40	113.42	17.32
MID * Turnover rate	2.51	1.92	0.98	1.06	2.91	1.05	1.92	1.08
Residuals	0.01	0.00	0.00	0.01	0.00	0.00	0.01	0.00

2.5.3 Effect of spatial reconfiguration in extreme LEMcommunities per habitat

We plotted the patch-size profile for future climate models under the A1B scenario to further explore the changes in the patch sizes and their impact on different LEM-communities in a more robust way (no assumption about the island-species-area-relationship, ISAR function). The projected patch-size profiles for the future (Figure 7 and Figure S6) showed three patterns of climate change impact: (1) biomes with significant displacement of their profiles to the lower-right corner (*i.e.* cryoturbated areas, evergreen forest and Paramo); (2) biomes that have a significant rise in their profiles, suggesting lower fragmentation and even greater habitat area (*i.e.* seasonal forest); and (3) biomes with uncertain futures, with some climate models projecting

an upward and others a downward curve displacement (*i.e.* xeric puna, montane shrubland, and prepuna). Humid puna habitat was largely projected to decline, but some future uncertainty remains – nine models projected habitat reduction and only one an increase.



Figure 7 Patch-size profile for projected future biomes under different climate models

(gray lines) for potential and remnant areas, assuming zero individual dispersal.

Black solid lines represent present profile. Evergreen forest (a) is the biome projected to be the most affected by climate change showing a projected reduction in total area while seasonal forest (b) is projected to have an increase in habitat size under future climate conditions. Both habitats have more uncertainty related to the total habitat area than to the largest patch size. Prepuna (c) shows a high uncertainty under future projected scenarios, with higher uncertainty in the largest patch size. Data represents A1B scenario for the period 2040-2069.

In some biomes, the future profiles diverged among them, indicating more uncertainty in the size of medium-sized patches that affect richness of disjoint and unitary communities. This is the case for potential evergreen forest, seasonal forest (Figure 7a-b), paramo and cryoturbated areas (Figure S6a-b), and implies that the future species survival of nested communities is less

uncertain (either positive or negatively) than for other communities in the habitat. Prepuna showed the opposite pattern (Figure 7c), where future profiles converged indicating more uncertainty in the richness of nested communities. Shrublands, xeric and humid puna had almost parallel profiles (Figure S6c-e), suggesting similar uncertainty levels for the richness of each LEM-community. Finally, human impact of the environment mostly constrained the potential future R-values, as can be seen on paramo, evergreen and seasonal forest. An unexpected pattern is observed in remnant evergreen forest (Figure 7a-remnant habitat), where more habitat is available for nested communities, but less for unitary ones (left and right hand of the profiles, respectively).

2.5.4 Sensitivity analysis

The effect of species turnover rate and dispersal capability on the uncertainty of species richness is robust for *z*-values in the range 0.1-0.5 (Figure S7a,b). This range of exponents corresponds to the observed range of *z*-values reported in natural systems in scales comparable to the areas studied here (Figure S4, and references in (Matthews et al. 2016)). For *z*-values closer to 1, the pattern changes and the uncertainty in each biome is mostly controlled by the source primarily driving total habitat modification. The ranking of the most important variable (either species turnover rate, dispersal capability and their interaction) depends on the function used to represent the change in the proportion of new species in each patch (Figure S7b).

2.6 Discussion

2.6.1 Describing Andean biomes richness

Our results show that the Andean biomes have different levels of fragmentation due to both, natural and human-driven factors that have impacted their richness. Cryoturbated areas and seasonal forest are strongly naturally patchy habitats (Armenteras et al. 2003), the first one having little evidence of human direct transformation. On the other hand, our results also show that land use change has negatively, and strongly, impacted richness in the paramo and the evergreen forest. These findings complement those of Tovar et al. (2013) whose analysis show a drastic reduction of habitat extent in both biomes.

Consistent with the findings of Fahrig (2017) we have also found that fragmentation per se is likely to cause an increase in richness, with a higher increase when turnover increases. This

should not be taken as a sign of a healthier environment though, because the model only refers to the number of species but does not predict the identity of the species, thus, the quality of the services or functions provided by the ecosystem based on the new species composition is uncertain (Sobral et al. 2016). In the equilibrium state, the number of species can increase from the migration and establishment of species from other habitats or from speciation (Chave et al. 2002), but immigration of non-native species will likely drive increases in the short term. For instance, a larger R-value caused by habitat loss can imply an increase in alien species, thanks to their plasticity, genetic variability, new climatic conditions variable enough as to encompass their original climatic constraints, among others (Hylander and Ehrlén 2013). However, the potential richness increase also provides some hope for native species survival, because fragmentation may be able to maintain the diversity that would otherwise be lost from habitat destruction. This seemingly contradictory result cannot occur in unitary or nested communities, which means that meta-population dynamics have to prevail and that extinction cannot be predictable. This mechanism is different from the one proposed to explain "extinction debt" (Hylander and Ehrlén 2013), because this higher richness level is caused by a long-term equilibrium according to the island-species-area relationship (ISAR) that underlies the model, and not a transient situation.

However, it is important to highlight that even though fragmented habitats with high turnover rates could support higher community richness than continuous habitats, the probability of randomly losing any single species in any patch is higher due to smaller population sizes and susceptibility to environmental stochasticity (MacArthur and Wilson 1967). Also, with each species present in less patches, the probability of recovery of any population from propagules from other populations is lower (Levins 1969). These two processes may threaten the stability of the system. Also, a larger proportion of alien species can increase the probability of one becoming invasive, and facilitating the replacement of local by other alien species as in an invasive meltdown (Simberloff and Holle 1999).

When we analyzed the effect of future climate change scenarios on species richness, overall, our results support previous projections of future species loss (Ibáñez et al. 2006, Feeley and Silman 2010, Dullinger et al. 2012) for several biomes (e.g. evergreen forest and paramo). However, our analyzes also show that there could be an increase in richness in others, particularly in the seasonal forest. Projections for this biome show all future patch-size profiles are above the

present profile for tolerant and intolerant species. These contrasting results match the conceptual model proposed by Young (2009), who suggests species adapted for dry and warm conditions might expand their distribution, while species better adapted for wet and cold conditions could show range contractions.

As expected, future climate change interacted with land use transformation. Extending the results obtained by Tovar et al. (2013), our future patch-size profiles indicate that biodiversity loss due to land use change will remain the major threat to biodiversity in evergreen forest, shrublands and humid puna. Interestingly, regardless the sign of the effect of climate change, human transformation of the habitats reduced the future species richness uncertainty in several habitats, a result consistent with the findings of Feeley et al. (2011) for montane tree species, who found a reduction in the variability of the species distribution when human activities were taken in account.

The unexpected result observed for evergreen forests where climate change would cause that nested communities increase in their equilibrium richness while unitary communities decrease is consistent with an upward displacement of the biome. This would be the case under the assumption that the connection between patches above the current tree line is allowed despite of human activities. Unfortunately, people in human settlements located in the higher grassland biomes (particularly humid puna) often burn trees to keep their traditional production system (Feeley et al. 2011), and, even now, upward seed dispersal is not as likely (Rehm and Feeley 2015). It is crucial to point out that we assumed that land use will remain constant despite future land use is expected to expand. If agriculture increases in areas similar to those where crops are currently more prevalent, as often assumed in land use change models (Verburg and Veldkamp 2001, Tovar et al. 2011), then future deforestation should reduce the total cover and fragment the evergreen forest more than any other habitat. Overall, the effects of changes in climate and land use represent a considerable threat to evergreen forest.

Generally, and despite the important challenge imposed to Andean ecosystems by land use and climate change, the R-values we obtained suggest that the largest source of uncertainty in calculating potential richness comes from our poor understanding of two biological properties of the communities: the predictability of the species survival caused by the interactions among species and the recolonization odds (Table 2). These two properties together control the species

turnover rate among patches (p_i values). Therefore, estimating empirical turnover patterns around the world can be as important as describing the ISAR itself (Brown and Kodric-Brown 1977, Baselga 2010, Matthews et al. 2016). A third property, the maximum individual dispersal, has a large impact on regional richness (Table 2; Figure S7), and is strongly related to recolonization. In this context, the largest efforts to address the risks related to climate change and habitat loss, should be focused on describing these interrelated properties. A better understanding of these properties can help to direct future research. For instance, if predictability of species survival and recolonization odds are high, the largest effort should be directed to characterize species traits and their impact on community assembly processes. Conversely, if either or both are low, more effort should be put to describe current species distribution.

2.6.2 General framework

The ISAR has received a lot of attention from ecologists (Pimm and Askins 1995, Thomas et al. 2004, Ewers and Didham 2006), and several mechanisms have been proposed to explain its shape (see for a review Whittaker and Triantis 2012). However, one of the most important limitations of its application is the lack of a clear framework to use ISARs in fragmented landscapes, an issue explicitly addressed in this paper. Here, we built on the species-area relationship to assess the impact of spatial configuration of a fragmented habitat on species richness of its community.

Our approach focused on the definition of hypothetical communities that we named Levins-McArthur communities (LEM-communities), in recognition to their contributions to spatial ecology and biogeography. Each LEM-community is conceptually associated with a habitat, and characterized by species with the same individual dispersal capability and tolerance to human impact. The LEM-community also has a particular ISAR and a species turnover rate among patches. Using the general shape of a ISAR we deduced that disjoint communities (i.e., complete species turnover among patches) and nested communities (i.e., the largest patch has every species) bracket the potential richness for any given dispersal capability, as can be seen in Figure 6. We also showed how the richness of these communities converge into a unitary community when the individual dispersal capability increases. Finally, we discussed how the species turnover rate among patches is a function of recolonization and of the predictability of the survival caused by interactions among species. Because species turnover is the largest source of

uncertainty (also shown by Matthews et al. 2016 for real datasets), understanding these two biological properties is fundamental to improve the predictability of richness changes.

In general, our approach is similar to surrogate species approaches, where species with certain characteristics (e.g., high dispersal or tolerant to human activities) can be used to represent broader communities (Watts et al. 2010). However, since the model works at the community level, community characteristics, such as species interactions, were implicitly included. Additionally, the model can be used in areas where species information is scarce. Some simple assumptions can be included to narrow the uncertainty related to the survival among species. For instance, the probability of an extinction could be inversely proportional to the overall abundance of a species, as in a neutral model (Hubbell 2001a), increasing the predictability of species survival among patches. Therefore, general patterns of species turnover and nestedness (Baselga 2010) could provide a strong basis to improve and validate these models with a minimum loss of generality. Another pattern to explore is the inclusion of explicit patterns of endemism (e.g. periphery, central area, particular areas), because where endemism is not as important, we can expect higher predictability in species survival among patches (more nestedness and less turnover). Also, nesting species distribution models into the projection of this habitat-oriented framework would open an interesting research agenda.

Some of the limitations of ISARs would still be applicable to our general framework. One important limitation is the assumption that the species from one habitat cannot exploit other habitats (Tjørve 2002, Proença and Pereira 2013). This assumption is also a limitation of our proposed approach. However, if the proportion of species coming from other environments remains relatively constant over time and for different human tolerance levels, the effect on the numerator and denominator of the R-values will be similar, and the R-values will be robust. Another limitation is that often the global distribution of the habitat is not included in the study area, and that could create a bias in the ISAR. In our particular study, this is not a major source of concern because the habitats were mostly fully contained in the area, but should be evaluated before extending the model to other regions.

Other limitations of the ISAR have been already incorporated into our framework, such as the need to incorporate species differences in the ISAR representation (Ewers and Didham 2006, Franzén et al. 2012) or the usage of different ISAR (Dengler 2009). We addressed the first step,

species perception of the environment, by changing the patch definition in a way consistent with the community, similar to the identification of an appropriate habitat scale proposed by Fahrig (2013), and not by changing the z exponent of the ISAR (as in Franzén et al. 2012). The second limitation is an important one, but only valid for the R-values because the patch-size profiles do not assume any particular ISAR shape.

A fundamental assumption of the R-index is that by using our extreme communities, composed of species that have ecologically similar individuals, we obtained extreme values of richness. Real communities, which are composed of species with mixed ecological strategies or characteristics, should have intermediate R-values. This assumption requires a more detailed analysis, but it seems consistent with the results of Kadmon and Allouche (2007) for landscapes composed of only one habitat, as it is the case in our framework, where the richness is assessed biome by biome.

Another fundamental assumption of the model is that the habitat is homogeneous among patches, so that the survival probability of a species does not depend on the patch identity. Directional turnover, when species are replaced in a predictable way along a gradient (Anderson et al. 2011) violates this assumption. A similar problem occurs if species richness is controlled by habitat heterogeneity in each patch instead of patch area (Freemark and Merriam 1986). In these two cases, the model can be used to test the underlying mechanisms by building a basic assumption of a nested pattern and exploring if the addition of extra environmental dimensions improves the prediction of the species richness, nestedness or turnover patterns.

Finally, the most important assumption of our framework is that ISAR is a community property, in other words, that the number of species relates to the area of the habitat such that the larger the area the more species will survive there, and that each new spatial unit will add the same or less species to the patch. As previously discussed, besides empirical evidence (Triantis et al. 2012, Matthews et al. 2016), some simulation models support this assumption, but may diverge in the equations that should be used to link the area and the species richness of a patch (see Appendix 2). Because this framework makes explicit the different sources of variation in species richness calculation for a given habitat it can help to test some of the fundamental models and theories about community assembly and their relative contribution in natural conditions.

2.7 Conclusion

Currently, species distribution models are the most common method to assess climate change impacts on biodiversity (e.g. Thomas et al. 2004, Feeley and Silman 2010). Despite evidence showing the relevance of biotic interactions even at large spatial scales, species distribution models usually lack species interactions and rely on species-specific characteristics, for which the necessary information can be scarce (Wisz et al. 2013, de Araújo et al. 2014). Our approach, grounded on community analyzes and how the species perceive their surroundings (habitat characteristics, individual dispersal capabilities, tolerance to human activities and recolonization odds) and on some community properties (species-area relationship and predictability of species survival), provides an alternative approach to estimate changes on potential richness under different scenarios of habitat configuration. This approach provides a new way to apply speciesarea relationships to patchy habitats, which are frequently present in nature and become more prevalent where human activities occur. It also highlights a strong difference of the effect of fragmentation and habitat loss on species richness contingent on community properties that should be further studied. Finally, our proposed framework adds a new tool to those used by managers to adapt management plans in the face of increasing land use, climate change and other threats that can reduce or fragment a habitat.

Acknowledgments

This study was part of the project "Vulnerability, Adaptation and Mitigation of Climate change effects in the Tropical Andes" implemented and funded by the Consortium for the Sustainable Development of the Andean Ecoregion (CONDESAN) and the General Secretariat of the Andean Community (SGCAN). In particular, we acknowledge the support provided by Francisco Cuesta. We also acknowledge the comments and ideas received by the members of the Cadotte Lab, at University of Toronto, Scarborough, and the support provided by the Cannaught scholarship at UTSC to CAA and the TD Professor of Urban Forest Conservation and Biology chair and Natural Sciences and Engineering Research Council of Canada (#386151) to MWC.

Chapter 3 Dominant plants in a meadow are more deterministically assorted than non-dominant plants

3.1 Abstract

Whether communities are assembled deterministically or stochastically has been a matter of debate in ecology for a long time. We argue that all species in a plant community are not governed by environmental constraints in the same way, and specifically that dominant species are influenced more by the environment than non-dominant species, but moderates the environment for non-dominant species, thus altering the role of deterministic mechanisms. Here, we removed two dominant species in temperate meadow plots and tested for differences in the mechanisms that organize plant communities. We assessed whether plots became more similar when we focus on the new dominant species that emerged after treatments were applied compared to the remaining non-dominant species. After removing dominant species, we found that non-dominant communities diverged over time while the new dominant species converged among plots over time. We found those trends despite plot level stochasticity and site-level changes in species abundances. Phylogenetic analyses suggest that the new dominant species were driven by habitat filtering and limiting similarity simultaneously. Conversely, nondominant species were more affected by limiting similarity. The larger role of limiting similarity for non-dominant species could imply that deterministic processes became less predictable due to, for instance, an increase on the importance of priority effects.

3.2 Introduction

The relative importance of deterministic and stochastic processes in determining community assembly mechanisms is a source of near constant debate in ecology (e.g. Hubbell 2001b, Tilman 2004, McGill et al. 2007). When deterministic processes prevail, traits that characterize a species (such as seed dormancy, height, frost tolerance, leaf thickness, nectar production, root structure, fungi associations) modify the survival odds of that species in its environment (MacArthur and Levins 1967, Chesson 2000, Tilman 2004). In contrast, stochastic mechanisms prevail when the number of seeds per species arriving to a place is more important than species characteristics. This dichotomy in mechanism is important because if local deterministic, or niche-based, processes are more important than stochastic ones, traits can provide predictive insights into

future community structure and composition under environnemental change. However, if stochastic processes prevail locally and regionally as in neutral models, more effort should be allocated to describing species distributions and dispersal patterns. Therefore, disentangling these processes is more relevant than ever as ecologists attempt to predict the potential impacts and consequences of global change on diversity and ecosystem services (Pearson and Dawson 2003, McGill et al. 2007, Buytaert et al. 2011, Laughlin et al. 2012).

Niche and neutral models predict an uneven species abundance distribution at local scales with few abundant species and a long tail of low abundance species (McGill et al. 2007), but they differ in the mechanisms resulting in dominance by few species. Observational studies have a notoriously difficult time disentangling these competing mechanisms and a dominance removal experiment can provide insights into the mechanisms structuring communities. Removing dominant species frees space that non-dominant species can occupy, but niche and neutral models differ on which species should occupy that space. In a neutral community propagules of dominant species are always more abundant than propagules of non-dominant ones. However, because the dominant species are constantly removed the abundance of the propagules that can actually survive (the non-dominant ones) is more even than if the dominant species were present. In turn, the more even distribution of propagules generates a larger effect of stochasticity added to the initial stochasticity among plots, reducing the similarity between plots (Figure 8a). In contrast, in a niche community the dominant removal causes a reorganization of competitive and facilitative networks among the remaining species, and therefore a new ranking is likely to occur (e.g. a previously outcompeted plant can became dominant). The new deterministic ranking increases the initial similarity between plots (Figure 8b).



Figure 8 Expected dynamic of several plots after dominant plants are removed (vertical dark gray bar) and some non-dominant species occupy the released space.

In a stochastic system the new dominants are likely to be those with higher density at the removal, increasing the dissimilarity of the system. In a deterministic system, the new dominants are likely to be those non-dominant species better suited for the new conditions, reducing the dissimilarity among the plots.

As with most theoretical studies on this topic, the predictions in Figure 8 assume that all the species in a community behave similarly and follow the same deterministic or random assembly mechanisms (e.g. Tilman 1980, 2004, Hubbell 2001b, Gilbert et al. 2009). However, dominant species confront environmental conditions more than non-dominants and thus dominant and non-dominant species likely face differing environmental constraints and might be subject to differing assembly mechanisms. For instance, dominant plants can reduce resources (e.g. shading, nutrients acquisition; Tilman 1980, Hubbell 2001b) but can also reduce environmental variability (e.g. decreased temperature or moisture variability; Brooker et al. 2008, McIntire and Fajardo 2014) (Figure 9). Thus, dominant species may be more deterministically assorted than non-dominants, or the other way around, independently of which species are the dominant ones in a given locality.



Figure 9 Dominant species interact more with the environment and with other dominant species (often intraspecific interaction) than with non-dominant species, while non-dominant species follow the opposite pattern.

Double-headed arrows represent biotic interactions, while single headed arrows represent interactions of species with the environment. Dominant plants also modify the environment under them by reducing some conditions (e.g. light, gray shade in the oscillating line at the left) and reducing the variability of others (e.g. wind, amplitude of the oscillation in the line at the left).

The convergence or divergence of plots over time is one way to identify the relative role of deterministic and stochastic processes. However, even if determinism is apparent because of increasing similarity over time, we cannot determine the mechanism responsible for this pattern from compositional patterns alone. However, these mechanisms create predictable community phylogenetic patterns based on the assumption that more closely related species are likely to be more similar (Webb et al. 2002, Weiher et al. 2011, Gerhold et al. 2015). With limiting similarity, strong interspecific interactions mean that the coexisting species will be more distantly related than expected by chance (overdispersion) because species compete more strongly when they are more similar to each other (MacArthur and Levins 1967). Strong positive interactions are also more frequent among distantly related species and can generate a similar overdispersed pattern (Valiente-Banuet and Verdú 2007). For habitat filtering, environmental constraints limit which species are able to thrive at a site, and those that do likely share some key traits or ecological strategies, increasing the relatedness of the coexisting species (clustered) (Mayfield and Levine 2010, Cadotte and Tucker 2017).

In this study we aimed to evaluate the difference between community assembly processes affecting dominant and non-dominant species. In particular, we want to test whether dominants are more deterministically sorted than non-dominant plants, and if deterministic, whether limiting similarity or habitat filtering prevails in each of these groups of plants. We evaluated the changes in community similarity after dominant species were removed over four years in five temperate meadows with varying species richness and composition, and explored the change in similarity among plots. At the end of the study period, we identified the dominant species in each plot and measured the changes in community similarity when only the new dominant species or only the remaining non-dominant species were used. We compared the dissimilarity among communities and compared these results with other treatments in the same sites to validate if the observed trends were caused by the effect of removing the dominant species, removing biomass or changing the species richness. We also compared the phylogenetic dispersion of each dominance group of species in each treatment to identify different structuring mechanisms. Finally, we explored the response of each species that became dominant in each site in the dominant removal treatments to see if they responded differently to different treatments or if they were just following site level trends.

3.3 Methods

3.3.1 Experimental design

In 2014, we identified and commenced the experiment in five old-field meadow sites at the Koffler Scientific Reserve, 40 Km North of Toronto. The Reserve was horse pasture more than 20 years ago, and the sites were located in fields in different areas of the reserve. In each site we visually identified areas with high species similarity and set 30 plots of 1x1m, leaving around 1m or more between them. By the end of the experiment, site species richness ranged between 13 and 54 species, and the dominant species were different combinations of three species: *Solidago altissima*, *Bromus inermis* and *Poa pratensis* (Table S2).

We measured cover and height of each species in each plot three times every year. On June 2014 we defined two dominant species per site as the species with largest coverage, height and frequency (proportion of plots where the species is present). At each site, we removed dominant species in five randomly selected plots (Dominant Removal plots, DR) and randomly matched them with another five plots. We randomly removed ramets in this second group of plots until we removed a similar amount of fresh biomass (Dominant Matching plots, DM) as in matching DR plots to account for the effect of dominant species from the removal of biomass. In another

five plots, we randomly selected two non-dominant species and removed them to account for the effect of changes in richness (Non-Dominant Removal plots, NDR). We also then matched these plots with five other plots and removed the same amount of biomass (Non-Dominant Matching plots, NDM). The other ten plots remained untreated as controls. The plants were removed mechanically cutting the stems or leaves as close to the ground as possible, to remove as much biomass without altering the soil structure. Removal treatments were reapplied at least twice a year after the plants measurements.

3.3.2 Phylogenetic information

We used the phylogeny of a nearby national park built by Jin (2015). Of the 61 species observed, seven species were identified at the genus level, so we replaced each of them with random congenerics in the tree for each one. Two species were not in Jin's tree, so we compare their location in Zanne et al. (2014) phylogeny and found a close match to each of them. We discarded species without a complete identification and non-tracheophytes (6 species in total).

All the analyses presented later were done with R 3.4.2 (R Core Team 2017). The phylogenetic analyses and data manipulation used the packages *picante* (Kembel et al. 2014).

3.3.3 Abundance and definition of new dominants

We estimated volume occupied by each species (height multiplied by cover) as a non-destructive proxy of species abundance. Because different species fill the space differently, the total volume cannot be used as a direct estimate of total biomass, but it provides an estimate of the tridimensional space occupied by each species. When field data was partially missing (e.g. height recorded, but not coverage, less than 1% of the records), we used other available information (e.g. pictures, count of ramets) and compared it with other records in the same site and treatment to fill the gap and prevent data loss.

We defined the dominant species at the end of the experiment in each plot as at least two species with the largest volume that occupied at least 50% of the total observed plot volume, discarding any species being removed from that plot. Frequency was one of the criteria to define pre-treatment dominance. However, high frequency is only a valid criterion if deterministic mechanisms prevail, so we did not include frequency as a criterion to define dominance at the end of the experiment. To identify the new dominant species we used the maximum volume of

each species in the final year in each plot. All together, we classified the species at each plot as: site dominants (pre-treatment criterion), new dominants (final year criterion) and non-dominants. It is important to highlight that the new dominant species in the DR treatment will be by definition different species than the site dominants. However, the new dominants and the site dominants should coincide in all the other treatments (DM removes a large amount of biomass, but because it is removed randomly the impact should reduce the abundance of all the species proportionally).

3.3.4 Statistical analysis

3.3.4.1 Treatment efficacy

To assess the efficacy of the treatments in each site, we estimated the average plot richness and total volume of species classified as initial site dominants in each treatment and site. We repeated those estimates for the new dominant species and for the non-dominant species to assess if treatments affected the community significantly.

3.3.4.2 Deterministic vs. stochastic trends

To assess whether the treatments increased the importance of deterministic or stochastic processes in the plant communities as a whole we measured the similarity among plots in each visit, treatment and site. Then, we used the pre-treatment conditions as a reference level and count the number of visits in which the similarity was higher or lower compared to the pre-treatment conditions. In a system that is not changing in time (deterministic and stochastic processes remain similar), the similarities should still oscillate because of small random fluctuations in measurement and local conditions, and in half of the cases the system will be more similar than before. However, if deterministic (or stochastic) processes increase, the proportion of times the plots are more similar should be higher (or lower) than 0.5.

We estimated the mean Bray-Curtis similarity (Magurran 2013) among plots in each visit, treatment and site using the *vegan* package (Oksanen et al. 2017). The similarity values range from one for identical communities to zero for communities with no species in common. Before measuring the similarity, we discarded the species removed in each plot in the DR and NDR treatments before and after the treatments started to compare the changes in the part of the community that was allowed to survive. We estimated for each site and treatment the proportion

of times the similarity increased compared to the pre-treatment condition and averaged the proportions among sites.

We repeated the previous analysis using the new dominant and the non-dominant species observed in each treatment to subset the community. In some cases, all the species in two or more plots in the same period, treatment and site were labeled as new dominant species. In those cases we assumed that the similarity between any two of these plots was one as the plots share a similar condition: all the species are new dominants.

3.3.4.3 Different deterministic processes: limiting similarity and habitat filtering

To test if limiting similarity or habitat filtering were more relevant as deterministic mechanisms to explain the changes in the communities we measured the average relatedness (mean phylogenetic distance, MPD) of the species in each plot at the end of the experiment using the package *picante* (Kembel et al. 2014). To control for different number of species in different plots we compared the observed value with 4999 runs obtained randomly by selecting the same number of species in the plot from all the species observed in each site at any moment during the experiment (null model "richness" in the package *picante*). We summarized the information estimating the standardized effect size of MPD (SES-MPD) in each plot at the end of the MPD of the random and the observed samples, divided by the standard deviation of the latter. When SES-MPD is negative, the observed species are more closely related than expected by chance, indicating limiting similarity. We tested whether the standardized effect size was similar among treatments using an ANOVA test, with site as a fixed term.

MPD is a metric that captures the overall relatedness among species. However, competitive interactions are expected to be stronger among closely related species than among distantly related ones. Therefore, we re-did the previous analyses using the standardized effect size of the mean nearest taxonomic distance (SES-MNTD). MNTD averages the relatedness of each species with its closest relative. In comparison with MPD, MNTD is more sensitive to the structure of the tips of the tree, while MPD is more sensitive to the topology of the base of the tree.

We repeated the last two analyses using different sub-set of species: only the new dominant species as identified in each treatment in each site and the non-dominants. To keep consistency among tests for communities defined in different ways, we used only presence-absence indices. We estimated the average treatment effects and the partial residuals and plotted them with the package *visreg* (Breheny and Burchett 2018).

3.3.4.4 Testing neutrality using species trends

We explored in more detail the transformation of the community in the dominant removal treatments to test if the observed pattern corresponded to neutral mechanisms. In particular, we focus on a prediction of the neutral community: after the dominant species are removed all the species increase their abundance, and the species more likely to occupy the space are those that were more abundant before the dominant species were removed. This should be the case because the local species abundance and the frequency rank are constrained by the regional species abundance rank for all the species and because the regional species abundance rank changes more slowly than the local abundance rank can. To validate the expectation with the observed data, three conditions should hold: first, the frequency ranking of the new dominant species at the end of the experiment should be correlated with the frequency ranking before the treatments were applied. Second, the species that became the new dominants should have had the higher frequency rankings before the treatments were applied. And third, the new dominant species in the dominant removal plots should occupy more volume than in the control plots at the end of the experiment because they should be filling the gap left by the site dominant species. These conditions are not exclusive of neutral mechanisms, but they should be true if neutral mechanisms prevail.

We measured the correlation between initial and final rankings of new dominant species (first condition) using Kendall's τ and tested the difference in initial ranks between new dominants and remaining non-dominants using Wilcoxon rank sum test with continuity correction (second condition). We used frequency for these two tests because neutral theory predict they should be correlated and to avoid duplicating a criterion already used to define local dominance (volume). Finally, we used generalized least squares models with different variances per treatment to test if each new dominant species increased their volume compared with the other two treatments (third condition). For this last test we included only sites in which the focal species became new

dominant because only in these sites we expected to find a treatment response. Site and treatment interaction was tested for each model and kept if significant.

3.4 Results

3.4.1 Treatment efficacy

In the five study sites, the treatments were effective in removing different amounts of biomass, small in non-dominant removal (NDR) and non-dominant matching (NDM) treatments, large in dominant removal (DR) and dominant matching (DM) treatments. Also, the site dominant species occupied most of the biomass in all treatments except dominant removal (DR), confirming that the DR treatments were effective in reducing drastically the volume occupied by the site dominant species (Figure S10). It is important to notice that the volume of the new dominants increased steadily in the DR plots, but never increased as much as to compensate for the removed biomass.

3.4.2 Deterministic vs. stochastic trends

As we expected, dominant and non-dominant species faced different degrees of change in stochasticity and determinism, even in DR, where all the original dominant species were replaced (Figure 10). In DR and NDR treatments the deterministic processes acting on dominant species became stronger, while the stochastic processes increased on the non-dominant species. When the biomass was removed randomly the pattern was reversed: stochasticity increased in dominant plants and determinism increased in non-dominant ones.



Figure 10 Effect of the treatments on community dissimilarity measured as the proportion of times the measured Bray-Curtis dissimilarity was lower than the dissimilarity before the treatments were applied.

Treatments are represented in individual panels: Ctrl: Control; NDM: Non-dominant matching, NDR: Non-dominant removal; DM: Dominant matching; DR: Dominant removal. Values higher than 0.5 indicate convergence after the impact (deterministic mechanisms driving a change in the community) while values lower than 0.5 indicate divergence after it (stochastic mechanisms driving a change in the community). Similarity values were measured using different sets of species: 'all' indicates all species were included in similarity estimation; 'New' indicates the dominant species defined in the last year in each site/treatment were used; and 'ND' indicates that the species not flagged as dominants were used. Dashed line shows the average value observed for control plots when all the species are included.

The difference between dominant and non-dominant species was not apparent in the control treatments (Ctrl). The control treatment also indicated an increase in deterministic processes among plots in most sites, similar for new dominant and non-dominant species. This site-level change in control plots was robust to the use of the period with maximum biomass only (around August) in each year (Figure S11), consistent with deterministic changes acting at a site scale.

3.4.3 Different deterministic mechanisms: Limiting similarity and habitat filtering

At the end of the experiment, the dominant and non-dominant species in the communities tended to differ in the relative effect of limiting similarity and habitat filtering. Mean phylogenetic distance (MPD) suggested that dominant removal treatments increased the clustering in the communities compared with the control treatment (Figure 11a) mostly because it increased the clustering of the dominant species (Figure 11c). However, we found the opposite effect, an increase in over-dispersion, when using mean nearest taxonomic distance (MNTD, Figure 11d).



Neither habitat filtering nor limiting similarity seemed to have a strong impact on non-dominant plants in the dominant removal (DR) treatments compared with the control plots.

Figure 11 Partial residuals (dots), confidence band (gray box), and prediction line (blue solid line) of the standardized effect size (z-scores) of the mean phylogenetic distance (SES-MPD) and mean nearest taxon distance (SES-MNTD) for each treatment and for all the community (All, panels a,b), dominant species as defined at the end of the experiment (panels c,d) and non-dominant species (panels e,f).

Dotted lines represent z-scores of ± 1.96 , and the dashed line the z-score of zero. Significance: *** < 0.001, ** < 0.01, * < 0.05, \cdot < 0.1.

Random biomass removal (DM and NDM treatments) had stronger effects on non-dominant plants: non-dominant plants clustered more in DM and NDM than in control plots when a large amount of biomass was randomly removed (MPD, p < 0.05; MNTD, p < 0.1). Random removal of small amounts of biomass had the opposite effect increasing the non-dominant overdispersion. Sites showed a sharp contrast in the relative weight of over and under-dispersion among them (Figure S12).

3.4.4 Testing neutrality using species trends

We found little support to the assumption that neutrality is an important mechanism structuring these communities. If species were neutral, the final frequency of the new dominant species should be correlated with the original correlation of those species (condition 1 in methods section testing neutrality). Even when the correlations were positive (Kendall's $\tau > 0.3$, Table 3), they were significant only in one site (p-value = 0.04). We also found that the species with initial higher frequency ranking were not more likely to became new dominant than those with lower initial frequency ranking (Wilcoxon rank test p > 0.1, condition 2). Finally, only 3 of the 17 new dominant species in different sites increased their volume in the dominant removal treatments compared with the other treatments (p < 0.05, Table 4, condition 3), 12 increased but nonsignificantly (p > 0.05), and, unexpectedly, the volume of two species decreased when they became dominant (*Vicia cracca*, p < 0.05; *Linaria vulgaris*, p > 0.1). These last two species were present in every site, but only became dominant in the site with lowest diversity. The DR treatment consistently increased the variance of 15 of the 17 new dominants species, probably because the species grew in some plots but not in others. In contrast, and as expected, the DM treatment created a really heterogeneous response in the volume mean and variance of the 17 species. The other treatments had no consistent effect on the new dominant species (Table S4).

Table 3 Changes in species ranking in the dominant removal treatments compared with the

expectations for a neutral model.

New dominant plants in a site were the two or more largest species needed to obtain at least 50% of total volume in each plot. Ranking was measured as frequency in the pre-treatment visit for the initial conditions, and for the spring and summer visit of the final year of the experiment. The expectations for a neutral model are that (1) the initial and final rankings are positively correlated (Kendall's τ test) and (2) that the initial ranking of the new dominant species will be higher (lower values) than the ranking of the other species present in the dominant removal plots still present in the control plots at the end of the experiment (Wilcoxon rank test).

Site		New dominant species		es	Initial rank of non-	Rank test of new dominant		
	Initial richness in DR plots	Initial ranking ^(a)	Final ranking ^{(b)†}	Initial-final ranking correlation (Kendall's τ)	dominant species present in control plots at the end ^{(c)++}	vs. non-dominant species (initial conditions) (between columns a and c)		
		-	_	(between columns a and b)		W	p-value	
Т	6	1,2,3,4	1,2,3,4	1.0*				
F	13	1,2,3,4,5,6	4,2,1,4,7,2	0.072	6	0.5	0.313	
С	13	1,4,7	1,5,5	0.816	1,3,5,5	6.5	1	
S	15	2,5,7, 8,12,12	1,7,4,1,9,6	0.357	1,3,4,10,12	19	0.519	
N	31	1,2,5,15, 20,25,32 [¥]	1,1,1,14, 10,23,20	0.720	2,2,6,7,8,9,10,10, 12,13,14,16(x3), 20(x4), 25(x4)	72	0.818	

Note: Species sorted by ranking can be found in Table S3. * p-value < 0.05. [†] Species sorted in the same order as in column (a). ^{††} (xN): N species with the same rank. [¥] Species absent at the beginning of the experiment.

Table 4 Effect of treatments in new dominant species in comparison with the control plots

at the final visit of the experiment.

All models were fitted using generalized least squares with different variance per treatment, and volume log+1 transformed using only sites where the species became dominant. Despite not all models fitted properly, we kept a homogeneous approach for consistency among them. Control treatment variance is always 1. Arrows indicate average positive (\uparrow) or negative (\downarrow) effect of the treatment. Significance: *** < 0.001, ** < 0.01, * < 0.05, · < 0.1. Effect of the other treatments and regression fit descriptors are presented in Table S4 and effect of site in Table S3.

Species	Dominant (-), new dominant (+) and other species (p) by site				Treatment variance rate to control variance		Treatment trend		
	Т	F	С	S	Ν	DR	DM	DR	DM
Symphyotrichum novae-	р	+	+	+	+	1.58	0.83	^ ***	\uparrow
angliae									
Poa pratensis [†]	-	-	+	+	+	1.18	1.43	\uparrow^{**}	1
Euthamia graminifolia		р	+	р	+	1.79	0.95	\uparrow^*	\uparrow
Bromus inermis	+	-	-	-	-	2.75	1.15	۲·	\checkmark
Daucus carota			р	+	р	1.78	2.41	Ť٠	Ť٠
Vitis riparia		р	р		+	++		\uparrow	
Carex sp.		+	р	р	р	4.39	1.80	\uparrow	Ť٠
Monarda fistulosa				+	+	3.54	2.11	\uparrow	\uparrow
Symphyotrichum ericoides		+	р			3.10	1.59	\uparrow	\uparrow
Solidago altissima	-	+	-	-	-	2.32	0.40	\uparrow	\downarrow^*
Cirsium arvense			р	+	р	1.85	0.81	\uparrow	\uparrow
Asclepias syriaca	+	р	р	+	р	1.58	0.19	\uparrow	1
Symphyotrichum lanceolatum			р	р	+	1.55	0.21	\uparrow	\checkmark
Crepis sp.					+	1.37	1.52	\uparrow	\uparrow
Hieracium caespitosum		+		р	р	1.32	0.09	\uparrow	\checkmark
Linaria vulgaris	+	р	р	р	р	0.85	1.01	\checkmark	\checkmark
Vicia cracca	+	+	р	р	р	0.67	0.71	\downarrow^*	\checkmark

Notes: [†]*Poa pratensis* was positively affected by DR treatment in every site but N. ^{††} The general approach used in this analysis generated an extremely large variance probably meaningless, caused by the presence of a large *Vitis riparia* in a single plot.

3.5 Discussion

The deterministic or stochastic nature of community assembly processes in ecosystems, and the relevance of species identity, is constantly studied and debated in ecology (McGill et al. 2007, Vellend et al. 2014, Tucker et al. 2016). By removing dominant species, we found that species responded differently suggesting that species characteristics are important, and therefore providing little support to neutral mechanisms, despite the existence of local stochastic mechanisms. Also, we found that deterministic mechanisms played a significant role but this role changed with the exposure of the plants to the environment: Dominant plants, by being more directly affected by environmental conditions, were strongly affected by habitat filtering and limiting similarity, consistent with an optimum competitive strategy controlled by the environment (Maire et al. 2012). In contrast, the less exposed non-dominant species diverged.

This divergence in non-dominant species created a less predictable pattern because divergence can occur in several ways.

3.5.1 The dominant species

We found evidence of three processes that could explain the observed deterministic trend in the new dominant species: an increase in the role of habitat filtering (Maire et al. 2012), an increase of limiting similarity (Figure 11) and site-level trends that prompted the increase in abundance of some species regardless the treatment (Figure 10a and b, Table 4).

The decrease in mean phylogenetic distance (MPD) of the new dominant species suggests that abiotic filters are more important than species interactions, and simultaneously, the increase in mean nearest taxonomic distance (MNTD) suggests the opposite pattern. This apparent contradiction can be the outcome if traits that control the general performance of a species in an environment are conserved and shared deeper in the phylogeny (e.g. stem and root architecture, cold tolerance), while traits that control performance in the presence of competitors are the product of trait divergence more recently (e.g. height, resource allocation strategies). This difference between MPD and MNTD is consistent with recent studies showing that competitive differences among plants appear to be under the control of relatively few traits, while many more traits influence species niches (Kraft et al. 2015b, Cadotte 2017), because as species diverge more of their traits are likely to diverge too. This is consistent with the original model of community assembly proposed by Weiher and Keddy (1995), as the species that cope with the new environmental conditions are a subset constrained by properties deep in the phylogeny, reducing MPD. Those species interact, reducing the survival odds of close relatives, which in turn increase MNTD.

Regional drivers also influenced the increase in similarity among new dominant species. Regional drivers include random site level changes in species abundance and changes of environmental conditions. Changes in environmental conditions can shift the optimum traits in the community (Laughlin 2014). Alternatively, a less stable environment can act as a stressor on the community (Chase 2010). However, regional trends are not synonym of neutral community assortment (Hubbell 2001b). On the one hand, the new dominant species were not always the second most dominant species (after the site dominants) at the beginning of the experiment.

Under a neutral assortment model we would expect the second most dominant species to be more likely to fill the gap after the removal of the site dominant species (Table 3).

Second, we found different species-responses (Table 4), which is by definition contradictory with neutrality. Five species exemplified the different species-responses: *Poa pratensis*, Symphyotrichum novae-angliae, Euthamia graminifolia, Vicia cracca and Asclepias syriaca. Poa pratensis was present in almost every plot at the beginning of the experiment and it was often the most abundant non-dominant species in sites dominated by *Bromus inermis* and *Solidago* altissima, thus its preponderance to becoming a new dominant can easily be explained by a either a neutral model or limiting similarity with *B. inermis*. In contrast, niche-related deterministic processes better explained S. novae-angliae and E. graminifolia abundance trends: the release of the space occupied by S. altissima allowed these species to occupy the emptied volume and, presumably, niche space. Interestingly, these three species belong to the same clade of the North American Astereae tribe (Noyes and Rieseberg 1999), suggesting that similarity and competition between the original dominant and these new dominant species could explain the contrasting response of MPD and MNTD (Mayfield and Levine 2010). The legume V. cracca, despite being frequent in every single site, only became dominant in the two less diverse sites. In these two sites the DR treatment reduced V. cracca volume (Table 4), suggesting a positive effect of the dominant species on V. cracca, and almost no other species were able to grow. Finally, most species responded like A. syriaca with marginally larger biomass in the DR treatments than in the control ones, and becoming dominant in only one site despite being present in all or most of the sites. The lack of strong response to the release of the dominant species can be caused by local stochasticity, historical contingency (Fukami 2015) or by species at their carrying capacity, in an scenario that species have limited impact on each other.

3.5.2 The non-dominant species

Non-dominant species responded more deterministically to the random removal of biomass (Figure 10). When large amounts of biomass were removed (dominant matching, DM) the phylogenetic dispersion of non-dominants decreased (Figure 11e,f), as expected if habitat filtering gains importance (Webb et al. 2002). The treatment itself is an artificial random disturbance that also created a more exposed (and variable) environment for the non-dominant species. This increase in variability (regardless the specific cause) can explain the increase in
habitat filtering. Because both MPD and MNTD decreased compared to the control treatment, we found little evidence for an increase in competition between close relatives as observed in dominant plants in DR plots. The small role of limiting similarity among non-dominant species can be caused by the large presence of site dominant species occupying most of the space, making the interactions between non-dominant and site dominant species more important than between non-dominant species.

In contrast, to the DM treatment that affected most plants in the plots, non-dominant matching (NDM) increased the randomness mostly of the non-dominant community without affecting the dominant species. This focalized effect of the NDM treatment occurred because often only small ramets were needed to match the biomass removed in the NDR treatment, driving the impact to mostly non-dominant species. The random removal of ramets probably opened new spaces among non-dominant species, and these new spaces were affected by stronger inter-specific interactions, explaining the relative higher impact of limiting similarity in the non-dominant species in the NDM treatment.

If non-dominant species are also affected by limiting similarity, why did non-dominant species appear more random in the DR treatment? We hypothesized that it occurred because limiting similarity with the dominant species ultimately means being different from those dominant species, but that difference can point in several directions and magnitudes. In the absence of a single optimum that maximizes fitness, historical contingency (Fukami 2015) and equalizing mechanisms (Chesson 2000) become more important, and both reduce the predictability of the community composition.

3.5.3 Long-term trends

The incomplete replacement of the site dominant species by new dominant species suggests that the patterns observed could change with time. However, a longer time should strengthen the interactions among species in the DR plots more than in the control plots, as the DR plots are still filling the volume and therefore defining the interactions among the species. This process should increase the discrepancy among new dominants and non-dominant species as both are driven by deterministic mechanisms. Also, longer time series could help to distinguish the role of ontogenetic variability, phenotypic plasticity and population dynamics in this process.

All together, even in this early years the information provided suggests a difference in the dynamic of dominant and non-dominant species: dominants being more deterministic than nondominants because the combined and simultaneous effect of limiting similarity and habitat filtering; in contrast, non-dominants are mostly affected by limiting similarity (Maire et al. 2012, Mariotte 2014). However, stochasticity and site level trends played an important role for dominant and non-dominant species. For instance, we expected the new dominant species observed in the dominant removal treatments to perform randomly in the control plots instead of converging (Figure S11), because they were originally non-dominant species and most of them were still non-dominants in the control plots at the end of the experiment. However, it is important to note that this conclusion presumes a relatively stable system. But the weather conditions were far from average: right before the experiment (2013) and during the first year (2014) the area experienced the two coldest consecutives winters in decades (CMOS-SCMO BULLETIN n.d.) followed by a warm and dry summer in 2016 and by a cold and rainy summer in 2017 (Phillips 2018). Changing environmental conditions can switch the equilibrium or just induce a more predictable community if inter-annual variability is a source of stress (Li and Shipley 2018). In both situations, the result is a switch in a subset of species that can outperform others and increase their dominance, reducing the dissimilarity among plots in a consistent way, as we observed. We have not enough information to rule out any of these mechanisms.

Finally, the removal of the dominant species aboveground and with limited soil disturbance in our experiment created a space that was not fully occupied even after four years, suggesting that underground species interactions can be important and long-lasting. Underground processes include legacies in microbial communities, the accumulation of resources in roots for subsequent growth as well as the survival of the root system of the removed species (Bardgett and Wardle 2010). An alternative hypothesis is that the removal of the dominant species increased the access to more aboveground resources but also increased the exposure to environmental stressors, as hypothesized in the introduction. The added stress caused by these variable conditions may offset some of the potential benefits until the plants in the community can cope with those new conditions either by arrival of new individuals better fitted for those conditions or by individual plastic responses. Longer experiments could help to clarify the role of these mechanisms.

3.5.4 A note on traits and phylogenies

Traits should provide a more deterministic way to represent species interactions and phylogenetic relatedness is assumed to be a proxy of traits similarity. However, it is not granted that phylogenies can properly describe trait similarity neither that the phylogenetic patterns observed can represent the limiting similarity and habitat filtering accurately (Gerhold et al. 2015, Cadotte et al. 2017). At the same time, it is not granted that any specific set of traits sampled will properly capture the hypothesized interaction mechanisms, nor that the chosen interaction mechanism is the most important one. Also, recording traits are time consuming and often destructive, so that only a limited set of traits can be sampled.

The phylogenetic approach used here (define the new dominant species, track their change along time and compare how MPD and MNTD of the new dominant species compare to non-dominants in control and dominant removal plots) provides two things: first, a way to narrow down questions. For instance, in the context of our specific experimental setting and results we can start asking which traits are conserved or not in the phylogeny, how deeply they are conserved, and how are the microclimatic or soil conditions changing with the different treatments? And second, this approach provides a replicable way to test for disparities in dominant and non-dominant mechanisms in any dominant removal experiment in plant communities. The later opens the door to asses the generality of these findings in previous dominant removal experiments using widely available phylogenetic trees (e.g. Zanne et al. 2014).

3.6 Conclusions

Community assembly theory transitioned from asking if communities were deterministic or stochastic, to assuming that communities can fluctuate between those two extremes, to more recently assuming that different mechanisms can act on different subsets of communities. Our results expanded our understanding of how processes can act on community components and showed that limiting similarity and habitat filtering, two contrasting deterministic mechanisms, act simultaneously in the same group of dominant species, while limiting similarity is more important for non-dominant species. These resulted in more deterministic trends for the dominant species than for non-dominant ones. In contrast with dominant species, non-dominants can be subject to more stochasticity, but also to a larger role of historical contingency.

Chapter 4

Dominant species converge, non-dominant species diverge: dominance, assembly rules and phylogenies interact in herbaceous ecosystems around the world

4.1 Abstract

Multiple lines of evidence suggest that dominant and non-dominant plants are different. For instance, the mass-ratio hypothesis predicts that plants transform their environment proportionally to their biomass. Hence, dominant species modify the environment in which non-dominant species live more than the opposite. Consistent with this disparity between dominants and non-dominants and with community assembly theories, in grasslands around the world dominant species were more closely related than expected by chance. Close relatedness is expected when a strong environmental filter allow only species sharing some traits to survive. Conversely, we found that non-dominant species were either randomly assorted or less related than expected by chance, showing disparity in the assembly processes of dominant and non-dominant species. Furthermore, several lineages scattered in the phylogeny had more non-dominant species, suggesting that traits related to non-dominants evolved multiple times and were phylogenetically conserved.

4.2 Introduction

Community assembly research aims to describe the mechanisms that drive the co-occurrence and co-existence of species and whole community patterns through time and across space (Weiher et al. 2011). However, usually unacknowledged is the fact that assembly mechanisms can vary for different subcomponents of a community (Pandit et al. 2009, Passy 2017) and that the difference between those subcomponents may not be random but predictable, and related to the abundance of the species in each subcomponent (Magurran and Henderson 2003).

One of oldest and most researched general patterns in ecology is that few species are very abundant, or dominant, and many more are not (Fisher et al. 1943). Both dominant and non-dominant groups of species are expected to have different characteristics with, for instance,

dominant species capturing more sunlight and other resources, while non-dominant species, especially in harsh environments, are often facilitated by dominant species (Lortie and Callaway 2006). Further, we should expect that local extinction of dominant species would be caused by deterministic processes and non-dominants are subject to stochastic extinctions (MacArthur and Wilson 1967). Dominant species fragment the environment where non-dominant species exist creating several small habitats adequate for smaller non-dominant plants (Aarssen et al. 2006). Finally, dominant species are more likely to modify the environment (e.g. light availability, wind exposure) than the non-dominants (sensu mass-ratio hypothesis, Grime 1998), despite some outstanding exceptions like legumes that can fix nitrogen from the air into the soil. These differences between dominant and non-dominant species can result in differences in how these two groups of species assemble into communities. In other words, the mechanisms that drive the co-existence of dominant plant species may not be the same as those for non-dominants.

Community assembly processes are often described as the interaction of three fundamental processes (Weiher and Keddy 1995, Weiher et al. 2011). First, dispersal limits the species that can physically reach a place. Second, for those species that reach a site, environmental constraints will only permit species that possess key traits to thrive there (e.g. frost tolerance). If those traits are shared by closely related species, the species that survive will be more closely related than expected by chance (clustered) (e.g. Webb et al. 2002, Gerhold et al. 2015). And third, species interactions (e.g. soil nutrient competition) will select for species that can coexist with one another, often selecting for species with niche differences. Thus, interspecific interactions can be manifest by the survival of species more distantly related (overdispersed) if the traits influencing niches are more similar in closely related species. These three processes, often called dispersal limitation, habitat filtering and limiting similarity, respectively, act simultaneously (Cadotte and Tucker 2017). Thus, clearly clustered and overdispersed phylogenetic patterns are only obvious when either habitat filtering or limiting similarity is stronger than the other, and if the relatedness of the species able to disperse to the area allow those trends to be observed (Chalmandrier et al. 2013).

Dispersal limitation, habitat filtering and limiting similarity are frequently employed to understand assembly mechanisms for a whole community, but since dominant and non-dominant plants are subject to different influences and ecological processes, we should expect that they will be undergoing different assembly mechanisms. In particular, we expect dominant species to

be more affected by habitat filtering than by limiting similarity because their larger volume should expose them more to the environmental conditions compared to non-dominants. Conversely, non-dominant species should be less affected by habitat filtering than by limiting similarity since the dominants shape and moderate local environmental conditions. This disparity in the strength of habitat filtering and limiting similarity in turn should imply a disparity in relatedness associated with dominance (hereafter relatedness disparity), with dominant species more phylogenetically clustered than non-dominant species.

This study explores the existence of relatedness disparity associated with dominance in herbaceous dominated ecosystems around the world. We used a global observational dataset (Borer et al. 2014a) that provides us with a unique opportunity to compare the relationship between species relatedness and dominance globally. In particular, (1) we determined if dominant species were more closely related to one another than non-dominant species within sites, and if this relatedness disparity was mostly driven by dominant species clustering or non-dominants overdispersion. Then, (2) we explored whether the dominant plant lineages in one site were more or less likely to be also dominants in different parts of the world. Because many of the sites are located in grasslands around the world, we expected to observe several grass species present in each site, some as dominants and others as non-dominants, and therefore a random pattern when the community is partitioned in dominant and non-dominant species.

Finally, (3) we searched for environmental variables and tree topology characteristics that could explain changes in the relatedness disparity around the world as a way to explore potential driving mechanisms (e.g. harsher environments increasing the role of facilitation and in turn increasing the disparity between dominants and non-dominants, or phylogenetic trees with even branches reducing it). For this last analysis we used the results of the first two questions to evaluate if a functional group (e.g. graminoids, forbs, woody species) explains changes in relatedness disparity among sites. Because we found that graminoids were often dominants, we added the working hypothesis that graminoid prevalence (proportion of biomass composed by graminoids) explains relatedness discrepancy and is explained by site productivity (represented as total standing biomass), and that legume biomass explains site productivity as a proxy of biological input of nitrogen (Figure 12). We assessed the role of these and other potential explanatory variables (descriptors of location, climate, human management, phylogenetic tree

topology) in explaining changes in relatedness disparity between dominant and non-dominant species.



Figure 12: Initial model representing the basic relationships (solid black arrows) of legume mass (LM), total biomass (BM), proportion of graminoids (pGr) and difference in the

dispersion of the dominants and non-dominant plants (Δ_{ses}).

The figure also represents the other potential interactions tested using stepbackwards variable selection (solid gray arrows). The gray box shows the interactions of the climatic variables and site location information. Dashed lines and boxes indicate that hemisphere (Hem), latitudinal distance to the Equator (lat) and elevation (Elev) are part of the site location description, and that precipitation (P), precipitation variance (PV), temperature (T) and temperature range (TR) are part of the climatic description.

4.3 Methods

4.3.1 Data sources

4.3.1.1 Phylogenetic information

We used the Zanne et al. (2014) phylogenetic tree (obtained from Phylomatic

http://phylodiversity.net/phylomatic/, accessed on 2017-08-16) to create a phylogenetic tree with

every vascular plant species present in the Nutrient Network (NutNet) dataset (Borer et al.

2014a). When a species was absent from the phylogenetic tree (44% of the observed species), we

looked for the other species of the same genus already in the tree, found the tip distance (branch length of the species terminal node to the closest node), identify the species closest to the median of the tip distances and included the missing species as a sister species of the species closest to the median. If the genus was absent, the new species was linked to the family node (4.6%). Overall, only in 8% of the 2355 genus-site combinations, a species absent in the phylogeny had one or more congeneric species in the same site, which supports the robustness of the replacement strategy.

To represent different phylogenetic topologies in different sites and assess their role in the observed relatedness patterns we pruned the tree to the species present in each site and estimated the number of species, Faith's Phylogenetic Diversity (hereafter PD) as a measure of phylogenetic history (Faith 1992), and three tree topology indices (Table S6). Phylogenetic manipulation of the data was done in R using functions in the APE (Paradis et al. 2017) and apTreeshape (Bortolussi et al. 2012) packages.

4.3.1.2 Site level data

We used pre-treatment data from the NutNet experiment as a global observational dataset. These data were downloaded on 2017-12-04, and any site with at least 30 sampled plots was included. In each site, plots were located at least 1.5m apart of each other. Each plot was composed by one 1x1m cover subplot and two 0.1x1m biomass subplots. The biomass subplots were located at less than 0.5m from the cover plot and the harvested biomass was pooled together. Species cover was measured for each species present in the plot, while biomass was measured as aboveground biomass for each functional group.

For each site, and using vascular plants cover data, we used three indices to rank species dominance: (1) the proportion of the plots where the species was present (*frequency*), (2) the mean species cover across all plots (assuming cover is 0 when the species was absent, *cover*), and (3) the mean species cover of the plots where the species was present only (*cover PO*). We focus on sites with a strong dominance of vascular plants (vascular plant cover larger than 90% of living mass cover).

We included site management descriptors (anthropogenic origin or not, presence of burning, presence of grazers) and biomass measurements from the NutNet dataset to identify variables

that could explain global changes in relatedness disparity (Table S6). In each site, we summed plot level functional groups biomass to estimate the total aboveground annual net productivity (BM), estimated legume aboveground annual net productivity (LM) and obtained the proportion of the living biomass that corresponded to graminoids (pGr). For this analysis we included only sites that had sampled functional groups biomass, and where the discrepancy in the proportion of graminoid cover and average proportion of graminoid biomass in each plot was less than 0.8, to prevent data inconsistencies. Climatic information was obtained from Worldclim 2 (Fick and Hijmans 2017).

We found 75 sites with at least 30 plots with cover before treatments were applied and 57 with complete biomass information (Figure S13). The sites were distributed in North America (39 total, 31 with biomass), Europe (11, 7), Australia (10, 10), South America (8, 3), Africa (4, 3) and Asia (3, 3). Overall, the data included 1784 species from 135 families of vascular plants.

4.3.2 Are dominant and non-dominant species similarly assorted? Estimating the relatedness of dominant, non-dominants and their disparity

To assess if the dominant species were more closely related to one another than species within the non-dominants we split the species in each site into three partitions (dominant, intermediate and non-dominant) each with a third of the species in the site according to the species rank. Even though forcing a symmetric partition may not be ecologically meaningful, we used this partition as it requires fewer assumptions and previous work has shown that it provides similar results to other ways to partition the community (Umaña et al. 2017). (See Appendix 2 for a comparison between partitions using this dataset, showing that all the partitions correlate and that partitioning the community in three symmetric groups provides conservative results.)

For these analyses, we ranked the species based on percent cover and focused on the first (dominant, D) and third (non-dominant, ND) partitions. For each site, we estimated the mean phylogenetic distance (MPD, average phylogenetic relatedness between all pair of species in the site) for dominants (D_{mpd}) and non-dominants (ND_{mpd}). To build the random expectations, we randomly sampled the same number of species observed in each partition from the species in the site. We repeated the randomizations 999 times and built a null MPD distribution. We estimated the standardized effect size for dominants and non-dominants (*e.g.* $D_{ses.mpd} = (D_{mpd} - D_{mpd})$

 MPD_{mean} /MPD_{sd}), where MPD_{mean} and MPD_{sd} are the mean and the standard deviation, respectively, of the observed and randomly generated MPD values together). Because of this normalization, the expected variance of each standardized effect size (SES) value is 1. The SES approaches zero when the observed set of species in a partition is random in relation to the species phylogeny, negative if the species are more closely related and positive if more distantly related.

We defined the relatedness disparity ($\Delta_{ses.mpd}$) as the difference between the SES measures of the dominant partition ($D_{ses.mpd}$) minus the non-dominant partition ($ND_{ses.mpd}$) for each site. A positive relatedness disparity ($\Delta_{ses.mpd} = D_{ses.mpd} - ND_{ses.mpd} > 0$) indicates that dominant species are more distantly related than non-dominants, as may be expected if limiting similarity (e.g. strong competitive interactions for soil nutrients) drives the assembly of the dominant species and habitat filtering (e.g. shade tolerance) is responsible for the assembly of non-dominant species.

We tested if dominant relatedness ($D_{ses.mpd}$) was different from zero globally, regardless the identity of the species in each site. We assumed each site's $D_{ses.mpd}$ value represents an independent observation and used a Shapiro-Wilks normality test first. If the site $D_{ses.mpd}$ were not normally distributed we used a Wilcoxon signed rank test to decide if the mean of the site $D_{ses.mpd}$ was lower, equal or higher than zero. If normal, we used a Kolmogorov-Smirnoff goodness-of-fit test (KS-test) to assess if the site $D_{ses.mpd}$ followed a normal distribution with mean 0 and variance 1, indicating random species sorting. If this KS-test was rejected, we kept the sign of the mean site $D_{ses.mpd}$ and did a second KS-test with the same parameters to confirm that the lack of fit was caused only by the mean. If the second KS-test was rejected we recorded if the variance was larger or smaller than 1. We repeated the analysis for ND_{ses.mpd} and for $\Delta_{ses.mpd}$. In the latter case, we used mean 0 and variance 2, assuming $\Delta_{ses.mpd}$ was the difference of two normally distributed variables with mean 0 and variance 1.

We used mean nearest taxonomic distance instead of MPD and repeated the previous analyses to estimate the SES ($D_{ses.mntd}$, $ND_{ses.mntd}$ and $\Delta_{ses.mntd}$). MNTD is the average phylogenetic relatedness of all the species with its closest relative. Because MPD includes all species pairs,

MPD is more sensitive to the basal structure of the tree while MNTD is more sensitive to the structure of the tips of the phylogeny (Cadotte and Davies 2016).

We redid the estimates using the species ranking provided by the other two abundance metrics (frequency and cover PO).

4.3.3 Are certain lineages more likely to be either dominant or non-dominant?

Using the previously described partition of species, we tested if the species in each of the partitions in different sites belonged to similar lineages. We expected all lineages to be equally likely to be in any partition (e.g., a grass lineage should be equally likely to have dominant or non-dominant species). For each partition, we estimated two phylogenetic indices of betadiversity among sites, the Sørensen-derived phylogenetic multisite and the pairwise-dissimilarity (Leprieur et al. 2012, Baselga et al. 2017). We also obtained the nestedness- and turnoverfractions of the Sørensen's phylogenetic dissimilarities. These indices allowed us to assess if the changes in the similarity between the partitions were driven by loss of certain branches of the phylogeny (nestedness) or by replacement of them (turnover). The observed dissimilarities for each partition were compared against 1497 random global datasets. Each random global dataset was built by selecting a third of the species in each site and then re-estimating the Sørensen's indices (multisite and mean pairwise for each partition) for each randomly generated dataset. The null hypothesis we tested was all species are equally likely to be in any dominance partition. To control for potential biogeographic bias, we repeated the analysis removing data from Australia, which tends to be unique in several biogeographic aspects, and North America, where 39 of the 75 sites are located.

To identify whether some lineages were more likely to be dominants we counted the number of sites in which any species of that lineage was dominant and compared that value with the total number of species of that lineage in any site. As in the previous test, we assumed that the probability that any taxon in any site of being in each dominance partition was identical (1/3), and ran a binomial test in each branch with 10 or more counts in that lineage. We repeated the analysis with the intermediate and non-dominant partitions.

4.3.4 Are there environmental conditions, topological characteristics of the phylogeny or biogeographic aspects that drive the relatedness disparity?

We explored whether some functional groups, tree topology, site productivity, climatic conditions, geographic location or human management of the site could help to explain changes in the relatedness disparity around the world. First, based on the results from the previous two questions, we identified graminoids as a functional group most likely affecting the relatedness disparity. Second, we ran preliminary backwards-stepwise regressions (using AIC as the model selection criterion) to identify the subset of variables that were more likely to be important in explaining biological sources of nitrogen (legumes aboveground biomass, LM), productivity (total aboveground biomass, BM), the prevalence of graminoids (proportion of BM composed by graminoids, pGr) and relatedness disparity (Δ_{ses} , representing either $\Delta_{ses.mptd}$ or $\Delta_{ses.mntd}$). Besides the paths between these four variables and the site descriptors, we included direct and indirect paths between these four variables (Figure 12).

The backwards-stepwise regressions resulted in four models that included predictor variables that were direct and indirectly connected to the four response variables (LM, BM, pGr, Δ_{ses}). Because regression models only test for correlations, these connections represent causal paths that can be direct or indirect and can go in either direction (Shipley 2016). Using the Figure 12 as a metamodel, we combined the LM, BM, pGr and Δ_{ses} models into a directed acyclic graph (DAG) and retained only direct paths. To retain only direct paths we removed any predictor variable from a model if it was already included as a predictor of any ancestor variable (e.g., if A was predicted by B and C, hereafter A \leftarrow B+C, and B \leftarrow C, we kept only A \leftarrow B \leftarrow C; notice that in A \leftarrow B \leftarrow ... \leftarrow P, variables B to P are ancestor variables of A). We tested for missing paths in the DAG using path analysis (Shipley 2016), and added the covariates as explanatory variables) to the descendent ones. The links were added only when they were statistically needed (p-value of the partial correlation < 0.05). We used Fisher's C index to validate the assumption that the DAG was a valid predictor of the observed correlation among the variables.

Lastly, we replaced Δ_{ses} in the final path model with D_{ses} (and ND_{ses}) and searched for missing paths to test whether any of the predictor variables not correlated to Δ_{ses} could directly affect the

dispersion of either partition. These analyses were done using piecewiseSEM (Lefcheck 2016). An evaluation of legumes biomass as a predictor of total biomass can be found in the Appendix 3 and a detailed review of the model results in the Appendix 4.

All analyses were done in R version 3.4.2 (R Core Team 2017). Besides the previously mentioned packages, we used ggplot2 (Wickham 2009). Across the study, we assumed a statistical significance level of 0.05.

4.4 Results

4.4.1 Are dominant and non-dominant subcomponents similarly assembled? Relatedness of dominant, non-dominants and their disparity

As expected, in herbaceous dominated systems around the world there is a relatedness disparity ($\Delta_{ses.mntd} < 0$, $\Delta_{ses.mpd} < 0$, Table 5, Figure S14), with dominant species more closely related than expected by chance ($D_{ses.mntd} < 0$, $D_{ses.mpd} < 0$). This result is consistent for different dominance metrics (frequency, mean cover of the plots where species were present only and overall mean cover, the later including the effect of the first two) and phylogenetic metrics (mean nearest taxon distance, MNTD, and mean phylogenetic distance, MPD). Non-dominant species were overdispersed using MNTD ($ND_{ses.mntd} > 0$), but randomly assorted according MPD ($ND_{ses.mpd} = 0$). To simplify the results, hereafter, we will present only the results based on mean nearest taxonomic distance and mean cover per plot because (1) mean cover per plot captures the effect of the other two dominance metrics, and (2) MNTD is more sensitive to recent history in the phylogeny than MPD, and therefore is probably more sensitive to competitive interactions, which is the expected mechanism driving non-dominant species patterns.

Table 5: Expected and observed global phylogenetic dispersion patterns of the dominant

(D_{ses}) and non-dominant (ND_{ses}) partitions and the relatedness disparity ($\Delta_{ses} = D_{ses}$ - ND_{ses}).

The analysis was performed using mean nearest taxonomic distance (MNTD) and mean phylogenetic distance (MPD), and three different dominance metrics (mean cover, frequency and mean cover of the plots where the species is present only, cover PO). We used the Shapiro-Wilk test for normality, the Wilcoxon test for not normaly distributed values and the Kolmogorov-Smirnoff to test for mean and variance. Symbols: (-) Negative mean, undersdispersion; (+) positive mean, overdispersion; (0) zero mean, stochastic assortment; (hv) variance higher than expected variance; (NN) non-normal distribution.

Dispersion metric	Dominance criteria	Δ_{ses}	D _{ses}	ND _{ses}
Expected distribution		Normal	Normal	Normal
Expect	ed mean	$\mu < 0$	$\mu < 0$	$\mu > 0$
Expected	d variance	$\sigma^2 = 2$	$\sigma^2 = 1$	$\sigma^2 = 1$
MNTD	Cover	-	-	+
	Frequency	-NN	-	+NN
	Cover (PO)	-	-NN	+
MPD	Cover	-	-NN	0
	Frequency	-	-NN	0NN
	Cover (PO)	-hv	-NN	0NN

4.4.2 Are certain lineages more likely to be either dominant or non-dominant?

Dominant species around the world were more closely related, *i.e.* less phylogenetic dissimilarity, than expected by chance (p-value < 0.001, Figure 13). A small turnover fraction of the total beta diversity explains the higher similarity of the dominant plants, which implies that some lineages were more commonly present than expected by random assembly. Conversely, the nestedness fraction of dissimilarity of dominants was larger than expected by chance. When nestedness fraction is measured using species, a large value indicates a strong reduction in the number of species, by extension, a large nestedness fraction in this phylogenetical turnover index indicates that in sites with fewer species, the species present belonged to fewer lineages than expected by random assembly. The pattern was reversed for non-dominant plants (*i.e.* against a random distribution, the observed values indicate more dissimilar lineages around the world, several lineages appearing in different sites and lineages getting lost more heterogeneously) while the intermediate partition was indistinguishable from random assembly. These patterns were robust to the exclusion of Australian or North American sites, and to the use of multisite and mean pairwise index (Figure S15).



Figure 13: Phylogenetic dissimilarities among sites when each site is partitioned into dominant, intermediate dominance and non-dominant species, each partition with a third of the species.

The total phylogenetic dissimilarity is measured as the multisite Sørensen (SOR), and decomposed in turnover (SIM) and nestedness (SNE) fractions. Dashed lines represent the observed values when species dominance is assigned base on mean cover per plot, while the black solid line represent the probability of a given dissimilarity value if the species were randomly distributed in the three partitions.

Grasses and sedges mostly comprised the lineages more often found in the dominant partition, and were also less likely to occur in the intermediate or non-dominants partitions (Figure 14). Within the grasses (113 genus in the family Poaceae), several genera were more often associated with dominant species (e.g. *Bromus, Elymus, Calamagrostis, Poa, Panicum, Sporobolus*), but only *Carex* of the 15 genera of sedges (family Cyperaceae) was more likely to be dominant. Among non-graminoids, only goldenrods (*Solidago* spp. and another genus in the same lineage) were likely to contain dominant species (p < 0.05). Acanthacea and *Hypochaeris radicata* may be more associated with dominant species (p < 0.05), but were reported in less than 15 sites, so their status is inconclusive.



Figure 14: Phylogenetic tree of species observed in the experiment showing the probability of a lineage to be dominant, intermediate or non-dominant.

The dominance partitions were determined at each site independently, with a third of species in each site in each of the partitions. The black solid outside arc indicate grasses, the black dotted arc legumes and the gray arcs indicate forbs. A gray edge indicates that the lineage was present in less than 10 sites (not enough cases to take a decision) or that the proportion is not different than 1/3 (p-value > 0.1). Red colors indicate proportion lower than expected, and green ones higher than expected. The edge's width indicates the significance of the difference.

On the other hand, more than a dozen different lineages were associated with non-dominant species more often than expected by chance (p < 0.05). In contrast to the strong dominance of the graminoid lineage, non-dominant lineages cover a large portion of the phylogenetic tree. Among the monocots, several genera in the orders Liliales and Asparagales (e.g. orchids) were more often non-dominants. The dicot lineages more likely to be identified as non-dominant included species of the genus *Brassica* (Brassicacea), *Geranium* (Geraniaceae), *Oxalis* (Oxalidaceae), *Ulmus* (Ulmaceae), *Polygala* (Polygalaceae), *Silene* (Caryophyllaceae), *Ratibida* (Asteraceae), *Taraxacum* (Asteraceae), *Myosotis* (Boraginaceae), *Gentiana* (Gentianaceae) and *Castilleja* (Orobanchaceae).

4.4.3 Are there environmental conditions, topological characteristics of the phylogeny or biogeographic aspects that drive the relatedness disparity?

Based on the results of the two previous questions, we decided to use the proportion of graminoids of the total biomass (pGr) as a predictor of relatedness disparity ($\Delta_{ses.mntd}$) to capture the role of this functional group. The regressions for legume biomass (LM), total biomass (BM),

pGr and $\Delta_{\text{ses mntd}}$ included two highly correlated variables: Faith's phylogenetic diversity (PD) and richness. We retained PD as it generated better individual regressions (higher r^2) than richness. Also, precipitation variability was dropped from each individual regression as a predictor in the backward-step procedure, so we removed it from the final model. Conversely, tests of missing paths indicated that pGr increased with temperature range (TR), so we restored this path after it was dropped when removing predictors present in ancestor variables. After performing all the adjustments, the final model properly represented the observed data (Fisher's C = 88.27, df = 78, p-value = 0.2 indicating good fit) and partially supported our original expectations: (1) legume biomass was positively related to total biomass (regression $r^2 = 0.47$. coefficient p = 0.007), (2) total biomass was positively related to the proportion of graminoids but not significantly (regression $r^2 = 0.38$, coefficient p = 0.135) and (3) the proportion of graminoid biomass was negatively related to $\Delta_{ses mntd}$ (regression $r^2 = 0.15$, coefficient p = 0.003) which implies that higher graminoid prevalence was associated with more closely related dominant plants compared to non-dominant plants (Figure 15). Marginal paths (p < 0.1, Table S5) appeared between $\Delta_{\text{ses.mntd}}$ and grazing (positive effect), productivity (positive effect), latitude (positive effect) and temperature range (negative effect). Other aspects of the final path model are discussed in the appendices 3 and 4.



Figure 15: Causal paths between site location (latitude, Lat; elevation, Elev; hemisphere, Hem), climate (precipitation, P; temperature range, TR; mean temperature, T), tree topology (Faith's PD, PD), management (anthropogenic, An; grazing, Gz), productivity (legume and total aboveground biomass, LM and BM, respectively) their effect on the

proportion of graminoids (pGr) and phylogenetic dissimilarity between dominants and

non-dominants ($\Delta_{ses.mntd}$).

Arrows are included for every path modelled, and point from the predictor to the response variable. Black arrows indicate positive impacts and gray arrows negative ones. Dashed arrows are non-significant paths (p-value > 0.05) and solid ones are significant paths (p-value < 0.05). The letters in the boxes represent the variables modelled (see Table S6) and the numbers represent the proportion of variability explained by the linear model (\mathbb{R}^2). Overall fit of the model indicated that it was unlikely to be any paths not represented in the model and that directions of the paths are consistent with the data: Fisher's C = 88.27, df = 78, p-value = 0.2

When replacing $\Delta_{\text{ses.mntd}}$ with dominant relatedness (D_{ses.mntd}) in the final model, the model explained the observed variance (Fisher's C = 76.33, df = 78, p-value = 0.532) indicating that the proportion of graminoids was the only modeled variable negatively affecting the dominant partition relatedness (regression $r^2 = 0.08$, coefficient p-value = 0.033). A consistent pattern was observed when we replaced $\Delta_{\text{ses.mntd}}$ with the non-dominant partition relatedness (ND_{ses.mntd}): the proportion of graminoids was the only variable directly affecting non-dominant dispersion (Fisher's C = 89.32, df = 78, p-value = 0.179), with a positive impact (regression $r^2 = 0.12$, coefficient p-value = 0.007). All together, these results indicate that the proportion of graminoids biomass is the key variable that controls the relatedness among species in the dominant and nondominant partitions, decreasing the former and increasing the latter.

4.5 Discussion

4.5.1 Dominant and non-dominant species follow different ecological mechanisms

The relatedness disparity between dominant and non-dominant species suggests different mechanisms acting on each of these partitions of the community (Ricotta et al. 2008, Lennon et al. 2011, Chai et al. 2016, Norden et al. 2017). For dominant species, the environment might provide a selective pressure that resulted in a single optimal strategy in each site that outperforms other species (Webb et al. 2002, Mayfield and Levine 2010, Kraft et al. 2015a, Cadotte and Tucker 2017). Alternatively, species from the most common dominant lineage, the graminoids, could interact more positively (or less negatively) amongst themselves than with forbs, facilitating the presence of other species in the same lineage compared with species from other lineages.

The non-dominant overdispersion was detectable when using the relatedness metric sensitive to tip distances (mean nearest taxonomic distance, MNTD) but not the metric sensitive to the basal structure (mean phylogenetic distance, MPD). Recent studies suggest that competitive differences are mostly associated to fewer traits than niche differences (Kraft et al. 2015b, Cadotte 2017). If we assume that fewer traits can change between close relatives than between distantly related species, the difference between MPD and MNTD suggests non-dominant species are less constrained by their niche differences (random MPD), and more by their competitive differences (overdispersed MNTD). This pattern is not consistent with dominant species creating several smaller fragments with heterogeneous environmental conditions, each with different optimal combinations of traits that relatively few species can occupy (Huston 1994, Aarssen et al. 2006). However, it is consistent with dominant species creating a new environment that equalizes the fitness of the non-dominant species, allowing species to coexist neutrally regardless of their ecological differences (Chesson 2000). This equalization is consistent with dominant plants engineering and homogenizing (but not depleting) the environment that non-dominants occupy (McIntire and Fajardo 2014). These results require to be further tested in local controlled conditions, and other derived implications tested more globally.

Biogeographic constraints could explain the large turnover fraction among sites observed for non-dominant species (Figure 13a). However, if biogeographic patterns were important to explain the relatedness disparity, dominants should be cosmopolitan and non-dominants should always have more restricted range, as is the case for graminoids. However, the pattern does not hold true because some lineages are non-dominant and cosmopolitan (e.g. orchids) while others are dominant despite having a restricted distribution (e.g. goldenrods).

4.5.2 Environmental drivers

The main mechanism controlling the magnitude of the relatedness disparity ($\Delta_{ses.mntd}$) was the prevalence of graminoids, measured as the proportion of biomass of that group. This pattern contrasts to the large number of variables that affected legumes biomass, total biomass and the proportion of graminoids. This is surprising as we may have expected a positive direct path between productivity and $\Delta_{ses.mntd}$ because facilitative interactions among distantly related species should become more important under harsh environmental conditions (Lortie and Callaway 2006). This path, as others, was only marginally significant (p-value = 0.071, Table

S5), which indicates that more research is needed to confirm that the proportion of graminoids is the only driver of $\Delta_{ses.mntd}$. We expect that descriptors relevant to species interaction at the plot level (e.g. soil nutrient availability, rooting depth, root structure, plant architecture) will explain the remaining variability in relatedness disparity, more than the site descriptors used here. It is also possible that the average phylogenetic distance between species in the dominant and the non-dominant partitions can provide more insights in the relatedness disparity between these partitions.

4.5.3 Non-dominance as a strategy

The more than a dozen clades with higher probability of being non-dominant than dominant species contrasts with classical formulations of theoretical ecology. Some classical life history frameworks, such as r-K (Reznick et al. 2002) and ruderal-competitor-stress tolerant (Grime 1974), were often used to identify specific traits that would allow a species to become dominant under specific conditions. These widely used frameworks are not explicit about non-dominant species life history, giving the impression that non-dominant species are those in the wrong environment given a set of traits. Consistent with this approach, Rabinowitz (1981a) suggested that non-dominant (rare) species are either failing, growing or strongly limited by other species, but that non-dominance (rarity) cannot be an "adaptive strategy" because if successful the higher fitness should reduce the evolutionary advantage of rarity, and because species more likely to be dominant should drive the non-dominants to extinction. If this is the case, and non-dominance is a transient state no lineage should be more likely to be non-dominant than dominant, unless the lineage can be dominant in a different biome. In contrast, we found more than a dozen genera distributed around the world more likely to have non-dominant species than dominant species. We hypothesize that there could be sets of traits associated with non-dominance that could configure "non-dominance" as a successful ecological strategy, that could explain why nondominant lineages are geographically widespread and repeatedly occur in the phylogeny. A superb example are the orchids: the family is composed almost entirely of non-dominant species in any biome they occupy, yet they are present on every continent and it is the most speciose plant family (27,801 of the 304,419 identified angiosperm species, compared with 11,554 species in Poaceae, The Plant List 2013).

Non-dominant species face some specific challenges, such as finding viable partners if sexual reproduction is needed (Farnsworth 2007, Vermeij and Grosberg 2018); but can also benefit from a more stable microclimatic environment created by dominant plants, reduced pest pressure or some other trade-offs (Rabinowitz et al. 1984, Aarssen et al. 2006). Non-dominance could also have evolutionary implications: smaller and more isolated populations could increase speciation rates or increase the odds of gene fixation. Because dominant and non-dominant species differ in the characteristics of the environment they face and in the restrictions on sexual reproduction, the origin of intraspecific trait variability (genetically driven vs. plasticity) may also differ. More work is required to confirm the presence of species traits relevant for non-dominant species and the importance of their role in the co-existence and evolution of non-dominant and dominant species.

4.5.4 Limitations and caveats

Our study sites are located in herbaceously dominated areas around the world (Borer et al. 2014a), which often has at least one graminoid dominant species, and frequently have several graminoid species. However, this bias in site selection cannot explain why in each site the dominant species tended to be more closely related among themselves, than expected under random assembly because under strong limiting similarity mechanisms we expect a single graminoid should outcompete most other graminoids from the dominant partition. The more parsimonious explanation is that the graminoids optimal conditions correspond to the local environmental conditions and, as a group, tend to be competitively superior to other life forms (sensu Mayfield and Levine 2010). Similarly, the selection bias cannot explain why the non-dominant species. Our findings are supported by similar results in other plant communities (Lennon et al. 2011) and by random patterns observed in communities comprised of species with little capability to modify their environment (e.g. chironomids in Siqueira et al. 2012).

Partitioning dominance in each site in two or three equally sized groups is simple but somewhat arbitrary (Umaña et al. 2017). Statistical criteria based on individual counts can distinguish core and resident animal species (Magurran and Henderson 2003, White and Hurlbert 2010), but they are difficult to apply to plants. A better approach could be to use traits to discriminate among plant species that are more likely to be dominants from those that are more likely to be non-

dominants. Such an approach could help to distinguish, for instance, a low-cover species with a less viable strategy from a species with a viable non-dominant strategy. A preliminary test of the relationship between probability of dominance and proportion of C3/C4 photosynthetic pathways in graminoids provides no additional explanation (see Appendix 5). We hypothesized that root and stem architecture may be better predictors of dominance strategy.

Furthermore, assigning species to dominant and non-dominant groups creates a dichotomy that may not reflect natural processes. Our results clearly suggest different dynamics operating in dominant and non-dominant species, but provides little information about the nature of the dynamics that drive the species in between these two extremes. Defining dominance and non-dominance strategies using traits could also help to distinguish between a real fully stochastic dynamic of species with intermediate dominance (as suggested by our results) from an apparent stochastic dynamic caused by a mix of dominant and non-dominant species with opposite deterministic responses.

4.5.5 Concluding remarks

Dominant plants in herbaceous ecosystems are more closely related than non-dominant species, indicating an asymmetry in the community dynamic. Dominant species show phylogenetic clustering, suggesting an important role for environmental filtering, while non-dominants show substantial phylogenetic variation suggesting that biotic interactions play a greater role. Our results show that these relatedness disparity associated to the dominance of the species is pervasive globally, and expand previous work that suggested dominant and non-dominant species were different (Huston 1994, Magurran and Henderson 2003, Laland et al. 2016, Chai et al. 2016, Umaña et al. 2017). We also found evidence that species dominance tend to be phylogenetically constrained suggesting that the traits that make a species more likely to be dominant are conserved in the phylogeny, and also, unexpectedly, that the traits that would make a species more likely to be non-dominant are conserved as well.

Chapter 5 Can partitioning plant communities by life history and scale dependent causality solve the productivity-biodiversity conundrum?

5.1 Abstract

Plant diversity and biomass production underpin our understanding of environmental functioning, and understanding their interaction is fundamental for managing natural systems and mitigating effects of global change. However, there is a wide diversity of theoretical and experimental studies with contrasting hypotheses about how diversity and productivity interact. In grasslands around the world we found that at the global scale, and after controlling for climate and human management, biodiversity appears to drive biomass and litter mass, two measures of annual productivity. However, at the smaller plot scale, the causality is reversed with biomass driving biodiversity. We show that these contrasting patterns are consistent with two different research questions relating biodiversity and biomass: how much biomass *can be* produced at a given locale given the biodiversity in the area vs. how much of this diversity *is* persisting locally given the biomass there. At both scales, these patterns were clearer when we partitioned the plant community into functional groups (legumes, forbs, graminoids). Different community assembly mechanisms among legumes, forbs and graminoids provide the basis for this dissimilar contribution.

Keywords: species richness, phylogenetic information, community assembly, ecosystem processes, causality, litter, grasslands, Nutrient Network, biomass

5.2 Introduction

The correlation between species diversity (i.e. species richness) and community productivity has long been observed globally across latitudinal and altitudinal gradients (Rosenzweig 1995 and references therein). However, to what extent this correlation is observed at local scales has been the subject of debate, with experimental evidence supporting the existence of a correlation between diversity and productivity (Cardinale et al. 2006, Tilman et al. 2014), but with contrasting results from natural systems (Mittelbach et al. 2001, Grace et al. 2007, Adler et al. 2011). One of the critical limitations in these correlations, is that not all species in a community

should be assumed to be equivalent, to the extent some species form coherent community subunits and are subject to different assembly rules (Grime 1998, Gibson et al. 1999). Therefore, acknowledging and characterizing the different groups of species comprising a community could describe ecosystem processes better than describing the diversity of the entire community as a single unit.

Most researchers acknowledge that species richness is of limited value to describe the correlation between productivity and biodiversity because different species accumulate biomass and litter in different magnitudes and qualities, and yet there is not a consensus on which biodiversity metric is best. One approach is to categorize species into general functional groups defined by shared attributes that strongly influence their interactions with resources and other species, and use the number of functional groups in a community as a predictor of biomass (Tilman et al. 1997). However, there is no clear consensus on how functional groups should be defined or why these provide mechanistic explanation for biomass production (Petchey 2004, Wright et al. 2006). An alternative approach would be to measure the phylogenetic relatedness among species in the community, under the assumption that phylogenetic information can capture the overall similarity of species and therefore provide more valuable information on how well the species overlap in the niche space and thus compete (Cadotte 2013, Gerhold et al. 2015, cf. Venail et al. 2015). A general limitation of these two approaches is that they assume that the community should be treated as a single combined unit. However, mounting evidence indicates that different species in a community can be subject to different assembly rules (Magurran and Henderson 2003, Henderson and Magurran 2014, Chai et al. 2016, Umaña et al. 2017).

There is also little consensus on the best way to represent the productivity of an ecosystem. Standing plant biomass is most commonly used for productivity estimates, but litter generated by deciduous plants and herbivory cause significant reductions in standing biomass and can bias productivity estimates. For that reason, some studies include litter mass in estimates of "total biomass" as drivers of diversity (Al-Mufti et al. 1977, Grace et al. 2016). However, the poor correlation between standing biomass and litter mass (O'Halloran et al. 2013) suggests that different mechanisms drive their accumulation.

Another strategy to improve the understanding of the biodiversity-productivity relationship in natural systems is to test different causal links in natural conditions using statistical tools, such as

path analysis and structural equation models, complementing manipulative experiments (Shipley 2016, Grace et al. 2016), to better understand how multiple factors influence productivity and diversity. Duffy et al. (2017) maintain that only with these statistical methods can we assess the impact of different mechanistic drivers of the biodiversity-productivity relationship. However, we argue that improving the statistical tools is not enough and that there is a need to formally distinguish two questions: the first is how much biomass and diversity can exist given the environmental conditions, that is the *potential* biomass-biodiversity relationship. The second question is how local interactions influence the correlation between biomass and diversity, namely the *realized* biomass-biodiversity relationship. For example, competitive interactions between species will cause a negative correlation between productivity and biodiversity (Huston 1994 and references therein), while a positive correlation can occur either by facilitation (Cardinale et al. 2002) or by warmer and moister conditions. In these examples, the first two mechanisms correspond to *realized* biomass production, while only the last one to the *potential* biomass production dictated by the environment (Table S12). In practice, the selection of predictors and response variables, and the spatial scale being measured, will strongly affect which relationship is actually being studied (Whittaker 2010).

Here we hypothesized that different functional groups do in fact follow different assembly rules and therefore convey different information about ecosystem processes and patterns. If this is the case, partitioning the community into functional groups should improve the understanding of the relation between diversity and ecosystem productivity, the latter represented as biomass and litter mass. Also, we hypothesized that the support from different theories to both directions in the causal links - biodiversity causes productivity and productivity causes biodiversity, are related to different spatial scales of the analysis (Whittaker 2010). We tested the partitioning of the community in functional groups and the potential scale-dependent-directionality using unmanipulated plots of a global grassland experiment, the Nutrient Network (NutNet, Borer et al. 2014a). The NutNet experiment measures biomass in forb, legume, graminoid and other functional groups that we have previously shown follow distinctive community assembly patterns and can therefore be biologically meaningful (Chapter 4). We assumed that comparing multiyear average values of unmanipulated (control) plots should provide a good estimate of the global correlation between diversity and potential biomass that is set by species pool, climatic, grazing and human management constraints (coarse focal scale). In contrast, we used within-site

variation of a single pre-treatment year to better capture the proximate interactions among species and therefore the correlation between diversity and realized biomass (fine focal scale). At this within-site scale, we expect the realized biomass will drive plot diversity. In both cases, we described the diversity using species richness and phylogenetic diversity, and partitioned the biomass and the diversity descriptors into graminoids, legumes and forbs. In all the analyses, we started with the assumption that litter accumulation was controlled by biomass and diversity. The best models at both scales were later expanded to test them more thoroughly.

5.3 Methods

5.3.1 Community information

Species percent cover, functional group biomass and litter mass was obtained from the Nutrient Network (NutNet) data repository on December 04th 2017. NutNet is a globally distributed grassland experiment, with almost a hundred sites around the world. At each site, plots are delimited and surveyed under natural conditions. Although many sites manipulate nutrients and grazing by large herbivores, here we focus on the unmanipulated plots, either control plots through time or pre-treatment plots in one year. First, we gathered data from sites with at least 3 control plots and with 5 years of data to characterize site level average productivity (dataset #1). Second, we used data before treatments were applied at sites with 30 or more plots to characterize the effects of species interactions (dataset #2).

To prevent bias in measurements, we (1) discarded sites with a large discrepancy in the ratio between graminoids biomass and cover (0.5 threshold difference) and (2) included only plots in which 95% of the total mass reported belonged to the key categories used in NutNet (i.e. litter, graminoid, woody, forb, bryophyte or legume). We focused on three functional groups for which biomass data was consistently available: legumes (family Fabaceae), graminoids (families Poaceae and Cyperaceae) and forbs (all other angiosperm species). Woody species and other tracheophytes were included for phylogenetic and richness estimates, but not for biomass because the sampling protocol was different. Of the 31 sites in dataset #1, only one had a large proportion of woody biomass (~40%) and one was intermediate (~12%); while all the other sites had less than 10% (correlation between total biomass and biomass of forbs, graminoids and legumes together was 0.99). In dataset #2, 10% of the 1892 plots (in 57 sites) had more than 10% woody biomass, and only one plot in an alpine grassland with scattered shrubs had a extreme

value of 92% woody biomass (overall correlation between total biomass and biomass without woody species was 0.99). We also found six sites with no within-site variability in litter production that we did not include in the dataset #2.

For broad scale environmental and human drivers we used location (North-South dummy, distance from equator in degrees and elevation in meters), climatic (Worldclim v2, Fick and Hijmans 2017) and management variables (NutNet dataset describing if the grasslands have anthropogenic origin, active burning management and if they are regularly grazed) to describe site characteristics in a global context (Table S6).

5.3.2 Phylogenetic information

We adapted the Zanne et al. (2014) phylogenetic tree to obtain a tree with all the species observed in the Nutrient Network sites. Any species absent in the original tree were assigned to a species with the median tip-depth value of the genus. If the genus was not present in the tree, we assigned the species to the family node as reported in the Zanne et al. tree.

We estimated Faith's PD (Faith 1992) to represent the total evolutionary history in the site for the full community and for the community subsets. Therefore, a zero represents no species of that clade living there, the clade depth will represent any monoculture of that clade and more species will increase the diversity of the site proportional the divergence between them. Mean phylogenetic distance (MPD) and mean nearest taxonomic distance (MNTD) represent overall community dissimilarity and average dissimilarity of the closest relative, respectively. MPD and MNTD were zero for monocultures (a species is identical to itself). However, MPD and MNTD are meaningless for empty communities because the dissimilarity of an empty community is meaningless. We estimated PD, MPD and MNTD using the total list of species present in control plots during the first four years of each site for dataset #1. For dataset #2, we estimated PD, MPD and MNTD for pre-treatment conditions in each plot. We estimated the same indices for legumes, graminoids and forbs. Phylogenetic trees were manipulated with the package *ape* (Paradis et al. 2017) and the indices were estimated with the package *picante* (Kembel et al. 2014). All the analyses were done in R (R Core Team 2017).

5.3.3 Statistical analysis

5.3.3.1 Potential biodiversity-biomass-litter correlation

We used the mean values of at least three control plots per site from 5-years (dataset #1) to estimate the effect of biodiversity and biomass on litter at a global scale. To assess the most likely causal direction in the correlation between diversity and biomass production we started building two models in which biomass was causally linked to diversity for each functional group, and litter biomass was caused by diversity and biomass (Figure 16b). Diversity was represented with species richness in the first model and with Faith's PD in the second. We built two more similar models in which diversity was caused by biomass (Figure 16d). The models were fitted using the package *piecewiseSEM* (Lefcheck 2016) and compared using Fisher's C. We also fitted the models in the package *lavaan* (Rosseel et al. 2017) using a bootstrap estimator, and compared the models with the robust χ^2 and p-values obtained using Bollen-Stine Bootstrap method (1000 samples) and other fit statistics of SE models (CFI, RMSEA and SRMR). The piecewise approach is robust to low sample size, but does not work when underlying causal links create cycles. The second method is robust to causal cycles, but requires larger sample sizes.



Figure 16 Basic models representing a situation where (a,b) diversity influences biomass and (c,d) biomass influences diversity.

Biodiversity and biomass are measured for the whole community in the left columns (a,c) and partitioned into functional groups on the right (b,d). In all cases litter is explained by the combination of diversity and biomass. Gray arrows are the only ones that change from one model to the other. Models a and c are fully saturated, and therefore the structural equation modelling cannot be used.

To assess the validity of partitioning the community into functional groups as a useful strategy we started using the litter part of the models previously fitted using functional groups and the piecewise approach, one with Faith's PD and the other with species richness, and compared them with models (1) using total PD and total biomass and (2) total richness and total biomass (Figure 16a). We also contrasted these four models with models with and without partitioning species into functional groups that included PD, MPD, MNTD, to represent total evolutionary history as well as overall similarity among species. To reduce collinearity in this comparison, we removed uninformative variables using a step-backwards approach, and compared the resulting models using BIC.

Using the best causal link directions, we expanded the structural equation models to represent the effects of climate, management and diversity on biomass and litter accumulation (Figure 17). We fitted models with and without functional groups, using phylogenetic distance and species richness. We added the biomass of the functional groups and compared the predictions of each model using Pearson R². Detailed specifications of this and other approaches, including one assuming the opposite causal connection between biomass and diversity, are described in the appendix.



Figure 17 Directed acyclic graph representing the effect of (1) location on climate; (2) climate, management and phylogenetic relatedness on functional group biomass production; and (3) management, phylogenetic relatedness and functional group biomass production on litter accumulation.

The coefficients are presented in the Table 8. Arrows indicate variables retained by the step-backwards selection process. Arrows inside the box were not tested part of the step-backwards process. Dashed and solid lines represent non significant (p-value > 0.05) and significant relationships (p-value < 0.05), respectively. Green lines indicate positive impact, while red lines indicate negative impact. The number in the bottom indicate the proportion of variance explained (R^2). Location variables: distance to equator (Lat), elevation (Elev), North-South hemisphere dummy (Hem). Climatic variables: precipitation, P, average temperature of the wettest quarter (TW), annual average temperature (T), and temperature range (TR). Management variables: grazed (Gz), burned (Br) and anthropic (An).

Biomass variables: legumes biomass (LM), graminoids biomass (GM) and forbs biomass (FM). Phylogenetic information was measured as Faith's PD (PD), mean phylogenetic distance (MPD) and mean nearest taxonomic distance (MNTD) for each functional group: legumes (lg), graminoids (gr) and forbs (fr).

5.3.3.2 Realized biodiversity-biomass-litter correlation

We used pre-treatment biomass and cover data at the plot level to explore the correlations between biomass and diversity and the relevance of partitioning the community into functional groups (dataset #2). As in the potential biodiversity-biomass-litter correlation, we tested first the most efficient approach for the direction of the causal link (among the options represented in Figure 16b and d) and about the usefulness of the information provided by the partitions using litter. Then, we expanded the best models into a more detailed analysis. We added site as a random term to capture mean differences among sites for the piecewise path analysis and as fixed component for the structural equation model. We did not use MPD or MNTD to prevent losing plots when species of any functional group were absent (a common problem for legumes). Mixed-effects models were fitted in R with the package *lme4* (Bates et al. 2014).

We analyzed in detail the effect of biomass as a driver of diversity at the local scale, which was the more parsimonious causal link obtained in the first step and is consistent with the models of realized diversity-biomass correlation. We started assuming that the diversity of a focal functional group (*FG*) in a plot of a given site ($DIV_{FG,site:plot}$) is a subset of the existent diversity in that site ($DIV_{FG,site:plot}$). The first mechanism that can explain a reduction in diversity is space constraints that decrease the number of species that survive in an average plot of that site (*SAMP*_{*FG,site*}). The second reason is related to intra-group species interactions: Increasing the biomass accumulated by the functional group should increase the diversity of the functional group that can co-occur in that plot unless some high productivity conditions make one species outcompete the others ($h(BM_{FG,site:plot})$), forming a hump-shaped relation (Al-Mufti et al. 1977, cf. Oksanen 1996). Finally, other functional groups interact one with each other, and can either increase, decrease or have a more complex impact on the focal group. Together, these four aspects can be represented by

$$DIV_{FG,site:plot} = DIV_{FG,site} + SAMP_{FG,site} + \sum_{j} h_{FG,j} (BM_{i,site:plot})$$

where *j* is any functional group, including the focal one, and $h_{FG,j}(BM_{i,site:plot})$ is the effect of the *j*-functional group biomass on the focal functional group. To capture a switch from positive to negative impact of biomass on diversity as predicted by hump-shaped diversity-productivity correlations we used linear and quadratic terms of the biomass after log(+1)-transforming the biomass. A site dummy captured the mean value which included $DIV_{FG,site} + SAMP_{FG,site}$, so that we could recover the spatial effect on diversity by subtracting the estimated value for the dummy minus the observed site diversity of the functional group. This difference could be absorbing other site-specific conditions. To compare the effect of spatial sampling among groups we calculated the ratio of diversity lost by functional group in each site ($SAMP_{FG,site}/$ $DIV_{FG,site}$).

We represented litter accumulation in a plot in a given site (*Litter*_{site:plot}) using the biomass and diversity descriptors used in diversity models. Litter results from the input provided by a functional group biomass from previous year that remains as litter until the next sampling season $(k(BM_{FG,site:plot}))$, while moisture, precipitation, bacterial load and other site level conditions affect the decomposition and consumption rate (D_{site}) . Also, plant species can differ in their decomposition and consumption rates $(d(DIV_{FG,site:plot}))$. Together, we represented the litter in a plot by

$$Litter_{site:plot} = \sum_{i \in FG} k(BM_{i,site:plot}) + D_{site} + \sum_{i \in FG} d(DIV_{i,site:plot})$$
(Eq. 2)

In correspondence with the biomass model, we assumed that the effect of biomass (k) was quadratic and the effect of diversity (d) was linear.

The four resulting individual linear models estimated using Faith's PD were combined in a single structural model and fitted using the package *lavaan* using a maximum likelihood (ML) estimator (Rosseel et al. 2017). Because of strong non-normality, we used bootstrap estimators (1000 replicates). We repeated the analysis using path analysis with the package *piecewiseSEM* (Lefcheck 2016) in R. We repeated the same analyses replacing Faith's PD with species richness.

5.4 Results

5.4.1 Biodiversity-productivity relationships among sites

Across the Nutrient Network dataset we found 31 sites with three or more control plots measured for 5 years, with biomass partitioned into functional groups and with cover records of legumes, graminoids and forbs. Global scale basic models indicated that partitioning always improved the resulting model (Table 6), and that phylogenetic information outperformed species richness as a predictor of litter, particularly when species similarity indices (MPD and MNTD) were included (Table 6). Further, biodiversity was more likely to cause biomass production than the opposite, regardless how diversity was measured (Table 7).

Table 6 Effect of richness and phylogenetic diversity metrics and of partitioning the

community into functional groups on litter at a global and local scale.

Diversity is described as either species richness (S), Faith's PD (PD), mean phylogenetic distance (MPD) and/or mean nearest taxonomic distance (MNTD) (n=31 for the global scale, 1892 for the local). Model's fit is described using Pearson's R², change in Akaike Information Criterion corrected for small sample size (Δ AICc), degrees of freedom used by the model (d.f.) and Bayesian Information Criteria (BIC). Each model can have different final predictors (not shown) because of the step-backwards optimization criteria. All the models without partitions at the global scale dropped all the initial predictors.

Partitions?	Initial predictors	\mathbf{R}^2 †	ΔAICe	d.f.	BIC
	Acceptance criteria	Closer to 1	Closer to 0		Lower
Global scale					
Whole community	Biomass, S	0.000	8.56	2	134.49
	Biomass, PD	0.000	8.56	2	134.49
	Biomass, PD, MPD, MNTD	0.000	8.56	2	134.49
Functional groups	Biomass, S	0.365	5.57	6	134.16
	Biomass, PD	0.355	2.96	5	131.22
	Biomass, PD, MPD, MNTD	0.524	0.00	7	128.66
Local scale					
Whole community	Biomass, S	0.020 [0.855]	92.2	4	4088.86
	Biomass, PD	0.023 [0.851]	85.1	5	4103.76
Functional groups	Biomass, S	0.033 [0.862]	32.9	8	4078.16
	Biomass, PD	0.031 [0.870]	0.0	9	4082.87

Notes: $\uparrow R^2$ values of the local models represent marginal (fixed terms only) and conditional (fixed and random terms combined, inside square brackets) R^2 values.

Table 7 Effect of the direction of the relationship between diversity and biomass

descriptors with basic structural equation models of diversity, biomass and litter globally

and locally.

Diversity is described as either species richness (S) or Faith's PD (PD). Models were tested using path analysis with Fisher's C (with degrees of freedom, d.f. and p-value testing that the structural model fits the model) and SE model (p-value, comparative fit index, CFI, root mean square error of approximation, RMSEA, and standardized root mean square residual, SRMR). All the models partition biomass and diversity in three functional groups (legumes, forbs and graminoids). Litter is always a function of the biomass and diversity descriptors (n=31).

	Path analyses				Structural Equation Models					
Model	AIC	AICc	Fisher's C	d.f.	p-value	χ^2	p-value [†]	CFI	RMSEA	SRMR
Acceptance criteria	Low	Low	Low		≥ 0.05	Low	≥ 0.05	≥ 0.9	< 0.08	< 0.08
Global scale										
S ← Biomass	87.2	244.9	41.2	6	0.000	36.6	0.000	0.318	0.601	0.147
PD ← Biomass	76.5	234.2	30.5	6	0.000	23.8	0.013	0.392	0.473	0.137
Biomass ← S	61.7	219.4	15.7	6	0.016	13.8	0.023	0.663	0.341	0.081
Biomass ← PD	54.7	212.4	8.7	6	0.194	7.3	0.087	0.797	0.216	0.067
Local scale										
S ← Biomass	98.6	99.5	44.63	6	0.000	29.7	1.000	0.998	0.072	0.0006
PD ← Biomass	73.9	74.8	19.88	6	0.003	11.9	1.000	0.999	0.042	0.0007
Biomass ← S	111.7	112.6	57.72	6	0.000	51.1	1.000	0.995	0.097	0.0014
Biomass ← PD	116.0	116.9	62.01	6	0.000	55.1	1.000	0.994	0.101	0.0015

Note: [†] Estimated using Bollen-Stine Bootstrap method.

The full structural equation model with partitions (Figure 17, Table 8) predicted total biomass (Figure 18, second row) and litter mass (Figure 18, third row) better than equivalent models without partitions (R² was 0.76 and 0.43 for total biomass and litter mass, respectively, vs. 64 and 40% in the second best model). The model shows that temperature range affected graminoids less than forbs or legumes, and only graminoids responded strongly to other climatic and management variables. We also found that once phylogenetic information was included, biomass of the competing groups provided no-extra information. In other words, the species relatedness was a better predictor of mean biomass of each of these functional groups than other group's biomass. Finally, litter mass increased with graminoids, but forb biomass decreased litter, which was probably caused by contrasting decomposition rates. Similarly, we found that graminoid PD decreased litter mass, suggesting a positive correlation of plant diversity with litter decomposition.

Table 8 Structural equation model representing the effect of (1) location on climate; (2) climate, management and phylogenetic relatedness on functional group biomass production; and (3) management, phylogenetic relatedness and functional group biomass production on litter accumulation.

Predictor	Temp. (°C)	Precip. (mm/year) †	Temp. range (°C)	Temp. wet quarter (°C)	Legumes (gr)†	Forbs (gr) †	Graminoid s (gr) †	Litter (gr)†
(Intercept)	32.07 ***	6.87 ***	32.24 ***	31.59 ***	-3.95	6.99 ***	-0.61	13.62 ***
North-South	-0.60 [-0.09]	0.07 [0.16]	1.37 * [0.51]	-0.77 [-0.08]				
Distance to equator (degrees)	-0.44 *** [-0.69]	-0.003 [-0.08]	-0.16 * [-0.59]	-0.334 . [-0.33]				
Elevation	-0.00 *** [-0.66]	-0.00 [-0.29]	-0.00 [-0.09]	-0.000 *** [-0.63]				
Temp. (°C)							-0.03 [-0.15]	
Temp. range (°C)					-0.19 * [-0.33]	-0.12 * [-0.31]	-0.10 ** [-0.22]	
Precip. (mm)†			-2.18 * [-0.36]		0.59 [0.17]		0.85 *** [0.33]	
Temp. wet quarter					LJ		0.08 *** [0.69]	
Anthropogenic							-0.95 *** [-0.41]	
Burned								
Grazed					1.43 * [0.39]			-1.29 [-0.25]
Forbs (gr)†								-0.97 ** [-0.46]
Graminoids (gr)†								0.94 ** [0.51]
Legumes (gr)†								
Graminoids: PD						-0.007		-0.03 ***
MDD					-0.02 *	[-0.639]	-0.02 ***	0.03 *
MIFD					[-0.42]	0.02	[-0.58]	[0.64]
D					[0.42]	[0.31]		
Forbs: PD						0.00 ** [0.53]		
MPD					0.03 *** [0.61]		0.01 * [0.21]	
MNT D					-0.02 *** [-0.66]		[0.21]	
Legumes: PD								
MPD							-0.01 * [-0.45]	0.03 [0.62]
MNT D							0.01 * [0.48]	-0.03 [-0.56]

Values represent coefficients; squared bracketed values represent standardized coefficients.

Predictor	Temp. (°C)	Precip. (mm/year) †	Temp. range (°C)	Temp. wet quarter (°C)	Legumes (gr)†	Forbs (gr) †	Graminoid s (gr) †	Litter (gr)†
SD	2.192	0.406	2.158	6.751	0.970	0.724	0.389	1.428
R^2	0.865	0.088	0.307	0.486	0.610	0.451	0.903	0.579
Normality test (p-value)	0.060	0.249	0.632	0.005	0.735	0.904	0.200	0.140

Notes: \dagger variable were log-transformed or log+1 transformed if zeros present. Cells' colors: Red: Variable not added. White: Included in the model, if no numbers, the variable was dropped by the step backwards algorithm. Orange: Variables fixed (not affected by step backwards algorithm). Green: Variables added after the piecewise SEM detected high correlation for the conditional variables. Model fit: Fisher's C = 182.48, d.f. = 184, p-value = 0.518.

Significance: . : p-value < 0.05, *: p-value < 0.05, **: p-value < 0.01, ***: p-value < 0.001.



Figure 18 Explanatory power of the different methods used in this study at the global scale.

Solid line represents the 1:1 relationship. The first row presents the explanatory power of each functional group model independently. In the second row, from left to right, the total biomass was estimated (left) as the sum of the
three functional groups, (centre) using a single community approach with Faith's PD and (right) without phylogenetic information. In the last two, species richness was dropped during the step-backwards procedure. In the third row, the studied variable is litter mass, and panels from left to right as in the second row. In this case, the litter model retained species richness in the right panel, but Faith's PD provided more information (centre panel). Alternative biomass models used in the second row are described in Appendix D.3.2, and alternative litter mass models in the third row in the Appendix D.3.7.

5.4.2 Local analysis

We found 51 sites with at least 30 plots with pre-treatment data and with biomass information categorized by functional groups (1697 plots in total). As in the global models, splitting the community into functional groups considerably improved explanatory power when predicting litter mass (Table 6). In contrast with the global models, we found that biomass production was more likely to cause diversity than the opposite (lower Fisher's C value, Table 7). Therefore, we analyzed in more detail the models with biomass causing diversity and partitioning the community. We fitted two models, one measuring diversity with Faith's PD and the other with species richness. Given non-normality issues, the maximum-likelihood estimates were based on Bollen-Stine Bootstrap (species richness model: $\chi^2 = 28.1$, df = 3, p-value = 0.0, BSB p-value = 1.0; Faith's PD model: $\chi^2 = 7.2$, df = 3, p-value = 0.066, BSB p-value = 1.0). The path also indicated a good model fit (species richness model: Fisher's C = 33.3, df = 6, p-value= 0.0; Faith's PD model: Fisher's C = 11.7, df = 6, p-value= 0.069). We did not add direct links between diversity indicators despite being identified as likely because it is not clear at this scale how diversity of a functional group can impact on another functional group diversity not mediated by a biomass change. The estimates obtained from the SE models and path analysis were similar, so we reported here only the ones obtained from the path analysis, but the ML results are detailed in the appendix D.3.1 of this chapter.

At the local scale, we found support for an intra-group hump-shape relationship between biomass and diversity in all cases using Faith's PD and but only for graminoids when using species richness (Figure S18). In all cases, the quadratic relation appears to be describing a plateau more than a hump-shaped relationship. Intergroup interactions were heterogeneous, often nonsignificant (p-value > 0.05). However, there was some evidence of interference of legumes mass on graminoids richness, as well as a hump-shaped effect of graminoids biomass on legumes diversity (both PD and richness). All the regression models were adequate ($R^2 > 0.6$), but the

largest part of the variance was captured by the site dummy, as expected given the global extent of the analysis. The spatial sampling effect was less predictable among sites for legumes compared to the other groups (wider range of values in Figure S19). Predictability of the spatial sampling effect changed considerably between species richness and Faith's PD, and also between graminoids and forbs.

Graminoid diversity and biomass increased litter mass, but forb and legume diversity decreased it, indicating a sharp contrast in the role of these groups (Figure S20). Legume biomass decreased litter mass too, but forb biomass had a quadratic response, first decreasing and then increasing litter accumulation.

5.5 Discussion

The global examination of natural herbaceous ecosystems provided strong support for our hypotheses related to causality, relatedness and partitioning. First, the direction of causality changed from the global to the local scale. Previous research has documented a change in the shape of the relationship between diversity and productivity, from linear at a large scale to notrend or hump-shaped at local scales (Chase and Leibold 2002, Gillman and Wright 2006, Bai et al. 2007). Here we found that not only does the shape of the relationship change, but also that changing the focal spatial scale of analysis reversed the actual direction of the underlying causal link. A general concern of the usage and interpretation of structural equation modelling is that other equivalent causal links are as likely to exist (Grace et al. 2010, Shipley 2016). However, we found consistency in our results in several tests of alternative models (see appendix D.3.9 for reversing the global model; local model expansion consistent with light availability mediating the effect of productivity on diversity (Borer et al. 2014b)). The change in directionality can be explained because the combination of several samples (averaging biomass or accumulation diversity) in time and space may disrupt the strong effect of short-scale interactions such as competition, allowing other processes as the portfolio effect (Figge 2004) to become apparent by increasing the odds of observing rare species for the current conditions. Even more, the two focal scales probably correspond to different research questions, that we have characterized as how much biomass/diversity can exist and how much actually occur, each of them related to different ecological processes (Whittaker 2010). For instance, the hump-shaped model proposed by Oksanen (1996) corresponds to the first question (richness is limited because no more individual

can exist in an area), while the one proposed by Al-Mufti et al. (1977) to the second one (richness is limited because species compete and some are excluded).

Second, we showed that phylogenetic information would improve our understanding of different processes mostly when metrics that estimate the pairwise relatedness were used. The high conceptual and statistical correlation between species richness and Faith's PD provides a basis for the similar and sometimes confusing results obtained when studies used one or the other (Cadotte 2013, 2015, Venail et al. 2015, Cadotte et al. 2017). However, phylogenetic information allows an indirect measure of similarity, something that species richness cannot provide. Therefore, phylogenetic information can outperform species richness because it can link community properties with theoretical expectation grounded in niche and evolutionary theories (Gerhold et al. 2015).

And third, we found strong evidence that partitioning the community into functional groups improved considerably the predictive capability of the models and showed that not all functional groups responded in the same way to environmental conditions, neither they affected litter accumulation or other groups similarly. As expected, forb and legume biomass increased with diversity, but in contrast to our expectations, graminoid biomass decreased with diversity (Maire et al. 2012). Therefore, the number of functional groups would be unlikely to provide information about ecosystem processes (Wright et al. 2006). Despite the fact that the number of functional groups is uninformative, the information contained within them, including biomass, richness or relatedness, is very pertinent for understanding community processes because it is likely that species within the different functional groupings might be subject to different responses to environmental conditions and interspecific interactions. In contrast to previous studies (cf. Steinmann et al. 2009), we used functional groups that have different assembly dynamics in this global experimental setting: grasses and sedges are under-dispersed and more likely to be dominant species, compared with forbs being over-dispersed and more likely to be sub-dominant species (Chapter 4). It is not clear if the functional groups used here represent the best strategy for partitioning the plant community. For instance, the negative effect of diversity on graminoid biomass can be related to strong habitat filtering mechanisms driving competitive interactions among grasses, but also to a poorly defined functional group. The latter may happen because only the genus Carex among the sedges was more likely to be dominant, compared with

several lineages of grasses; therefore, a random replacement of one sedge species by a grass species will always decrease phylogenetic diversity but will likely increase biomass.

Partitioning the community into functional groups also helped us to identify differences in their response to climatic and geographic drivers, despite the potential issues with the categories. Interestingly, graminoids were more affected by climatic conditions than forbs or legumes. Because graminoids are more likely to be dominants in this study system they are likely to modify local conditions, such as wind or evaporation losses, for the other species present on those ecosystems (Spasojevic and Suding 2012, Odling-Smee et al. 2013). It is also possible that the large diversity included in forbs hamper the identification of any climatic pattern, as different lineages could respond to different climatic drivers. The use of functional groups also improved the explanatory power of biomass-biodiversity models (cf. Grace et al. 2007, 2016), helped to find global trends in the relation between biomass productivity, diversity and litter accumulation that were not evident before (O'Halloran et al. 2013), and showed that interactions between functional groups is frequent, but sometimes better captured by phylogenetic information than by biomass.

In our study we focused on two spatial scales to study the mechanisms that underlie the biodiversity-productivity correlation, the plot scale allowed to capture short-scale interactions in $1m^2$ areas, and the site scale that by combining multiple plots allowed us capturing climatic and management gradients. A more comprehensive study could use continuous approaches (McGlinn et al. 2018) to identify the key aspects that drive the change in the directionality, and test if similar results are valid in more or less complex biomes, such as forests or deserts. The shifts in the directionality should also be placed in context of the macro gradient of productivity apparently driving diversity when different biomes are compared (Rosenzweig 1995).

The different causal directions in the biomass-biodiversity relation indicate differences in the key mechanisms acting at different scales. Because different functional groups can be affecting but also be affected by the environment in different ways, we propose that identifying relevant functional groups will deeply increase our understanding of global trends in plant communities regardless the spatial and temporal scale of the analysis. Partitioning the communities and describing causal links properly will most likely prove also useful to better comprehend other community assembly mechanisms and ecosystem processes.

Chapter 6 Conclusions

Community assembly theory is concerned with the mechanisms that specify which group of species co-exist at a specific place (Weiher et al. 2011). As in other realms of science, these mechanisms can be either deterministic or stochastic, and a long debate in ecology exists around the deterministic nature of community assembly (Vellend et al. 2014). The debate is academically interesting, but it is unclear if it is relevant in the context of impending threats to the environment from climate change and land use transformation (Millennium Ecosystem Assessment 2005). In other terms, whether communities are deterministic or stochastic is a really important question for understanding the future of ecosystems and of ecosystem services in the Anthropocene, when massive environmental changes threaten to totally transform vast regions of the earth.

I proposed two criteria that have to be met in order for community assembly studies to be important to predict ecosystem services. First, biodiversity information, the information contained in the genes and species in the world, has to be meaningful to predict ecosystem processes in general, and ecosystem services in particular. If biodiversity is meaningless for these purposes, then the deterministic or stochastic nature of the interactions among species is also meaningless in any practical sense. And second, to understand the future of diversity, the deterministic-stochastic nature of the interactions among species has to be as important, or more important, than current drivers of global change, such as climate change or land use change. These two problems are addressed in the fourth and first chapters of the thesis, respectively.

Mounting evidence indicates dominant and non-dominant species have systematic differences (Grime 1998, Magurran and Henderson 2003, Maire et al. 2012, Mariotte 2014), and therefore, can help to understand if communities are more deterministic or stochastic. I hypothesized that dominant species not only capture more resources than non-dominant species, but also change the environment where other species thrive more than non-dominant species. The effect of dominance in the determinism of herbaceous communities is studied in the second chapter using a local experiment, and in the third chapter with a global experiment.

6.1 Key results

In the second chapter of my thesis, I explored the effect of the deterministic-stochastic nature of species interactions in a patchy landscape and extended the island species-area relationship (sensu Whittaker and Triantis 2012) to fragmented landscapes. With this model, I showed that the degree of determinism or stochasticity in assembly mechanisms controls species turnover rate among patches, subsequently controlling regional richness. I applied this model to the tropical Andes and showed that the degree of determinism or randomness in species survival was the single largest source of uncertainty in predicting species loss over the next 50 years, followed by dispersal ability. In general, these two aspects, dispersal and determinism in species survival, generate more uncertainty in the total richness than the uncertainty caused by land use change or climate change.

Interestingly, purely stochastic and purely deterministic mechanisms can generate similar community patterns (McGill et al. 2007) and different deterministic mechanisms can generate very different communities (Weiher and Keddy 1995), making them difficult to distinguish. Yet, it is underappreciated that different species in the same community might be influenced by different mechanisms at the same time (Chalmandrier et al. 2013). Finally, it is not often considered that plants modify environmental conditions where other plants live (Jones et al. 1997), and because dominant plants affect the community more than the non-dominant plants (sense mass-ratio-hypothesis Grime 1998), the dominant plants should be more affected by the environment than by interspecific interactions. Conversely, non-dominants should be more affected by interspecific interactions than the dominant species, and less by habitat filtering. Chapters three and four explore these asymmetries between dominant and non-dominant species in a meadow with a local and a global experiment.

In the third chapter of my thesis I examined how dominance shapes communities with a fouryear dominant removal experiment in five meadows near Newmarket, Ontario, Canada. The results indicate that the non-dominant species were randomly assorted prior to the treatments, but exposure to environmental conditions or to random disturbance increased the similarity among them. My results are also consistent with (a) an asymmetry in the environmental conditions faced by dominant and non-dominant species that caused different assembly rules acting on them, (b) multiple traits influencing niche differences and therefore habitat filtering and (c) few traits

controlling fitness differences among species and therefore limiting similarity (Maire et al. 2012, Kraft et al. 2015b, Cadotte 2017). Trends for individual species ranged from being negatively to positively affected by the dominant species. Further, stochasticity was important, and temporal change in some species could be better explained by fluctuations in species abundance at the site level, but some apparently stochastic patterns could also be explained by deterministic assembly mechanisms acting simultaneously with opposite effects. All together I concluded that there is little support for neutral dynamics despite large stochasticity at the local level because of species-specific and dominance-related responses.

In my fourth thesis chapter I assessed the generality of the local experiment results of the second chapter using natural conditions data from 75 meadows and grasslands around the world from the Nutrient Network experiment (Borer et al. 2014a). This analysis confirmed a difference between dominant and non-dominant dynamics: dominant species were more phylogenetically clustered to one another, while non-dominants were over-dispersed. This pattern is consistent with dominant species being more affected by habitat filtering than limiting similarity, while the opposite is true for the non-dominants (Mariotte 2014). A couple of lineages were likely to contain dominants, but more than a dozen lineages scattered throughout the phylogeny were more likely to contain non-dominant species. This is not consistent with common understanding of evolutionary processes (Rabinowitz 1981b), in which the survival of rare species is seen as a conceptual difficulty (Aarssen et al. 2006). Our findings of several lineages more likely to be non-dominant suggest that remaining as a non-dominant species can be a successful evolutionary strategy.

Finally, my fifth chapter explored how dominance influences two basic environmental functions, biomass productivity and litter decomposition, using data from un-manipulated plots in 51 sites obtained from the Nutrient Network (Borer et al. 2014a). First, I reviewed several proposed mechanisms that may explain the causal connections between diversity and productivity, and found that diversity drives productivity at a global scale, but productivity drives diversity locally. Then I used graminoid biomass as a proxy of dominant productivity and forbs and legumes as subdominants, and modeled how the plant community 1) responded to the environment and 2) influenced ecosystem functioning. Partitioning the community into functional groups (graminoids, forbs and legumes) always improved the performance of the models and provided other insights such as different responses of functional groups to temperature range, the role of

competitive interactions between and within functional groups, and the effects of plant diversity on decomposition rates.

Together, the results show that biodiversity matters in meadows because it is an important predictor of biomass productivity, that disentangling the stochastic-deterministic nature of community assembly mechanisms is important to predict regional biodiversity, and that dominant species transform the habitat, decreasing the role of habitat filtering and increasing the role of limiting similarity of the non-dominant species.

6.2 Moving forward

These results put forth new questions and new ideas that can be used to solve old questions about the implications of stochasticity and determinism in natural systems, and how to use this basic conceptual knowledge to better understand two key aspects of ecosystem services: diversity and productivity.

6.2.1 Ideas coming from a landscape perspective

I found that species spatial-turnover rates are the most important unknown to predict landscape level extinctions of species. If this is the case, what are the turnover rates in natural conditions? Are there global trends? Several trends (e.g. latitudinal, real islands vs. naturally fragmented habitat vs. artificially fragmented habitat) have been reported for the slope of the species area relationship (Triantis et al. 2012), but the turnover rates have not been as thoroughly studied.

The richness model presented in chapter two shows that the parameter describing determinism and turnover has a large impact on regional species richness. However, some values of the turnover parameter can be unrealistic (like assuming that every patch in a habitat have a different set of species), while specific theoretical models support other values. The model can be extended to identify the parameter value ranges that match observed turnover patterns (Wang et al. 2012), observed slopes of ISARs and SACs (Triantis et al. 2012), the tendency to find that several patches contain more diversity than a single large patch (Fahrig 2017) or to the test validity of some theories (Hubbell 2001a).

Finally, a critical assumption of the landscape model is that the species in each habitat follow similar assembly rules. However, I have proven in chapters three and four that this assumption is

not true in meadows and grasslands, and is probably not true in other systems too (Aranda et al. 2013, Mariotte 2014). Therefore, I need to critically assess the robustness of the model to that assumption and how to adapt it to multiple dispersal and turnover patterns occurring in the same habitat for different groups of species.

6.2.2 Partitioning the community using a dominance criterion

Partitioning the community in dominant and non-dominant species, or proxies for them, was useful to describe the dynamics of the meadows that I studied. However, it is not clear if this is a peculiarity of the grasslands and meadows. In particular, does this relation hold in other biomes? Does it hold at other spatial scales? Can fertilization change the differences in mechanisms organizing dominant and non-dominant species?

As mentioned in this thesis, the asymmetry between dominants and non-dominant species can be caused, among others, by (1) similar random fluctuations acting on very abundant vs. non-abundant species (Hanski 1982), (2) differences in the amount of resources available to smaller and larger plants (Grime 1998), (3) differences in the variability (Jones et al. 1997) or (4) predictability of those resources, or (5) changes in the rules that control the organization of the plants (from a single optimum constrained by the habitat conditions, to multiple optimums characterized by a large distance in any direction from a specific point) (Maire et al. 2012). Each hypothesis may have different implications in the generality of the asymmetry, and distinguishing among them can also help to identify better criteria than dominance to partition the communities in an informative way.

The possibility of non-dominance as an evolutionary strategy is particularly interesting, and I think it could be confirmed if (1) across biomes similar clades are more likely to be non-dominant and (2) some traits converge among non-dominant species. There is enough evidence to support that reproductive traits will converge (e.g. low abundance reduces the odds of non-directional pollen dispersion) (Vermeij and Grosberg 2018), but this evidence points to common challenges of non-dominance, and not to common benefits. The existence of benefits is a potential hypothesis for multiple non-dominant lineages occurring in the phylogeny, and could explain why rarity is such common and predict effects of losing rare species in the wake of global changes. Differences in dominant and non-dominant species could also give us insights on the delivery of ecosystem services in a community. If non-dominant plants converge in a small

set of non-reproductive trait values, are those values related to particular ecosystem processes besides productivity? For instance, given that non-dominants potentially rely more heavily on pollinators, are non-dominants particularly important to support a more diverse community of pollinators (Garibaldi et al. 2016)?

Mean phylogenetic distance (MPD), a measure of similarity among species sensitive to basal relatedness, provided a similar response consistent with habitat filtering for dominant plants in the local and global experiment. But an index more sensitive to recent relatedness, the mean nearest taxonomic distance (MNTD), indicated an impact of limiting similarity on the dominant species in chapter three, and detected an effect of habitat filtering, not of limiting similarity, in chapter four. The inconsistency can be related to the difference in the definition of dominant species in the two analyses (2 or more species needed to occupy at least 50% of volume in chapter three, the third of species with highest average cover in chapter four), or in the differences between the experimental setting in chapter three and the observational approach in chapter four.

Differences between MPD and MNTD can also indicate a difference in the evolution of traits related to competition and those related to niche characteristics. If this is true, these two measures could be used simultaneously to identify the relative role of habitat filtering and limiting similarity in a community (Chalmandrier et al. 2013). To do so, it will have to be proven in different biomes that traits related to habitat filtering mechanisms are older in the phylogenies than traits related to limiting similarity mechanisms.

6.2.3 From diversity to ecosystem processes and services

I showed that the scale at which the productivity-diversity relationship is studied matters (Whittaker 2010), as it can actually capture different questions and different mechanisms. An immediate task should be to complete the characterization of the different mechanisms that could relate diversity and productivity started in chapter five, and assess how closely they match the reported patterns for herbaceous communities. Further, this analysis could be extended towards forests and other biomes to confirm the generality of these findings.

I also showed that partitioning the communities into functional groups was a useful strategy to predict biomass and litter accumulation using biological information. In fact, once functional

group diversity information was included into models, climatic drivers of litter accumulation were either uninformative or had counter-intuitive impacts on litter. Clearly, temperature has an impact on decomposition rates (e.g. Zimmermann et al. 2009), but temperature also has an impact on litter production and on the decomposer community (Bardgett and Wardle 2010). I acknowledge that the lack of relation can be misleading or caused by sampling bias. However, disentangling if different functional groups, temperature or humidity are more likely to mediate the litter accumulation could be important to understand litter decomposition and can also provide further insights on other soil related processes, such as water retention and nutrient cycling. These mechanisms can be important to improve climate change impacts predictions.

I suggest that a similar improvement in the understanding of other ecosystem processes and ecosystem services could be obtained by partitioning the communities into functional groups. The assumption that diversity can, and should, be reduced to the number of species or the number of years since divergence that underlies the use of "species richness" and "phylogenetic diversity" is probably misleading, as I found in chapter five. As previously indicated, finding the right scale and strategy to partition the communities and to characterize those partitions requires further research.

6.3 Final remarks

My thesis supports that understanding the deterministic-stochastic nature of community assembly is not only an interesting theoretical endeavor, but also an urgent task to better comprehend the ongoing global changes and their impacts. In my study I found that convergence, a simple deterministic rule, drives the most abundant species. Therefore, services closely related to biomass accumulation such as carbon sequestration should be more deterministic. Conversely, either purely stochastic mechanisms or divergence, a deterministic mechanism that creates multiple optimums, drive the non-dominant plants. Hence, services related to species diversity, such as pollination, should be more difficult to predict. I also found that dominant and non-dominant species are not randomly distributed in the phylogeny, and therefore the importance of deterministic and stochastic mechanisms should also differ among lineages.

The studies presented here covered a large range of scales, from landscape to local interactions, and some assumptions have been used across different scales and were needed to integrate the

document. In particular, I assumed that determinism at a large scale in the form of the predictability of species survival is a function of the determinism or stochasticity of local species interactions. Similarly, I assumed that large-scale provision of ecosystem processes and ecosystem services is a cumulative function of local processes and services. These two assumptions should be thoroughly tested in different biomes and with different processes and services.

Species contain different information acquired during millions of years of evolution and coexistence, and partitioning the community in ecologically meaningful units is a simple way to recover that information. This historically acquired information is useful to understand how the species respond to the environment as much as how the species change the environment. For instance, the changes in the environment caused by dominant plants change the relative importance of deterministic assembly rules from underdispersion to oversdispersion for the non-dominant species. Together, these changes in the environment modify for good or for worse the conditions for other species not randomly, affect ecosystem processes such as productivity and litter accumulation in different and sometimes contrasting ways and are also the basis for the ecosystem services that we benefit from.

References

- Aarssen, L. W., B. S. Schamp, and J. Pither. 2006. Why are there so many small plants? Implications for species coexistence. Journal of Ecology 94:569–580.
- Adler, P. B., E. W. Seabloom, E. T. Borer, H. Hillebrand, Y. Hautier, A. Hector, W. S. Harpole, L. R. O'Halloran, J. B. Grace, T. M. Anderson, J. D. Bakker, L. A. Biederman, C. S. Brown, Y. M. Buckley, L. B. Calabrese, C.-J. Chu, E. E. Cleland, S. L. Collins, K. L. Cottingham, M. J. Crawley, E. I. Damschen, K. F. Davies, N. M. DeCrappeo, P. A. Fay, J. Firn, P. Frater, E. I. Gasarch, D. S. Gruner, N. Hagenah, J. H. R. Lambers, H. Humphries, V. L. Jin, A. D. Kay, K. P. Kirkman, J. A. Klein, J. M. H. Knops, K. J. L. Pierre, J. G. Lambrinos, W. Li, A. S. MacDougall, R. L. McCulley, B. A. Melbourne, C. E. Mitchell, J. L. Moore, J. W. Morgan, B. Mortensen, J. L. Orrock, S. M. Prober, D. A. Pyke, A. C. Risch, M. Schuetz, M. D. Smith, C. J. Stevens, L. L. Sullivan, G. Wang, P. D. Wragg, J. P. Wright, and L. H. Yang. 2011. Productivity Is a Poor Predictor of Plant Species Richness. Science 333:1750–1753.
- Al-Mufti, M. M., C. L. Sydes, S. B. Furness, J. P. Grime, and S. R. Band. 1977. A Quantitative Analysis of Shoot Phenology and Dominance in Herbaceous Vegetation. Journal of Ecology 65:759–791.
- Anderson, M. J., T. O. Crist, J. M. Chase, M. Vellend, B. D. Inouye, A. L. Freestone, N. J. Sanders, H. V. Cornell, L. S. Comita, K. F. Davies, S. P. Harrison, N. J. B. Kraft, J. C. Stegen, and N. G. Swenson. 2011. Navigating the multiple meanings of β diversity: a roadmap for the practicing ecologist. Ecology Letters 14:19–28.
- Aranda, S. C., R. Gabriel, P. A. V. Borges, A. M. C. Santos, J. Hortal, A. Baselga, and J. M. Lobo. 2013. How do different dispersal modes shape the species–area relationship? Evidence for between-group coherence in the Macaronesian flora. Global Ecology and Biogeography 22:483–493.
- de Araújo, C. B., L. O. Marcondes-Machado, and G. C. Costa. 2014. The importance of biotic interactions in species distribution models: a test of the Eltonian noise hypothesis using parrots. Journal of Biogeography 41:513–523.
- Armenteras, D., F. Gast, and H. Villareal. 2003. Andean forest fragmentation and the representativeness of protected natural areas in the eastern Andes, Colombia. Biological Conservation 113:245–256.
- Bai, Y., J. Wu, Q. Pan, J. Huang, Q. Wang, F. Li, A. Buyantuyev, and X. Han. 2007. Positive linear relationship between productivity and diversity: evidence from the Eurasian Steppe. Journal of Applied Ecology 44:1023–1034.
- Bardgett, R. D., and D. A. Wardle. 2010. Aboveground-belowground linkages: biotic interactions, ecosystem processes, and global change. Oxford University Press Oxford.
- Baselga, A. 2010. Partitioning the turnover and nestedness components of beta diversity. Global Ecology and Biogeography 19:134–143.
- Baselga, A., D. Orme, S. Villeger, and J. D. B. and F. Leprieur. 2017. betapart: Partitioning Beta Diversity into Turnover and Nestedness Components.

- Bates, D., M. Maechler, B. Bolker, S. Walker, R. H. B. Christensen, H. Singmann, and B. Dai. 2014. lme4: Linear mixed-effects models using Eigen and S4.
- Bertuzzo, E., F. Carrara, L. Mari, F. Altermatt, I. Rodriguez-Iturbe, and A. Rinaldo. 2016. Geomorphic controls on elevational gradients of species richness. Proceedings of the National Academy of Sciences 113:1737–1742.
- Borer, E. T., W. S. Harpole, P. B. Adler, E. M. Lind, J. L. Orrock, E. W. Seabloom, and M. D. Smith. 2014a. Finding generality in ecology: a model for globally distributed experiments. Methods in Ecology and Evolution 5:65–73.
- Borer, E. T., E. W. Seabloom, D. S. Gruner, W. S. Harpole, H. Hillebrand, E. M. Lind, P. B. Adler, J. Alberti, T. M. Anderson, J. D. Bakker, L. Biederman, D. Blumenthal, C. S. Brown, L. A. Brudvig, Y. M. Buckley, M. Cadotte, C. Chu, E. E. Cleland, M. J. Crawley, P. Daleo, E. I. Damschen, K. F. Davies, N. M. DeCrappeo, G. Du, J. Firn, Y. Hautier, R. W. Heckman, A. Hector, J. HilleRisLambers, O. Iribarne, J. A. Klein, J. M. H. Knops, K. J. La Pierre, A. D. B. Leakey, W. Li, A. S. MacDougall, R. L. McCulley, B. A. Melbourne, C. E. Mitchell, J. L. Moore, B. Mortensen, L. R. O'Halloran, J. L. Orrock, J. Pascual, S. M. Prober, D. A. Pyke, A. C. Risch, M. Schuetz, M. D. Smith, C. J. Stevens, L. L. Sullivan, R. J. Williams, P. D. Wragg, J. P. Wright, and L. H. Yang. 2014b. Herbivores and nutrients control grassland plant diversity via light limitation. Nature 508:517–520.
- Bortolussi, N., E. Durand, M. Blum, and O. Francois. 2012. apTreeshape: Analyses of Phylogenetic Treeshape.
- Breheny, P., and W. Burchett. 2018. visreg: Visualization of Regression Models.
- Brooker, R. W., and R. M. Callaway. 2009. Facilitation in the conceptual melting pot. Journal of Ecology 97:1117–1120.
- Brooker, R. W., F. T. Maestre, R. M. Callaway, C. L. Lortie, L. A. Cavieres, G. Kunstler, P. Liancourt, K. TielböRger, J. M. J. Travis, F. Anthelme, C. Armas, L. Coll, E. Corcket, S. Delzon, E. Forey, Z. Kikvidze, J. Olofsson, F. Pugnaire, C. L. Quiroz, P. Saccone, K. Schiffers, M. Seifan, B. Touzard, and R. Michalet. 2008. Facilitation in plant communities: the past, the present, and the future. Journal of Ecology 96:18–34.
- Brown, J. H. 1984. On the relationship between abundance and distribution of species. American naturalist:255–279.
- Brown, J. H., J. F. Gillooly, A. P. Allen, V. M. Savage, and G. B. West. 2004. Toward a metabolic theory of ecology. Ecology 85:1771–1789.
- Brown, J. H., and A. Kodric-Brown. 1977. Turnover rates in insular biogeography: effect of immigration on extinction. Ecology 58:445–449.
- Buytaert, W., M. Vuille, A. Dewulf, R. Urrutia, A. Karmalkar, and R. Célleri. 2011. Uncertainties in climate change projections and regional downscaling: implications for water resources management. Hydrology and Earth System Science 14:1247–1258.
- Cadotte, M., C. H. Albert, and S. C. Walker. 2013. The ecology of differences: assessing community assembly with trait and evolutionary distances. Ecology Letters 16:1234–1244.

- Cadotte, M. W. 2013. Experimental evidence that evolutionarily diverse assemblages result in higher productivity. Proceedings of the National Academy of Sciences 110:8996–9000.
- Cadotte, M. W. 2015. Phylogenetic diversity and productivity: gauging interpretations from experiments that do not manipulate phylogenetic diversity. Functional Ecology 29:1603–1606.
- Cadotte, M. W. 2017. Functional traits explain ecosystem function through opposing mechanisms. Ecology Letters 20:989–996.
- Cadotte, M. W., and T. J. Davies. 2016. Phylogenies in ecology: a guide to concepts and methods. Princeton University Press.
- Cadotte, M. W., T. Jonathan Davies, and P. R. Peres-Neto. 2017. Why phylogenies do not always predict ecological differences. Ecological Monographs.
- Cadotte, M. W., and C. M. Tucker. 2017. Should Environmental Filtering be Abandoned? Trends in Ecology & Evolution 32:429–437.
- Cardinale, B. J., M. A. Palmer, and S. L. Collins. 2002. Species diversity enhances ecosystem functioning through interspecific facilitation. Nature 415:426–429.
- Cardinale, B. J., D. S. Srivastava, J. Emmett Duffy, J. P. Wright, A. L. Downing, M. Sankaran, and C. Jouseau. 2006. Effects of biodiversity on the functioning of trophic groups and ecosystems. Nature 443:989–992.
- Chai, Y., M. Yue, X. Liu, Y. Guo, M. Wang, J. Xu, C. Zhang, Y. Chen, L. Zhang, and R. Zhang. 2016. Patterns of taxonomic, phylogenetic diversity during a long-term succession of forest on the Loess Plateau, China: insights into assembly process. Scientific Reports 6:27087.
- Chalmandrier, L., T. Münkemüller, L. Gallien, F. de Bello, F. Mazel, S. Lavergne, and W. Thuiller. 2013. A family of null models to distinguish between environmental filtering and biotic interactions in functional diversity patterns. Journal of Vegetation Science 24:853–864.
- Chase, J. M. 2007. Drought mediates the importance of stochastic community assembly. Proceedings of the National Academy of Sciences 104:17430–17434.
- Chase, J. M. 2010. Stochastic Community Assembly Causes Higher Biodiversity in More Productive Environments. Science 328:1388–1391.
- Chase, J. M., and M. A. Leibold. 2002. Spatial scale dictates the productivity–biodiversity relationship. Nature 416:427–430.
- Chave, J., H. C. Muller-Landau, and S. A. Levin. 2002. Comparing classical community models: theoretical consequences for patterns of diversity. The American Naturalist 159:1–23.
- Chesson, P. 2000. Mechanisms of maintenance of species diversity. Annual review of Ecology and Systematics:343–366.
- Clark, J. S., M. Dietze, S. Chakraborty, P. K. Agarwal, I. Ibanez, S. LaDeau, and M. Wolosin. 2007. Resolving the biodiversity paradox. Ecology Letters 10:647–659.
- CMOS-SCMO BULLETIN. (n.d.). Highlights From Top Ten Weather Stories in 2015.

- Cornwell, W. K., and D. D. Ackerly. 2010. A link between plant traits and abundance: evidence from coastal California woody plants. Journal of Ecology 98:814–821.
- Csardi, G., and T. Nepusz. 2006. The igraph software package for complex network research. InterJournal Complex Systems:1695.
- De Deyn, G. B., J. H. C. Cornelissen, and R. D. Bardgett. 2008. Plant functional traits and soil carbon sequestration in contrasting biomes. Ecology Letters 11:516–531.
- Dengler, J. 2009. Which function describes the species-area relationship best? A review and empirical evaluation. Journal of Biogeography 36:728–744.
- Diamond, J. M., N. P. Ashmole, and P. E. Purves. 1989. The present, past and future of humancaused extinctions. Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences 325:469–477.
- Dirnböck, T., S. Dullinger, and G. Grabherr. 2003. A regional impact assessment of climate and land-use change on alpine vegetation. Journal of Biogeography 30:401–417.
- Dirnböck, T., F. Essl, and W. Rabitsch. 2011. Disproportional risk for habitat loss of highaltitude endemic species under climate change. Global Change Biology 17:990–996.
- Drake, J. A. 1991. Community-assembly mechanics and the structure of an experimental species ensemble. American Naturalist:1–26.
- Duffy, J. E., C. M. Godwin, and B. J. Cardinale. 2017. Biodiversity effects in the wild are common and as strong as key drivers of productivity. Nature 549:261–264.
- Dullinger, S., A. Gattringer, W. Thuiller, D. Moser, N. E. Zimmermann, A. Guisan, W. Willner, C. Plutzar, M. Leitner, T. Mang, M. Caccianiga, T. Dirnböck, S. Ertl, A. Fischer, J. Lenoir, J.-C. Svenning, A. Psomas, D. R. Schmatz, U. Silc, P. Vittoz, and K. Hülber. 2012. Extinction debt of high-mountain plants under twenty-first-century climate change. Nature Climate Change 2:619–622.
- Ellenberg, H. 1979. Man's influence on tropical mountain ecosystems in South America. The second Tansley lecture. Journal of Ecology 67:401–416.
- Ewers, R. M., and R. K. Didham. 2006. Confounding factors in the detection of species responses to habitat fragmentation. Biological Reviews 81:117–142.
- Fahrig, L. 2003. Effects of habitat fragmentation on biodiversity. Annual Review of Ecology, Evolution, and Systematics 34:487–515.
- Fahrig, L. 2013. Rethinking patch size and isolation effects: the habitat amount hypothesis. Journal of Biogeography 40:1649–1663.
- Fahrig, L. 2017. Ecological Responses to Habitat Fragmentation per se. Annual Reviews of Ecology, Evolution and Systematics 48:1–23.
- Faith, D. P. 1992. Conservation evaluation and phylogenetic diversity. Biological conservation 61:1–10.
- Farnsworth, E. J. 2007. Plant life history traits of rare versus frequent plant taxa of sandplains: Implications for research and management trials. Biological Conservation 136:44–52.
- Feeley, K. J., and M. R. Silman. 2010. Land-use and climate change effects on population size and extinction risk of Andean plants. Global Change Biology 16:3215–3222.

- Feeley, K. J., M. R. Silman, M. B. Bush, W. Farfan, K. G. Cabrera, Y. Malhi, P. Meir, N. S. Revilla, M. N. R. Quisiyupanqui, and S. Saatchi. 2011. Upslope migration of Andean trees. Journal of Biogeography 38:783–791.
- Fick, S. E., and R. J. Hijmans. 2017. WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. International Journal of Climatology 37:4302–4315.
- Figge, F. 2004. Bio-folio: applying portfolio theory to biodiversity. Biodiversity & Conservation 13:827–849.
- Fischlin, A., G. F. Midgley, J. Price, R. Leemans, B. Gopal, C. Turley, M. Rounsevell, P. Dube, J. Tarazona, and A. Velichko. 2007. Ecosystems, their properties, goods, and services. Climate Change 2007: Impacts, Adaptation and Vulnerability. Contribution of Working Group II to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. http://www.ipcc.ch/publications and data/ar4/wg2/en/ch4.html.
- Fisher, R. A., A. S. Corbet, and C. B. Williams. 1943. The Relation Between the Number of Species and the Number of Individuals in a Random Sample of an Animal Population. Journal of Animal Ecology 12:42–58.
- Franzén, M., O. Schweiger, and P.-E. Betzholtz. 2012. Species-area relationships are controlled by species traits. PLoS ONE 7:e37359.
- Freemark, K. E., and H. G. Merriam. 1986. Importance of area and habitat heterogeneity to bird assemblages in temperate forest fragments. Biological Conservation 36:115–141.
- Fukami, T. 2015. Historical Contingency in Community Assembly: Integrating Niches, Species Pools, and Priority Effects. Annual Review of Ecology, Evolution, and Systematics 46:1– 23.
- Garibaldi, L. A., L. G. Carvalheiro, B. E. Vaissiere, B. Gemmill-Herren, J. Hipolito, B. M. Freitas, H. T. Ngo, N. Azzu, A. Saez, J. Astrom, J. An, B. Blochtein, D. Buchori, F. J. C. Garcia, F. Oliveira da Silva, K. Devkota, M. d. F. Ribeiro, L. Freitas, M. C. Gaglianone, M. Goss, M. Irshad, M. Kasina, A. J. S. P. Filho, L. H. P. Kiill, P. Kwapong, G. N. Parra, C. Pires, V. Pires, R. S. Rawal, A. Rizali, A. M. Saraiva, R. Veldtman, B. F. Viana, S. Witter, and H. Zhang. 2016. Mutually beneficial pollinator diversity and crop yield outcomes in small and large farms. Science 351:388–391.
- Gause, G. F. 1932. Experimental studies on the struggle for existence i. Mixed population of two species of yeast. Journal of Experimental Biology 9:389–402.
- Gerhold, P., J. F. Cahill, M. Winter, I. V. Bartish, and A. Prinzing. 2015. Phylogenetic patterns are not proxies of community assembly mechanisms (they are far better). Functional Ecology 29:600–614.
- Gibson, D. J., J. S. Ely, and S. L. Collins. 1999. The core–satellite species hypothesis provides a theoretical basis for Grime's classification of dominant, subordinate, and transient species. Journal of Ecology 87:1064–1067.
- Gilbert, B., R. Turkington, and D. S. Srivastava. 2009. Dominant Species and Diversity: Linking Relative Abundance to Controls of Species Establishment. The American Naturalist 174:850–862.
- Gillman, L. N., and S. D. Wright. 2006. The influence of productivity on the species richness of plants: a critical assessment. Ecology 87:1234–1243.

- Gonzalez, A. 2000. Community relaxation in fragmented landscapes: the relation between species richness, area and age. Ecology Letters 3:441–448.
- Gould, S. J. 1990. Wonderful life: the Burgess Shale and the nature of history. WW Norton & Company.
- Grace, J. B., T. M. Anderson, H. Olff, and S. M. Scheiner. 2010. On the specification of structural equation models for ecological systems. Ecological Monographs 80:67–87.
- Grace, J. B., T. M. Anderson, E. W. Seabloom, E. T. Borer, P. B. Adler, W. S. Harpole, Y. Hautier, H. Hillebrand, E. M. Lind, M. Pärtel, J. D. Bakker, Y. M. Buckley, M. J. Crawley, E. I. Damschen, K. F. Davies, P. A. Fay, J. Firn, D. S. Gruner, A. Hector, J. M. H. Knops, A. S. MacDougall, B. A. Melbourne, J. W. Morgan, J. L. Orrock, S. M. Prober, and M. D. Smith. 2016. Integrative modelling reveals mechanisms linking productivity and plant species richness. Nature 529:390–393.
- Grace, J. B., T. Michael Anderson, M. D. Smith, E. Seabloom, S. J. Andelman, G. Meche, E. Weiher, L. K. Allain, H. Jutila, M. Sankaran, J. Knops, M. Ritchie, and M. R. Willig. 2007. Does species diversity limit productivity in natural grassland communities? Ecology Letters 10:680–689.
- Grime, J. P. 1974. Vegetation classification by reference to strategies. Nature 250:26–31.
- Grime, J. P. 1998. Benefits of plant diversity to ecosystems: immediate, filter and founder effects. Journal of Ecology 86:902–910.
- Guo, Z., X. Xiao, and D. Li. 2000. An assessment of ecosystem services: water flow regulation and hydroelectric power production. Ecological Applications 10:925–936.
- Hanski, I. 1982. Dynamics of Regional Distribution: The Core and Satellite Species Hypothesis. Oikos 38:210–221.
- Henderson, P. A., and A. E. Magurran. 2014. Direct evidence that density-dependent regulation underpins the temporal stability of abundant species in a diverse animal community. Proc. R. Soc. B 281:20141336.
- Hijmans, R. J., J. van Etten, M. Mattiuzzi, M. Sumner, J. A. Greenberg, O. P. Lamigueiro, A. Bevan, E. B. Racine, and A. Shortridge. 2015. raster: Geographic Data Analysis and Modeling.
- Holdridge, L. R. 1967. Life zone ecology. Costa Rica.
- Hubbell, S. P. 2001a. The unified neutral theory of biodiversity and biogeography. Princeton University Press, Princeton.
- Hubbell, S. P. 2001b. The unified neutral theory of biodiversity and biogeography. Princeton University Press, Princeton.
- Huston, M. A. 1994. Biological diversity: the coexistence of species on changing landscapes. Cambridge University Press, Cambridge, UK.
- Hutchinson, G. E. 1978. An introduction to population ecology.
- Hylander, K., and J. Ehrlén. 2013. The mechanisms causing extinction debts. Trends in Ecology & Evolution 28:341–346.

- Ibáñez, I., J. S. Clark, M. C. Dietze, K. Feeley, M. Hersh, S. LaDeau, A. McBride, N. E. Welch, and M. S. Wolosin. 2006. Predicting biodiversity change: outside the climate envelope, beyond the species-area curve. Ecology 87:1896–1906.
- Jin, L. 2015, November. A Spatio-temporal Phylogenetic Approach to Community Ecology. Thesis.
- Jones, C. G., J. H. Lawton, and M. Shachak. 1997. Positive and Negative Effects of Organisms as Physical Ecosystem Engineers. Ecology 78:1946–1957.
- Kadmon, R., and O. Allouche. 2007. Integrating the effects of area, isolation, and habitat heterogeneity on species diversity: A unification of island biogeography and niche theory. The American Naturalist 170:443–454.
- Kembel, S. W., D. D. Ackerly, S. P. Blomberg, W. K. Cornwell, P. D. Cowan, M. R. Helmus, H. Morlon, and C. O. Webb. 2014. picante: R tools for integrating phylogenies and ecology.
- Kissling, W. D., C. F. Dormann, J. Groeneveld, T. Hickler, I. Kühn, G. J. McInerny, J. M. Montoya, C. Römermann, K. Schiffers, and F. M. Schurr. 2011. Towards novel approaches to modelling biotic interactions in multispecies assemblages at large spatial extents. Journal of Biogeography 39:2163–2178.
- Körner, C., and M. Ohsawa. 2005. Mountain systems. Pages 681–716 in R. Hassan and R. Scholes, editors. Millennium ecosystem assessment. Ecosystems and human well-being: current state and trends. Island Press, Washington DC.
- Kraft, N. J. B., P. B. Adler, O. Godoy, E. C. James, S. Fuller, and J. M. Levine. 2015a. Community assembly, coexistence and the environmental filtering metaphor. Functional Ecology 29:592–599.
- Kraft, N. J. B., O. Godoy, and J. M. Levine. 2015b. Plant functional traits and the multidimensional nature of species coexistence. Proceedings of the National Academy of Sciences 112:797–802.
- Laland, K., B. Matthews, and M. W. Feldman. 2016. An introduction to niche construction theory. Evolutionary Ecology 30:191–202.
- Laughlin, D. C. 2014. Applying trait-based models to achieve functional targets for theorydriven ecological restoration. Ecology Letters 17:771–784.
- Laughlin, D. C., C. Joshi, P. M. van Bodegom, Z. A. Bastow, and P. Z. Fulé. 2012. A predictive model of community assembly that incorporates intraspecific trait variation. Ecology Letters 15:1291–1299.
- Lavorel, S., and E. Garnier. 2002. Predicting changes in community composition and ecosystem functioning from plant traits: revisiting the Holy Grail. Functional Ecology 16:545–556.
- Lavorel, S., K. Grigulis, P. Lamarque, M. Colace, Garden, J. Girel, G. Pellet, and R. Douzet. 2010. Using plant functional traits to understand the landscape distribution of multiple ecosystem services. Journal of Ecology 99:135–147.
- Lawton, J. H. 1999. Are There General Laws in Ecology? Oikos 84:177–192.
- Lefcheck, J. 2016. piecewiseSEM: Piecewise Structural Equation Modeling.

- Lennon, J. J., C. M. Beale, C. L. Reid, M. Kent, and R. J. Pakeman. 2011. Are richness patterns of common and rare species equally well explained by environmental variables? Ecography 34:529–539.
- Leprieur, F., C. Albouy, J. De Bortoli, P. F. Cowman, D. R. Bellwood, and D. Mouillot. 2012. Quantifying Phylogenetic Beta Diversity: Distinguishing between 'True' Turnover of Lineages and Phylogenetic Diversity Gradients. PLoS ONE 7:e42760.
- Leroux, S. J., C. H. Albert, A.-S. Lafuite, B. Rayfield, S. Wang, and D. Gravel. 2017. Structural uncertainty in models projecting the consequences of habitat loss and fragmentation on biodiversity. Ecography 40:36–47.
- Levins, R. 1969. Some demographic and genetic consequences of environmental heterogeneity for biological control. Bulletin of the Entomological Society of America 15:237–240.
- Lewis, O. T. 2006. Climate change, species–area curves and the extinction crisis. Philosophical Transactions of the Royal Society B: Biological Sciences 361:163–171.
- Li, S., T. Guo, M. W. Cadotte, Y. Chen, J. Kuang, Z. Hua, Y. Zeng, Y. Song, Z. Liu, W. Shu, and J. Li. 2015. Contrasting effects of phylogenetic relatedness on plant invader success in experimental grassland communities. Journal of Applied Ecology 52:89–99.
- Li, Y., and B. Shipley. 2018. Community divergence and convergence along experimental gradients of stress and disturbance. Ecology 99:775–781.
- Loehle, C., and B.-L. Li. 1996. Habitat destruction and the extinction debt revisited. Ecological Applications 6:784–789.
- Lomolino, M. V. 1996. Investigating causality of nestedness of insular communities: selective immigrations or extinctions? Journal of Biogeography 23:699–703.
- Lomolino, M. V. 2000. Ecology's most general, yet protean pattern: the species-area relationship. Journal of Biogeography 27:17–26.
- Loreau, M., and A. Hector. 2001. Partitioning selection and complementarity in biodiversity experiments. Nature 412:72–76.
- Lortie, C. J., and R. M. Callaway. 2006. Re-analysis of meta-analysis: support for the stressgradient hypothesis. Journal of Ecology 94:7–16.
- Lowe, W. H., and M. A. McPeek. 2014. Is dispersal neutral? Trends in Ecology & Evolution 29:444–450.
- MacArthur, R. 1955. Fluctuations of animal populations and a measure of community stability. ecology 36:533–536.
- MacArthur, R. 1960. On the relative abundance of species. American Naturalist:25–36.
- MacArthur, R. H., and E. O. Wilson. 1967. The theory of island biogeography. Princeton University Press.
- MacArthur, R., and R. Levins. 1964. Competition, habitat selection, and character displacement in a patchy environment. Proceedings of the National Academy of Sciences of the United States of America 51:1207.
- MacArthur, R., and R. Levins. 1967. The limiting similarity, convergence, and divergence of coexisting species. American naturalist:377–385.

Magurran, A. E. 2013. Measuring biological diversity. John Wiley & Sons.

- Magurran, A. E., and P. A. Henderson. 2003. Explaining the excess of rare species in natural species abundance distributions. Nature 422:714–716.
- Maire, V., N. Gross, L. Börger, R. Proulx, C. Wirth, L. da Silveira Pontes, J. Soussana, and F. Louault. 2012. Habitat filtering and niche differentiation jointly explain species relative abundance within grassland communities along fertility and disturbance gradients. New Phytologist 196:497–509.
- Malcolm, J. R., C. Liu, R. P. Neilson, L. Hansen, and L. Hannah. 2006. Global warming and extinctions of endemic species from biodiversity hotspots. Conservation Biology 20:538– 548.
- Manning, P., F. van der Plas, S. Soliveres, E. Allan, F. T. Maestre, G. Mace, M. J. Whittingham, and M. Fischer. 2018. Redefining ecosystem multifunctionality. Nature Ecology & Evolution 2:427–436.
- Mariotte, P. 2014. Do subordinate species punch above their weight? Evidence from above- and below-ground. New Phytologist 203:16–21.
- Matthews, T. J., K. A. Triantis, F. Rigal, M. K. Borregaard, F. Guilhaumon, and R. J. Whittaker. 2016. Island species–area relationships and species accumulation curves are not equivalent: an analysis of habitat island datasets. Global Ecology and Biogeography 25:607–618.
- Mayfield, M. M., and J. M. Levine. 2010. Opposing effects of competitive exclusion on the phylogenetic structure of communities. Ecology Letters 13:1085–1093.
- McGarigal, K., and B. Marks. 1995. FRAGSTATS: Spatial pattern analysis program. Forest Science Dept., Oregon State University.
- McGill, B. J., R. S. Etienne, J. S. Gray, D. Alonso, M. J. Anderson, H. K. Benecha, M. Dornelas, B. J. Enquist, J. L. Green, F. He, and others. 2007. Species abundance distributions: moving beyond single prediction theories to integration within an ecological framework. Ecology Letters 10:995–1015.
- McGlinn, D. J., X. Xiao, F. May, N. J. Gotelli, S. Blowes, T. Knight, O. Purschke, J. Chase, and B. McGill. 2018. MoB (Measurement of Biodiversity): a method to separate the scaledependent effects of species abundance distribution, density, and aggregation on diversity change. bioRxiv.
- McIntire, E. J. B., and A. Fajardo. 2014. Facilitation as a ubiquitous driver of biodiversity. New Phytologist 201:403–416.
- Meehl, G. A., C. Covey, K. E. Taylor, T. Delworth, R. J. Stouffer, M. Latif, B. McAvaney, and J. F. B. Mitchell. 2007. The WCRP CMIP3 Multimodel Dataset: A new era in climate change research. Bulletin of the American Meteorological Society 88:1383–1394.
- Millennium Ecosystem Assessment. 2005. Ecosystems and human well-being: synthesis. Island Press, Washington, DC.
- Mittelbach, G. G., C. F. Steiner, S. M. Scheiner, K. L. Gross, H. L. Reynolds, R. B. Waide, M. R. Willig, S. I. Dodson, and L. Gough. 2001. What is the observed relationship between species richness and productivity? Ecology 82:2381–2396.

- Mooers, A. O., and S. B. Heard. 1997. Inferring evolutionary process from phylogenetic tree shape. The quarterly review of Biology 72:31–54.
- Myers, N., R. A. Mittermeier, C. G. Mittermeier, G. A. B. da Fonseca, and J. Kent. 2000. Biodiversity hotspots for conservation priorities. Nature 403:853–858.
- Norden, N., V. Boukili, A. Chao, K. H. Ma, S. G. Letcher, and R. L. Chazdon. 2017. Opposing mechanisms affect taxonomic convergence between tree assemblages during tropical forest succession. Ecology Letters 20:1448–1458.
- Noyes, R. D., and L. H. Rieseberg. 1999. ITS sequence data support a single origin for North American Astereae (Asteraceae) and reflect deep geographic divisions in Aster s.l. American Journal of Botany 86:398–412.
- Odling-Smee, J., D. H. Erwin, E. P. Palkovacs, M. W. Feldman, and K. N. Laland. 2013. Niche Construction Theory: A Practical Guide for Ecologists. The Quarterly Review of Biology 88:3–28.
- O'Halloran, L. R., E. T. Borer, E. W. Seabloom, A. S. MacDougall, E. E. Cleland, R. L. McCulley, S. Hobbie, W. S. Harpole, N. M. DeCrappeo, C. Chu, J. D. Bakker, K. F. Davies, G. Du, J. Firn, N. Hagenah, K. S. Hofmockel, J. M. H. Knops, W. Li, B. A. Melbourne, J. W. Morgan, J. L. Orrock, S. M. Prober, and C. J. Stevens. 2013. Regional Contingencies in the Relationship between Aboveground Biomass and Litter in the World's Grasslands. PLoS ONE 8:e54988.
- Oksanen, J. 1996. Is the Humped Relationship between species Richness and Biomass an Artefact due to Plot Size? Journal of Ecology 84:293–295.
- Oksanen, J., F. G. Blanchet, M. Friendly, R. Kindt, P. Legendre, D. McGlinn, P. R. Minchin, R. B. O'Hara, G. L. Simpson, P. Solymos, M. H. H. Stevens, E. Szoecs, and H. Wagner. 2017. vegan: Community Ecology Package.
- Olson, D. M., E. Dinerstein, E. D. Wikramanayake, N. D. Burgess, G. V. N. Powell, E. C. Underwood, J. A. D'amico, I. Itoua, H. E. Strand, J. C. Morrison, C. J. Loucks, T. F. Allnutt, T. H. Ricketts, Y. Kura, J. F. Lamoreux, W. W. Wettengel, P. Hedao, and K. R. Kassem. 2001. Terrestrial ecoregions of the world: a new map of life on Earth. BioScience 51:933–938.
- Paine, R. T. 1969. A Note on Trophic Complexity and Community Stability. The American Naturalist 103:91–93.
- Pandit, S. N., J. Kolasa, and K. Cottenie. 2009. Contrasts between habitat generalists and specialists: an empirical extension to the basic metacommunity framework. Ecology 90:2253–2262.
- Paradis, E., S. Blomberg, B. Bolker, J. Brown, J. Claude, H. S. Cuong, R. Desper, G. Didier, B. Durand, J. Dutheil, R. J. Ewing, O. Gascuel, C. Heibl, A. Ives, B. Jones, F. Krah, D. Lawson, V. Lefort, P. Legendre, J. Lemon, R. McCloskey, J. Nylander, R. Opgen-Rhein, A.-A. Popescu, M. Royer-Carenzi, K. Schliep, K. Strimmer, and D. de Vienne. 2017. ape: Analyses of Phylogenetics and Evolution.
- Parmesan, C. 2006. Ecological and evolutionary responses to recent climate change. Annual Review of Ecology, Evolution, and Systematics 37:637–669.

- Passy, S. I. 2017. Framework for community functioning: synthesis of stress gradient and resource partitioning concepts. PeerJ 5:e3885.
- Patterson, B. D., and W. Atmar. 1986. Nested subsets and the structure of insular mammalian faunas and archipelagos. Biological Journal of the Linnean Society 28:65–82.
- Pearson, R. G., and T. P. Dawson. 2003. Predicting the impacts of climate change on the distribution of species: are bioclimate envelope models useful? Global Ecology and Biogeography 12:361–371.
- Petchey, O. L. 2004. On the statistical significance of functional diversity effects. Functional Ecology 18:297–303.
- Phillips, D. 2018, January 7. Canada's Top Ten Weather Stories 2017.
- Pimm, S. L., and R. A. Askins. 1995. Forest losses predict bird extinctions in eastern North America. Proceedings of the National Academy of Sciences 92:9343–9347.
- Proença, V., and H. M. Pereira. 2013. Species–area models to assess biodiversity change in multi-habitat landscapes: The importance of species habitat affinity. Basic and Applied Ecology 14:102–114.
- Pybus, O. G., and P. H. Harvey. 2000. Testing macro–evolutionary models using incomplete molecular phylogenies. Proceedings of the Royal Society of London B: Biological Sciences 267:2267–2272.
- R Core Team. 2016. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.
- R Core Team. 2017. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.
- Rabinowitz, D. 1981a. Seven forms of rarity. Page *in* H. Synge, editor. The biological aspects of rare plant conservation. Wiley-Blackwell, New York, NY.
- Rabinowitz, D. 1981b. Seven forms of rarity. Pages 205–217 in H. Synge, editor. The biological aspects of rare plant conservation. Wiley-Blackwell, New York, NY.
- Rabinowitz, D., J. K. Rapp, and P. M. Dixon. 1984. Competitive Abilities of Sparse Grass Species: Means of Persistence or Cause of Abundance. Ecology 65:1144–1154.
- Ramirez-Villegas, J., F. Cuesta, C. Devenish, M. Peralvo, A. Jarvis, and C. A. Arnillas. 2014. Using species distributions models for designing conservation strategies of Tropical Andean biodiversity under climate change. Journal for Nature Conservation 22:391–404.
- Rehm, E. M., and K. J. Feeley. 2015. The inability of tropical cloud forest species to invade grasslands above treeline during climate change: potential explanations and consequences. Ecography 38:1167–1175.
- Reznick, D., M. J. Bryant, and F. Bashey. 2002. r- and K-selection revisited The role of population regulation in life-history evolution. Ecology 83:1509–1520.
- Ricotta, C., S. Godefroid, and L. Celesti-Grapow. 2008. Common species have lower taxonomic diversity Evidence from the urban floras of Brussels and Rome. Diversity and Distributions 14:530–537.

- Rosenzweig, M. 1995. Species diversity in space and time. Cambridge University Press, Cambridge; New York.
- Rosseel, Y., D. Oberski, J. Byrnes, L. Vanbrabant, V. Savalei, E. Merkle, M. Hallquist, M. Rhemtulla, M. Katsikatsou, M. Barendse, M. Chow, and T. Jorgensen. 2017. lavaan: Latent Variable Analysis.
- Rybicki, J., and I. Hanski. 2013. Species–area relationships and extinctions caused by habitat loss and fragmentation. Ecology Letters 16:27–38.
- Saura, S., and L. Pascual-Hortal. 2007. A new habitat availability index to integrate connectivity in landscape conservation planning: Comparison with existing indices and application to a case study. Landscape and Urban Planning 83:91–103.
- Scott-Phillips, T. C., K. N. Laland, D. M. Shuker, T. E. Dickins, and S. A. West. 2013. The Niche Construction Perspective: A Critical Appraisal. Evolution 68:1231–1243.
- Shipley, B. 2016. Cause and Correlation in Biology: A User's Guide to Path Analysis, Structural Equations and Causal Inference with R. Cambridge University Press.
- Simberloff, D. 2000. What do we really know about habitat fragmentation? The Texas journal of science 52:5–22.
- Simberloff, D., and B. V. Holle. 1999. Positive Interactions of Nonindigenous Species: Invasional Meltdown? Biological Invasions 1:21–32.
- Simberloff, D. S., and L. G. Abele. 1976. Island biogeography theory and conservation practice. Science 191:285–286.
- Siqueira, T., L. M. Bini, F. O. Roque, S. R. Marques Couceiro, S. Trivinho-Strixino, and K. Cottenie. 2012. Common and rare species respond to similar niche processes in macroinvertebrate metacommunities. Ecography 35:183–192.
- Soberón, J. 2007. Grinnellian and Eltonian niches and geographic distributions of species. Ecology Letters 10:1115–1123.
- Sobral, F. L., A. C. Lees, and M. V. Cianciaruso. 2016. Introductions do not compensate for functional and phylogenetic losses following extinctions in insular bird assemblages. Ecology Letters 19:1091–1100.
- Spasojevic, M. J., and K. N. Suding. 2012. Inferring community assembly mechanisms from functional diversity patterns: the importance of multiple assembly processes. Journal of Ecology 100:652–661.
- Steinmann, K., H. P. Linder, and N. E. Zimmermann. 2009. Modelling plant species richness using functional groups. Ecological Modelling 220:962–967.
- Stern, N. 2007. The economics of climate change: the Stern review. cambridge University press.
- Storch, D., E. Bohdalková, J. Okie, and D. Gravel. 2018. The more-individuals hypothesis revisited: the role of community abundance in species richness regulation and the productivity–diversity relationship. Ecology Letters 21:920–937.

The Plant List. 2013. The Plant List. http://www.theplantlist.org/.

- Thomas, C. D., A. Cameron, R. E. Green, M. Bakkenes, L. J. Beaumont, Y. C. Collingham, B. F. N. Erasmus, M. Ferreira de Siqueira, A. Grainger, L. Hannah, and others. 2004. Extinction risk from climate change. Nature 427:145–148.
- Thuiller, W., C. Albert, M. B. Araújo, P. M. Berry, M. Cabeza, A. Guisan, T. Hickler, G. F. Midgley, J. Paterson, and F. M. Schurr. 2008. Predicting global change impacts on plant species' distributions: Future challenges. Perspectives in Plant Ecology, Evolution and Systematics 9:137–152.
- Tilman, D. 1977. Resource competition between plankton algae: an experimental and theoretical approach. Ecology 58:338–348.
- Tilman, D. 1980. Resources: A graphical-mechanistic approach to competition and predation. The American Naturalist 116:362–393.
- Tilman, D. 2004. Niche tradeoffs, neutrality, and community structure: A stochastic theory of resource competition, invasion, and community assembly. Proceedings of the National Academy of Sciences of the United States of America 101:10854–10861.
- Tilman, D., F. Isbell, and J. M. Cowles. 2014. Biodiversity and Ecosystem Functioning. Annual Review of Ecology, Evolution, and Systematics 45:471–493.
- Tilman, D., J. Knops, D. Wedin, P. Reich, M. Ritchie, and E. Siemann. 1997. The Influence of Functional Diversity and Composition on Ecosystem Processes. Science 277:1300–1302.
- Tilman, D., R. M. May, C. L. Lehman, and M. A. Nowak. 1994. Habitat destruction and the extinction debt. Nature 371:65–66.
- Tjørve, E. 2002. Habitat size and number in multi-habitat landscapes: a model approach based on species-area curves. Ecography 25:17–24.
- Tovar, C., C. A. Arnillas, F. Cuesta, and W. Buytaert. 2013. Diverging responses of tropical Andean biomes under future climate conditions. PLoS ONE 8:e63634.
- Tovar, C., C. A. Arnillas, M. Peralvo, and G. Galindo. 2011. Regional Scenarios of Biodiversity State in the Tropical Andes. Pages 265–285 Land Use, Climate Change and Biodiversity Modeling: Perspectives and Applications. IGI Global, Hershey.
- Travis, J. M. J. 2003. Climate change and habitat destruction: a deadly anthropogenic cocktail. Proceedings of the Royal Society of London. Series B: Biological Sciences 270:467–473.
- Triantis, K. A., F. Guilhaumon, and R. J. Whittaker. 2012. The island species-area relationship: biology and statistics: The island species-area relationship. Journal of Biogeography 39:215–231.
- Tucker, C. M., L. G. Shoemaker, K. F. Davies, D. R. Nemergut, and B. A. Melbourne. 2016. Differentiating between niche and neutral assembly in metacommunities using null models of β-diversity. Oikos 125:778–789.
- Tylianakis, J. M., R. K. Didham, J. Bascompte, and D. Wardle. 2008. Global change and species interactions in terrestrial ecosystems. Ecology Letters 11:1351–1363.
- Umaña, M. N., C. Zhang, M. Cao, L. Lin, and N. G. Swenson. 2017. A core-transient framework for trait-based community ecology: an example from a tropical tree seedling community. Ecology Letters 20:619–628.

- Valiente-Banuet, A., and M. Verdú. 2007. Facilitation can increase the phylogenetic diversity of plant communities. Ecology Letters 10:1029–1036.
- Vellend, M., D. S. Srivastava, K. M. Anderson, C. D. Brown, J. E. Jankowski, E. J. Kleynhans, N. J. B. Kraft, A. D. Letaw, A. A. M. Macdonald, J. E. Maclean, I. H. Myers-Smith, A. R. Norris, and X. Xue. 2014. Assessing the relative importance of neutral stochasticity in ecological communities. Oikos 123:1420–1430.
- Venail, P., K. Gross, T. H. Oakley, A. Narwani, E. Allan, P. Flombaum, F. Isbell, J. Joshi, P. B. Reich, D. Tilman, J. van Ruijven, and B. J. Cardinale. 2015. Species richness, but not phylogenetic diversity, influences community biomass production and temporal stability in a re-examination of 16 grassland biodiversity studies. Functional Ecology 29:615–626.
- Verburg, P. H., and A. Veldkamp. 2001. The role of spatially explicit models in land-use change research: a case study for cropping patterns in China. Agriculture, ecosystems & environment 85:177–190.
- Vermeij, G. J., and R. K. Grosberg. 2018. Rarity and persistence. Ecology Letters 21:3-8.
- Wang, Z., J. Fang, Z. Tang, and L. Shi. 2012. Geographical patterns in the beta diversity of China's woody plants: the influence of space, environment and range size. Ecography 35:1092–1102.
- Watts, K., A. E. Eycott, P. Handley, D. Ray, J. W. Humphrey, and C. P. Quine. 2010. Targeting and evaluating biodiversity conservation action within fragmented landscapes: an approach based on generic focal species and least-cost networks. Landscape Ecology 25:1305–1318.
- Webb, C. O. 2000. Exploring the phylogenetic structure of ecological communities: an example for rain forest trees. The American Naturalist 156:145–155.
- Webb, C. O., D. D. Ackerly, M. A. McPeek, and M. J. Donoghue. 2002. Phylogenies and community ecology. Annual Review of Ecology and Systematics 33:475–505.
- Weiher, E., D. Freund, T. Bunton, A. Stefanski, T. Lee, and S. Bentivenga. 2011. Advances, challenges and a developing synthesis of ecological community assembly theory.
 Philosophical Transactions of the Royal Society B: Biological Sciences 366:2403–2413.
- Weiher, E., and P. A. Keddy. 1995. Assembly rules, null models, and trait dispersion: new questions from old patterns. Oikos 74:159–164.
- White, E. P., and A. H. Hurlbert. 2010. The Combined Influence of the Local Environment and Regional Enrichment on Bird Species Richness. The American Naturalist 175:E35–E43.
- Whittaker, R. J. 2010. Meta-analyses and mega-mistakes: calling time on meta-analysis of the species richness—productivity relationship. Ecology 91:2522–2533.
- Whittaker, R. J., and K. A. Triantis. 2012. The species–area relationship: an exploration of that 'most general, yet protean pattern.' Journal of Biogeography 39:623–626.
- Wickham, H. 2009. ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York.
- Wisz, M. S., J. Pottier, W. D. Kissling, L. Pellissier, J. Lenoir, C. F. Damgaard, C. F. Dormann,
 M. C. Forchhammer, J.-A. Grytnes, A. Guisan, R. K. Heikkinen, T. T. Høye, I. Kühn, M. Luoto, L. Maiorano, M.-C. Nilsson, S. Normand, E. Öckinger, N. M. Schmidt, M. Termansen, A. Timmermann, D. A. Wardle, P. Aastrup, and J.-C. Svenning. 2013. The

role of biotic interactions in shaping distributions and realised assemblages of species: implications for species distribution modelling. Biological Reviews 88:15–30.

- Wright, D. H., B. D. Patterson, G. M. Mikkelson, A. Cutler, and W. Atmar. 1997. A comparative analysis of nested subset patterns of species composition. Oecologia 113:1–20.
- Wright, J. P., S. Naeem, A. Hector, C. Lehman, P. B. Reich, B. Schmid, and D. Tilman. 2006. Conventional functional classification schemes underestimate the relationship with ecosystem functioning: Functional groups and ecosystem function. Ecology Letters 9:111–120.
- Young, K. R. 2009. Andean land use and biodiversity: humanized landscapes in a time of change. Annals of the Missouri Botanical Garden 96:492–507.
- Zanne, A. E., D. C. Tank, W. K. Cornwell, J. M. Eastman, S. A. Smith, R. G. FitzJohn, D. J. McGlinn, B. C. O'Meara, A. T. Moles, P. B. Reich, D. L. Royer, D. E. Soltis, P. F. Stevens, M. Westoby, I. J. Wright, L. Aarssen, R. I. Bertin, A. Calaminus, R. Govaerts, F. Hemmings, M. R. Leishman, J. Oleksyn, P. S. Soltis, N. G. Swenson, L. Warman, and J. M. Beaulieu. 2014. Three keys to the radiation of angiosperms into freezing environments. Nature 506:89–92.
- Zimmermann, M., P. Meir, M. Bird, Y. Malhi, and A. Ccahuana. 2009. Litter contribution to diurnal and annual soil respiration in a tropical montane cloud forest. Soil Biology and Biochemistry 41:1338–1340.

Appendices

A Appendices of chapter 2

Data and code descriptions available in FigShare (doi: 10.6084/m9.figshare.5466331)

A.1 Supplementary table

Table S1 Global climate models (GCMs) used in the study for each gas emission scenario

Institution	CMIP3 I.D.	Gas emission scenarios	
		A2	A1B
Bjerknes Centre for Climate Research, Norway	BCCR-BCM2.0	Х	Х
CSIRO Atmospheric Research, Australia	CSIRO-MK3.0	х	х
CSIRO Atmospheric Research, Australia	CSIRO-MK3.5	х	х
Institute for Numerical Mathematics, Russia	INM-CM3.0	х	Х
Center for Climate System Research (The University of Tokyo), National Institute for Environmental Studies, and Frontier Research Center for Global Change (JAMSTEC), Japan	MIROC3.2 (medres)	x	x
Center for Climate System Research (The University of Tokyo), National Institute for Environmental Studies, and Frontier Research Center for Global Change (JAMSTEC), Japan	MIROC3.2 (hires)		х
National Center for Atmospheric Research, USA	CCSM3	х	х
US Dept. of Commerce / NOAA / Geophysical Fluid Dynamics Laboratory, USA	GFDL-CM2.0	х	Х
US Dept. of Commerce / NOAA / Geophysical Fluid Dynamics Laboratory, USA	GFDL-CM2.1	Х	Х
NASA / Goddard Institute for Space Studies	GISS-AOM		Х





Figure S1 Effect of individual dispersal and tolerance to human activities on the species

perception of the landscape.

In the top, the two maps represent the area as if human activities do not exist (left) and with human activities (right). If species are tolerant to human activities, their relevant space is the potential habitat. Intolerant species avoid transformed areas and will perceive their habitat as the remnant habitat only. Remnant habitats will always have less or equal total area than potential habitats, but the number of fragments can either increase (as in the example) or decrease. Groups of fragments can act as an individual patches according the dispersal capability of the individuals in the LEM-community. Therefore, the dispersal capability will define the size and number of patches in each of the habitats. Unrestricted dispersal implies that individuals move freely among fragments, and the habitat respond as a single patch. Restricted dispersal implies individuals are able to move freely in a non-contiguous patch (black solid arrows), but only partially among patches (gray dashed arrows). Therefore, fragments linked with black solid arrows

form discontinuous patches and their area should be pulled together before using the ISAR to estimate the species richness, while fragments linked with gray dashed arrows conform independent patches. Meta-population dynamics, characterized by local extinctions and recolonizations, appear among the independent patches. Isolation happens when no dispersal is possible at all, preventing any recolonization event so that the boundaries of each patch can be represented as barriers. No meta-population dynamic occurs at unrestricted dispersal or isolated patches.



Figure S2 From fragments to patches to habitat richness.

a) First, individual dispersal capability and tolerance to human impact of the species inhabiting the green habitat (here shown only for tolerant species) define the habitat configuration, i.e. patch sizes and distances between them. In this example, we represent four levels of connectivity, each representing a different habitat configuration, as perceived by the species. Dark arrows connect fragments that act as a single patch (labelled A-D). Null dispersion (dashed borders in the patches) inhibits recolonization. b) Second, if the area of the patch controls the number of species in a patch, then the number of species expected for each patch can be estimated using an ISAR. c) Third, the identity of the species (here numbers 1-14) depends on the predictability of species survival and the recolonization odds. Recolonization odds are a function of the individual dispersal capability so it is represented on the vertical axis. The predictability of species survival is represented here with two extremes: Always the same species survive and always a different set of species survives. When dispersal is unrestricted the full area (all fragments) acts as a single patch, and species turnover is meaningless. We name this a "unitary community" and in this case the total habitat richness depends only on the species-area relationship. If there is more than one patch, the set of species that survives is always the same and recolonization is possible, then there is no species turnover and we observe a species nested pattern among patches; we call this case a "nested community". However, if there is more than one patch, species can recolonize them but each patch has always a different set of species, the species turnover is total. We call this case a "disjoint community". Similarly, if the patches are isolated, the long-term expectation is that species turnover will be total and a disjoint community forms, even when in the short-term the species extinction can be predictable. If survival is not predictable and dispersal is restricted, the expected community is intermediate (in terms of total richness) between a nested and a disjoint community.



Figure S3 Map of the study area, showing the Andean biomes and the area of the lowland biomes that were modelled but not included in the analysis.

The potential map (a) represents the expected distribution of biomes if there were not land use transformation. The remnant map (b) shows the distribution of the biomes excluding human transformed areas. The remaining areas can have extensive usage by humans living in their surroundings.



Figure S4 Estimated exponent (z-value) of power-law expressions that represent the number of species per unit of area (from Matthews et al. 2016) as a function of the ratio between the largest and the smallest patch in the study area.

The blue and green vertical lines represent the ratio of the largest to smallest patch in different habitats in this study and the red dashed line represents the 0.25 exponent, often reported in several species-area relationship studies. Notice how the z-values range between 0.1 and 0.5 for the range of habitats in our study area, and converge towards 0.25 when the ratio increases.



Figure S5 Patch-size profile for current potential (black lines) and remnant biomes (gray

lines), assuming null dispersal.

Patch size profiles show the cumulative area when the patches are sorted from the largest (order 1) to the smallest. Dots show the threshold between patches smaller and larger than 10 and 1 km². Among these habitats, paramo (a) has the highest area loss caused by human impact, affecting particularly the largest patch (parallel lines, see Figure 3 in the main document). In the humid puna (b), the largest patch has been reduced by human activities, and the slightly convergent profiles suggest that a section of the largest patch has been lost and another fragmented. Glacier and cryoturbated areas (c) show no impact of land use change (black line is covered by gray one). In contrast, seasonal forest (d) shows a reduction in the area of the second and other mid-size patches, leaving the largest patch almost unchanged. Finally, prepuna (e) has suffered a small reduction in the three largest patches, and is also showing a degree of fragmentation.



Figure S6 Patch size profile for projected biomes under different climate models for

potential and remnant areas, assuming null dispersal.

Black lines represent present patch size profile and gray lines different future climate models. Paramo (a) and glacier and cryoturbated areas (b) are projected to be the most threatened biomes by future climate change, with strong reductions in total habitat, largest patch size, mid-size patches and even in the amount of patches. (e) Humid puna shows a low overall uncertainty (most models predict a decrease in the area), but montane shrublands (c), and xeric puna (d) present high uncertainty (half the lines above, half below the present profile).



Figure S7 Sensitivity analysis showing the proportion of variance explained by different variables in our richness calculation and how this variance changes with different exponents of the power-law expression (z-value).

Assuming that (a) the proportion of new species decreases with each new patch $(p_i = p^{i-1})$ and (b) the largest patch has all new species $(p_1=1)$ and any other patch has the same proportion $(p_{i>1}=p)$. The vertical dashed and dotted lines represent z=0.25 and 0.4, respectively, to represent the most common range of observed z-values (Figure S4). In that range, dispersal, turnover and their interaction explain most of the variability in most of the cases. The order of the variables is sensitive to the change in the probability of finding a new species (p_i) , with dispersal increasing its importance when every patch has a similar proportion of new species.

A.3 ISAR curves in nature and models: the prevalence of powerlaw

The original studies that characterized the ISARs often focus on the shape of the function and the parameters that describes it, without explicitly discussing the spatial arrangement of the patches (e.g. Connor and McCoy 1979, Pimm and Askins 1995, Rosenzweig 1995 and references there). Therefore, although ISAR curves have been widely used to study habitat loss and fragmentation, only habitat loss is explicitly described by ISARs. However, the observed patterns of the species-area relationship can implicitly capture other aspects such as fragmentation, isolation, species interactions, stochasticity and heterogeneity of the habitat where the species survive, and, at a different scale, possibly speciation and extinctions. These implicit natural aspects can be captured either as fitted parameters or as residual noise. For example, the exponent in the power-law relationship, a fitted parameter, was often interpreted as a sign of fragmentation (Connor and McCoy 1979, Triantis et al. 2012).

The simplicity and generality of the power-law ISARs was later used to predict long-term changes in species richness with mixed results, (Pimm et al. 1995, Ewers and Didham 2006 and references therein). These studies often assumed OF-SAR for one or more habitats (e.g. Thomas et al. 2004, Proença and Pereira 2013), which implies that individuals have a perfect dispersal capability in their current range and disregard the actual fragmented structure of the area.

On the other hand, simulation models have also been used to understand the origin of ISARs and the impact of fragmentation. As expected, each of these simulation models simplified the natural processes and complexities in different ways. For instance, Hubbell's (2001) model assumes that species interact identically in a homogeneous habitat where speciation occurs, while the Rybicki & Hanski (2013) model explicitly assumes non-interacting dissimilar species in a heterogeneous habitat without speciation. Interestingly, both models predict a power-law that describes the OF-SAR, but with different exponents: 0.25 for Rybicki & Hanski's model and 1 for Hubbell's (see proof in Appendix A2.1). Rybicki & Hanski's (2013) extended the ISAR to capture fragmentation patterns in highly fragmented habitats: the species-fragmented-area relationship (SFAR). As we show below, their proposed formula is consistent with the addition of a correction factor to the exponent of a power-law ISAR (Appendix A2.2). The correction factor is
itself a function of immigration and emigration rates, fragment shapes and distance among them. Unfortunately, this estimation requires a large amount of information about the species that is often unavailable. Also, the SFAR main application occurs when species dispersal among fragments is high. When species dispersal is heterogeneous, the fragments can be clustered in different subsets, as in Hanski et al. (2017). Therefore, each set of fragments, as defined by Rybicky and Hanski, corresponds to a patch in our proposed framework (fragments linked by black arrows in Appendix 1: Figure S1).

A.3.1 Neutral theory

Neutral theory predicts a tri-phasic species accumulation curve (SAC, also known as continental-SAR) (Hubbell 2001). Also, neutral theory predicts that species richness, the number of species in equilibrium (*S*) in a neutral community without migration and perfect mixing, can be approximated by the Fisher formula $S \approx \theta . \ln(1+J/\theta)$, where $\theta = 2Jv$, with *J* the number of individuals and *v* the speciation rate. That creates a non-linear relationship between *S* and *J* for a given θ (Hubbell 2001). However, replacing θ by 2Jv, gives $S=2Jv.\ln(1+1/2v)$. Assuming a homogeneous individual density across the habitat d, J = dA, and $S = cA^z$, with $c=2vd.\ln(1+1/2v)$ and z = 1.

Because Fisher equation is only approximated for large number of individuals, this relation is asymptotic when J increases. To confirm that the species-area curve proposed by Hubbell predicts a z = 1 when the area of the metacommunity increases, we simulated 40 communities using different values of area and speciation rate with the package UNTB in R (Hankin 2007). As expected, the lines become almost parallel as the area increases even for the relatively small metacommunities (Figure S8).



Figure S8 Simulated number of species expected for a given area and speciation rate using

Hubbell neutral model.

The light blue lines represent the lines with exponent 1. Note how in the log-log space, the relation between area and richness converge towards a z=1 as the area increases for all the speciation rates. Density is 1 individual per area unit. The red line represents the richness for a given area when the product of speciation rate and area (θ) is fixed, as in Hubbell (2001).

A.3.2 Species fragmented area relationship: An ISAR for highly fragmented habitats

To combine the effects of fragmentation and area in a highly fragmented habitat where individuals can still disperse among fragments, Rybicky and Hanski (2013) proposed the following expression to calculate species richness:

 $S = cA^z e^{-b/\lambda_M}$

where z is the species-area-relationship, b is an estimated parameter and λ_M is the metapopulation capacity of the fragmented landscape. Defining $b=b'\log(A)$, the richness is

 $S = cA^{z-b'/\lambda_M}$.

Framed in this way, *b*' can be interpreted as the product of the effect of the fragmented landscape on *z* for the particular group of species under study times a unit standardization component for λ_M . Therefore, Rybicky and Hanski (2013) provide a strong theoretical argument for the change in the exponent of a power-law regression in a fragmented landscape when individual dispersal among fragments is high.

A.3.3 References

- Connor, E. F., and E. D. McCoy. 1979. The statistics and biology of the species-area relationship. The American Naturalist 113:791–833.
- Ewers, R. M., and R. K. Didham. 2006. Confounding factors in the detection of species responses to habitat fragmentation. Biological Reviews 81:117–142.
- Hankin, R. K. S. 2007. Introducing untb, an R Package For Simulating Ecological Drift Under the Unified Neutral Theory of Biodiversity. Journal of Statistical Software 22:1–15.
- Hanski, I., T. Schulz, S. C. Wong, V. Ahola, A. Ruokolainen, and S. P. Ojanen. 2017. Ecological and genetic basis of metapopulation persistence of the Glanville fritillary butterfly in fragmented landscapes. Nature Communications 8.
- Hubbell, S. P. 2001. The unified neutral theory of biodiversity and biogeography. Princeton University Press, Princeton.
- Pimm, S. L., and R. A. Askins. 1995. Forest losses predict bird extinctions in eastern North America. Proceedings of the National Academy of Sciences 92:9343–9347.
- Pimm, S. L., G. J. Russell, J. L. Gittleman, and T. M. Brooks. 1995. The Future of Biodiversity. Science 269:347–350.
- Proença, V., and H. M. Pereira. 2013. Species–area models to assess biodiversity change in multi-habitat landscapes: The importance of species habitat affinity. Basic and Applied Ecology 14:102–114.
- Rosenzweig, M. 1995. Species diversity in space and time. Cambridge University Press, Cambridge; New York.
- Rybicki, J., and I. Hanski. 2013. Species–area relationships and extinctions caused by habitat loss and fragmentation. Ecology Letters 16:27–38.
- Thomas, C. D., A. Cameron, R. E. Green, M. Bakkenes, L. J. Beaumont, Y. C. Collingham, B. F. N. Erasmus, M. Ferreira de Siqueira, A. Grainger, L. Hannah, and others. 2004. Extinction risk from climate change. Nature 427:145–148.
- Triantis, K. A., F. Guilhaumon, and R. J. Whittaker. 2012. The island species-area relationship: biology and statistics: The island species-area relationship. Journal of Biogeography 39:215–231.

A.4 Proof of maximum richness when a patch is symmetrically divided

To study the effect of cutting a patch symmetrically (even sections) we used first an analytical approach for the case of two patches using a general form of a species-area relationship. Then, we assumed the Arrehnius SAR and generalize the results for multiple patches.

A.4.1 General ISAR shape: Two even patches can foster more species than a single large one

If *A* is the area of a patch, we assume that any species-area relationship, hereafter *S*(*A*), will have zero species when area is zero (*S*(0) = 0), will have more species when area increases (d*S*/d*A* = S'(A) > 0) and will have a decreasing slope, so that in a larger patch, more area is needed to add a new species ($d^2S/dA^2 = S''(A) < 0$). Under those conditions, we want to evaluate the difference between two habitat configurations, a single patch of area *A* and two patches of area *qA* and (1 – *q*)*A*, where *q* is any number between 0 and 1. In particular, we want to know if there is any value of q that minimizes the function F(q) = S(A) - [S(qA) + S(A - pA)].

The first and second derivative of F with respect to q are

F'(q) = A[S'(A - pA) - S'(pA)]

$$F''(q) = -A^{2}[S''(A - pA) + S''(pA)]$$

Since dS/dA > 0, *F'* will be zero when A = 0 and when S'(A - pA) = S'(pA). The second condition is the interesting one. In this case, the left and right hand terms will always be different because S' is always decreasing (third condition), unless that A-pA = pA, which occurs when p = 0.5. Also, F" will always be positive because $d^2S/dA^2 < 0$, which proofs that for any given A, p = 0.5 generates the minimum value for F. This implies that two patches of similar size have the larger number of species possible for a disjoint community regardless the specific shape of the SAR, as far as the SAR follows the three criteria previously outlined.

A.4.2 Arrhenius ISAR: Several even patches can foster more species than uneven ones

To study the effect of evenly cut patches on total richness for nested and disjoint communities we cut an area in a hundred of smaller pieces. The pieces were cut in different sizes according to

the parameter ω . When ω is 0, all the patches have the same area (Figure S9a). Larger ω values increase the asymmetry in the area of the patches, as represented by the colours. The minimum area of a patch is set to 0.1% of the total area. Then, we estimated the richness of disjoint and nested communities. Note that for any value of *z* the maximum richness of a disjoint community (Figure S9b) and the minimum richness of a nested community (Figure S9c) occurs when all the patches have the same area ($\omega = 0$).



Figure S9 Effect of patch size asymmetry on habitat richness for different exponents of the

species-area relationship.

(a) Proportion of the area covered by each of 100 patches of a fictitious habitat characterized by ω , sorted vertically from smallest to largest. The species richness (*S*) of the habitat is a function of ω (representing different patch-size

profile), the exponent of the species-area relationship (z) and the species turnover among patches. If turnover is maximum (b, disjoint community) each patch has a different group of species ($S = c \Sigma A_i^z$) and S is maximum, and if turnover is null (c, nested community), the largest patch (A_1) has all the species ($S = c A_i^z$), and S is minimum. Notice that the maximum (minimum) richness in a disjoint (nested) community occurs when habitat area is evenly distributed among the patches ($\omega = 0$).

B Appendices of chapter 3

B.1 Supplementary tables

Table S2 Description of the experimental sites.

The descriptions include the total number of species observed in each site during the experiment, the species originally identified as dominant species at the site level (and therefore removed in the dominant removal treatment), and the other species identified as dominant species at the plot level by the end of the experiment in the control plots and in the dominant removal plots. A general description of the site is also included.

Site	Species	Dominant species	New dominant species in dominant	Description
	richness	initial definition	removal plots base on volume, and	
		(additions at the end	rank change base on frequency	
		of the experiment)	(initial rank \rightarrow final rank)	
Т	13	Poa pratensis Solidago altissima	Asclepias syriaca $(4 \rightarrow 4)$ Bromus inermis $(3 \rightarrow 3)$ Linaria vulgaris $(2 \rightarrow 2)$ Vicia cracca $(1 \rightarrow 1)$	Slope facing west, next to a horse track, and probably kept as a lawn during several years. Management stopped 15 years ago.
F	29	Bromus inermis Poa pratensis (Solidago altissima)	Carex sp. $(4 \rightarrow 4)$ Hieracium caespitosum $(3 \rightarrow 1)$ Solidago altissima $(2 \rightarrow 2)$ Symphyotrichum ericoides $(5 \rightarrow 7)$ Symphyotrichum novae-angliae $(6 \rightarrow 2)$ Vicia cracca $(1 \rightarrow 4)$	Flat area. Probably used as a pasture during several years. Management stopped 15 years ago.
С	34	Bromus inermis Solidago altissima (Poa pratensis) (Vitis riparia)	Euthamia graminifolia $(4 \rightarrow 5)$ Poa pratensis $(1 \rightarrow 1)$ Symphyotrichum novae-angliae $(7 \rightarrow 5)$	Protected by a forested creek in the South. Probably used as a pasture during several years. Management stopped 15 years ago.
S	35	Bromus inermis Solidago altissima (Poa pratensis)	Asclepias syriaca $(12 \rightarrow 9)$ Cirsium arvense $(5 \rightarrow 7)$ Daucus carota $(8 \rightarrow 1)$ Monarda fistulosa $(12 \rightarrow 6)$ Poa pratensis $(2 \rightarrow 1)$ Symphyotrichum novae-angliae $(7 \rightarrow 4)$	Slope facing South, protected by forests and hills in most directions. Probably used as a pasture during several years. Management stopped 15 years ago.
N	54	Bromus inermis Solidago altissima (Crepis sp.) (Monarda fistulosa) (Poa pratensis) (Symphyotrichum novae-angliae)	Crepis sp. $(^+ \rightarrow 20)$ Euthamia graminifolia $(20 \rightarrow 10)$ Monarda fistulosa $(2 \rightarrow 1)$ Poa pratensis $(1 \rightarrow 1)$ Symphyotrichum lanceolatum $(14 \rightarrow 14)$ Symphyotrichum novae-angliae $(5 \rightarrow 1)$ Vitis riparia $(25 \rightarrow 23)$	Slope facing North, protected by a hill by the west. Probably used for hay production between 1940 and 1995. Area mowed nearby a few years before the beginning of the experiment.

Note: † Species absent in the first visit in that site.

Table S3 New dominant species initial ranking and site trend.

Rankings correspond to the rankings in Table 3, site trend compared with a reference site (R), as in Table 4. Arrows indicate average positive (\uparrow) or negative (\downarrow) effect of the treatment. Significance: *** < 0.001, ** < 0.01, * < 0.05, · < 0.1.

Constitution and the	Т	site	F	site	C site		S site		N site	
Species name	Rank	Trend	Rank	Trend	Rank	Trend	Rank	Trend	Rank	Trend
Vicia cracca	1	R	1	\downarrow^{***}						
Linaria vulgaris	2	R								
Bromus inermis	3	R								
Asclepias syriaca	4	R					11	\mathbf{v}		
Solidago altissima			2	R						
Hieracium caespitosum			3	R						
Carex sp.			4	R						
Symphyotrichum ericoides			5	R						
Symphyotrichum novae-angliae			6	R	7	\uparrow	7	^ **	5	个***
Poa pratensis					1	R	2	\uparrow^*	1	\uparrow
Euthamia graminifolia					4	R			20	\downarrow^{**}
Monarda fistulosa							12	R	2	个***
Cirsium arvense							5	R		
Daucus carota							8	R		
Symphyotrichum lanceolatum									15	R
Vitis riparia									25	R
Crepis sp.									32	R

Table S4 Effect of treatments in new dominant species in the final visit of the experiment in

comparison with the control plots.

All models were fitted using generalized least squares with different variance per treatment, volume log+1 transformed, and using only sites where the species became dominant. Despite not all models fitted properly, we kept a homogeneous approach for consistency among them. Control treatment variance is always 1. Arrows indicate average positive (\uparrow) or negative (\downarrow) effect of the treatment. Significance: *** < 0.001, ** < 0.01, * < 0.05, · < 0.1.

	Treatment variance rate to						Regression descriptors				
Scientific name	meat	control	variance		Treatment trend				Res.	Normality	
Sciencine name							Obs.	d.f	test		
	DR	DM	NDR	NDM	DR	DM	NDR	NDM			
Symphyotrichum novae- angliae	1.58	0.83	1.00	0.85	^ ***	\uparrow	\uparrow	\downarrow	117	109	<0.001
Poa pratensis [†]	1.18	1.43	1.29	0.77	\uparrow^{**}	\mathbf{v}	\downarrow	\downarrow	85	70	0.384
Euthamia graminifolia	1.79	0.95	1.00	0.94	\uparrow^*	\uparrow	\uparrow	\uparrow	60	54	0.043
Bromus inermis	2.75	1.15		2.81	\uparrow ·	\checkmark		\uparrow	25	21	0.014
Daucus carota	1.78	2.41	1.88		\uparrow ·	\uparrow ·	\uparrow		25	21	< 0.001
Vitis riparia	++			++	\uparrow			\uparrow	20	17	<0.001
Carex sp.	4.39	1.80	0.65	0.94	\uparrow	\uparrow ·	\checkmark	\checkmark	30	25	0.008
Monarda fistulosa	3.54	2.11	0.87	0.50	\uparrow	\uparrow	\uparrow	\downarrow	59	53	< 0.001
Symphyotrichum ericoides	3.10	1.59		3.67	\uparrow	\uparrow		\uparrow	25	21	<0.001
Solidago altissima	2.32	0.40	0.42	1.22	\uparrow	\downarrow^*	\checkmark	\uparrow	27	22	0.619
Cirsium arvense	1.85	0.81	0.80	0.55	\uparrow	\uparrow	\checkmark	\checkmark	28	23	0.033
Asclepias syriaca	1.58	0.19	0.29	0.50	\uparrow	\mathbf{v}	\checkmark	\downarrow	58	52	< 0.001
Symphyotrichum Ianceolatum	1.55	0.21	0.50	0.36	\uparrow	\downarrow	\downarrow	\downarrow	30	25	0.002
Crepis sp.	1.37	1.52	1.94	1.12	\uparrow	\uparrow	\uparrow	\uparrow	30	25	< 0.001
Hieracium caespitosum	1.32	0.09	1.07	0.94	\uparrow	\checkmark	\uparrow	\uparrow	29	24	< 0.001
Linaria vulgaris	0.85	1.01	0.36	0.68	\checkmark	\checkmark	\downarrow^{**}	\mathbf{v}	29	24	0.376
Vicia cracca	0.67	0.71		1.40	\downarrow^*	\downarrow		\uparrow	50	45	0.337

Notes: [†]*Poa pratensis* was positively affected by DR treatment in every site but N. ^{††} The general approach used in this analysis generated an extremely large variance probably meaningless, caused by the presence of few large *Vitis riparia* in few plots in each treatment.



B.2 Supplementary figures

Figure S10 Average change in volume and richness of the plots in each site and treatment.

Site dominant species are the two species identified as dominants at the beginning of the experiment, new dominants are at least two species needed to occupy 50% of the volume of any plot in the dominant removal treatment at the end of the experiment, and non-dominant is any other species. In the x-axis, B indicates the volumes before treatments were applied, while '14 to '17 represent the four years of the experiment.



Figure S11 Bray-Curtis dissimilarity changes along time for each study site in control plots,

using peak-season biomass only to assess site level trends.

Values lower than the first-year peak-season dissimilarity indicate convergence (deterministic mechanisms driving a change in the community) while values higher indicates divergence (stochastic mechanisms driving a change in the community). Similarity values were measured using different sets of species: 'All' indicates all the species were included; 'New' indicates the dominant species defined in the last year in the control treatment in each site were used; 'ND' indicates that the species not flagged as dominants were used and 'new in DR' is like 'new', but using the dominant species defined in the dominant removal treatment.





Dotted lines represent z-scores of ± 1.96 , and the dashed line the z-score of zero. Significance: *** < 0.001, ** < 0.01, * < 0.05, \cdot < 0.1.

C Appendices of chapter 4

C.1 Supplementary tables

Table S5: Marginally significant missing paths in the final DAG.

Abbreviations as in Table S6. Assuming that the DAG is valid, each missing path test identify the set of variables that have to be held constant in order to confirm that there is no statistical correlation between a pair of variables of interest. The missing path column indicates the pair of variables analysed in each case. In all the cases, the hypothesis is that the estimate value of the correlation is zero. The variables held constant are not shown. Only missing paths with p-value between 0.05 and 0.1 are shown.

Missing path	Estimate	Standard error	df	p-value
$\Delta_{ses} \sim Gz$	0.9792	0.4908	54	0.0511
$\Delta_{ m ses}$ ~ Lat	0.0362	0.0194	54	0.0677
TR ~ T	0.2253	0.1135	51	0.0524
BM ~ Δ_{ses}	0.0935	0.0506	49	0.0707
$TR \simeq \Delta_{ses}$	-0.3003	0.1770	50	0.0960

Table S6: Variables used to model the difference in the dispersion of dominants and non-

dominant plants and	the sources	of information	used to	measure them.

Variables	Definition	Source*
Response v	ariables	
$\Delta_{ses.mntd}$	Difference in the phylogenetic dispersion (measured as the standardized effect size of the mean nearest taxonomic distance) of the third most dominant minus the third lest dominant plant species in each site. Dominance defined as the mean cover per plot.	Based on NutNet cover and Zanne et al. 2014
pGr	Proportion of graminoid biomass in the site.	Based on NutNet biomass
BM	Total aboveground annual net production by m ² (log transformed)	Based on NutNet biomass
LM	Legume aboveground annual net production by m2 (log + 1 transformation)	Based on NutNet biomass
Climatic va	ariables*	
Т	Mean annual temperature.	Worldclim 2
TR	Range of the annual temperature.	Worldclim 2
Р	Total annual precipitation (log transformed)	Worldclim 2
PV	Variance of the monthly precipitation	Worldclim 2
Location v	ariables	
Elev	Altitude in meters above see level.	Worldclim 2
Lat	Distance from the equator in degrees.	NutNet description
Hem	North/South hemisphere dummy	NutNet description
Manageme	nt	
Gz	Is the site subject to grazing?	NutNet description
An	Is the site of anthropogenic origin? (e.g. restored)	NutNet description
Fr	Is the site subject to prescriptive fire?	NutNet description
Phylogenet	ic information	
Rec	Proportion of the phylogenetic information of the tree contained in the tree tips.	Based on NutNet cover and Zanne et al. 2014
S	Number of species present in the site.	Based on NutNet cover
PD	Site level Faith's Phylogenetic Diversity (Faith 1992), obtained as the sum of the branches of the site tree	Based on NutNet cover and Zanne et al. 2014
Gm	Site level Gamma statistics (Pybus and Harvey 2000), obtained as the mean of a hundred randomly generated hinary trees based on the site tree	Based on NutNet cover and Zanne et al. 2014
Bal	Site level Colless statistics (Mooers and Heard 1997), obtained as the mean of a hundred randomly generated binary trees based on the site tree.	Based on NutNet cover and Zanne et al. 2014

Notes: * All Worldclim 2 (Fick and Hijmans 2017) variables were included in the model, but discarded because they (1) were correlated among each other or (2) they were poor predictors. Variables listed here but not present in the final model were discarded for similar reasons.

C.2 Supplementary figures



Figure S13: Global distribution of the study sites. In blue, sites with cover information only, in red sites with cover and biomass info.



Figure S14: Comparing the difference in dispersion between dominants and non-dominant

plants when the community is divided in two or three partitions.

The difference was assessed as the standardize effect size of the dominant minus the non-dominant partition (Δ_{ses}). The first row shows the results when dispersion is measured using the mean nearest taxon distance (MNTD) and the second using mean pairwise distance (MPD). The columns represent different ways to measure the dominance of the

species: mean cover, frequency (proportion of plots where the species is present) and mean cover of the plots where the species is present only.



Figure S15: Phylogenetic dissimilarities among sites when each site is partitioned in two or

three dominance partitions.

Two dominance partitions include dominant (D) and non-dominant (ND) species, each with half of the species. Three dominance partitions include also an intermediate dominance (ID), each with a third of the species in the site. The phylogenetic dissimilarity is measured as the multisite Sørensen (SOR – multi) and as the mean pairwise Sørensen dissimilarity (SOR – pairs). In both cases, the indices were decomposed in their turnover (SIM) and nestedness (SNE) fractions. Dashed lines represent the observed value when dominance is assigned base on observed mean species cover per plot, while the densities represent the expected value when the species are randomly distributed in the two or three partitions. Rows 1 and 3 include all sites (All), 2 and 4 exclude Australian sites (nAu), and 3 and 6 exclude sites in North America (nNA).

C.3 Alternative definitions of disparity in phylogenetic dispersion

We used five different ways to estimate the difference on phylogenetic dispersion between dominants and non-dominants using mean-nearest taxonomic distance (MNTD). Here we describe the different ways we did the partitions and the analysis done to compare them. Because the partitions would be composed of different number of species, we estimated the standardized effect size of MNTD (SES.MNTD) for each of the partitions ($\Delta_{ses.mntd}$) in each site. In every case, we pruned the phylogenetic tree to the species present in each site before running the analysis. In that way, the standardized effect sizes for any dominance partition used the same species pool that only includes the species observed in that site at the beginning of the experiment, before nutrients were added to the site. Given that the SES estimates included only the species present, the $\Delta_{ses.mntd}$ is a measure of the difference in how dominant and non-dominant plants' performance correlate with the environment.

C.3.1 Detecting Partition methods

C.3.1.1 Partitions with similar number of species

This method assumed that each partition is composed of a similar number of species. The partition with the most dominant species will be composed of the highest ranking species. The next partition will be composed by the next higher ranking species, etc. We define partitions for two and three groups of dominance. Δ_{ses} was then the difference in the phylogenetic dispersion metric between the highest ranking minus the lowest ranking partition. Each site dispersion trend was then classified as either "D_{SES} < ND_{SES}", "D_{SES} > ND_{SES}" or "NONE" if Δ_{ses} was significantly less, more or equal to 0, respectively.

C.3.1.2 Partitioning by Jenks criteria

We used Jenks' natural break optimization to find two classes of species, dominant and nondominant, in each site. Each class composed of species with similar abundance. The optimization algorithm cuts the community using an abundance criterion, minimizing the variability inside each class and maximizing the variability between the classes. The class with the largest abundance values corresponded to the dominant species. Δ_{ses} and the site dispersion trend was then estimated as when partitioning in groups with same number of species (see C.3.1.1). The underlying assumption of this method is that a variance criterion can effectively distinguish biologically meaningful processes.

C.3.1.3 Thresholds

Because different sites can have different number of dominant species, we explored several cut values in a community to test the hypothesis that at least in one of them $\Delta_{\text{ses.mntd}}$ was higher or lower than zero. We defined dominant partitions in each site as a group of species that include either 10, 20, ..., 90% (hereafter cutoffs) of the most dominant species in the site. For each one of these cutoffs we split the community in the dominant (D) and non-dominant partition (ND) assigning every species to one of these partitions.

To decide if the dominant and non-dominant partitions were differently assorted, we estimated $\Delta_{ses.mntd}$ for each combination of site, variable and cutoff. We assumed the $\Delta_{ses.mntd}$ values were normally distributed with variance 2 and, if both partitions were equally dispersed, mean 0. At any cutoff, a value significantly lower (higher) than zero indicated dominants were more clustered (overdispersed) than non-dominants. In the absence of a meaningful way to identify a relevant cutoff for every site, we identify the site dispersion trend (either "D_{SES} < ND_{SES}", "D_{SES} > ND_{SES}", "NONE" or "BOTH") if any difference between dominants and non-dominants was significant in any cutoff smaller or equal to a given threshold. For instance, at a 30% threshold, we assumed "D < ND" if any negative significant difference existed at a 10, 20, or 30% cutoff. If, for instance, "D < ND" at 10% and "D > ND" at 30%, we marked that trend as "BOTH". The actual value $\Delta_{ses.mntd}$ was then assumed to be the one at which the minimum difference existed, unless there was a significant positive difference only. In that case, we used the maximum value.

As expected, the probability of detecting a difference increased with threshold. We opted for the 50% threshold as a compromise between detection capability and error prone of the criteria (Figure S16). We found no evidence of both trends acting simultaneously.



Figure S16: Proportion of sites with different dispersion trends at different thresholds that capture any partition of species in dominants and non-dominants at a cut-off equal or lower than the given threshold.

Each panel represents a different dominance measure. We consider four potential dispersal trends: At least in one cut-off lower than the threshold: the dominants are less dispersed than non-dominants (D < ND), the dominants are more dispersed than non-dominants (D > ND), both conditions hold (D < ND at one cut-off and D > ND at another cut-off) and none conditions hold at any cut-off.

C.3.1.4 Extremes

In this method to partition the communities we defined the dominant and non-dominant partitions as the ones that would create a minimum or a maximum $\Delta_{\text{ses.mntd}}$. We started with the ses.mntd values for each site, partition and cutoff generated in C.3.1.3. Then, we obtained the difference between each combination of the dominant and non-dominant partitions. While doing so, each species was assigned to either the dominant or the non-dominant partition, or to none of them. We choose the minimum and maximum $\Delta_{\text{ses.mntd}}$, and estimated their significance. We identified the site trend as either "D_{SES} > ND_{SES}", "D_{SES} < ND_{SES}", "NONE" or "BOTH" according the rules in Table S7. The final $\Delta_{\text{ses.mntd}}$ used to compare with the other partitioning methods was estimated base on the observed trend (Table S7). We opted to focus in in the min($\Delta_{\text{ses.mntd}}$) unless strong evidence against it existed because the minimum curved seems farther from zero, indicating it is less likely to be randomly generated (Figure S17).

Table S7: Criteria used to define the site level dispersion trend and the measure of $\Delta_{ses.mntd}$ used to compare with the other partitioning methods.



Figure S17: Observed density of the $\Delta_{ses.mntd}$ globally when obtaining the maximum (max) and minimum (min) difference in the dominant and non-dominant dispersion.

Each panel represents a different dominance measure.

C.3.2 Comparison among partition methods

To compare the different partition methods we estimated the correlation between the $\Delta_{ses.mntd}$ obtained in each site with each method and compared the number of sites classified in each trend (Table S8).

Table S8: Comparison between the different partition methods in terms of the trends in the

partitions detected (a) and the correlation between $\Delta_{ses.mntd}$ values (b).

The dominants dispersion can either be lower, equal or higher than non-dominants (D < ND, None, D > ND). Two partition methods (thresholds and extremes) can also detect that a site has a dominant/non-dominant partition which is less than zero and another one that is more than zero simultaneously (Both). The numbers in the trends table (a) indicate the number of sites classified in each trend by each method. The numbers in the correlations table (b) indicate the Pearson's correlation value (**: p-value < 0.01, ***: p-value < 0.001, n=75 sites).

a. Trends					
	Eq.rich3	Eq.rich2	Jenks	Thresholds	Extremes
Both	-	-	-	0	2
D <nd< td=""><td>16</td><td>17</td><td>13</td><td>33</td><td>50</td></nd<>	16	17	13	33	50
D>ND	0	1	0	3	7
None	59	57	62	39	16
b. Correlation	ons				
	Eq.rich3	Eq.rich2	Jenks	Thresholds	Extremes
Eq.rich2	0.720 ***				
Jenks	0.412 ***	0.342 **			
Thresholds	0.630 ***	0.720 ***	0.620 ***		
Extremes	0.714 ***	0.745 ***	0.528 ***	0.760 ***	

T 1

All the partition methods were positively correlated (p-value < 0.01, most p-value < 0.001). Despite it was not the most sensitive partition method; we used the partitioning in groups of similar species richness because it had fewer assumptions.

C.4 Testing other legumes' proportion, biomass and their interaction with climatic variables

The naïve interpretation of legumes biomass as a predictor of total biomass relies in the assumption that nitrogen fixation provides proportional benefits to every group of plants independent of climatic conditions and of the total biomass. However, the benefits provided by nitrogen fixers (legumes) could be strongly limited under dry conditions because of limited amount of water could deter the nitrogen from being transferred from one plant to another. Similarly, given a general positive correlation between productivity of different plants, we may expect the proportion of legumes in the community to be more informative of the benefits provided by the legumes than the total legume's biomass.

To test those alternative hypothesis we: (1) compared the amount of variance that could be explained of the proportion of legumes and of the legume's biomass by the step-backwards model using richness and Faith's phylogenetic diversity, (2) compared the amount of variance of

the total biomass that could be explained using the proportion and total amount of legumes and (3) added interaction terms between total (or proportion of) legumes biomass and temperature and precipitation (linear and quadratic terms). When doing the final test with legumes' proportions, we started with the total biomass model that use the legumes biomass as a predictor and replaced it with the legumes' proportion because the step-backwards regression dropped the proportion in the test (2).

C.4.1 Legumes as a response variable

Total legume's biomass and proportion of legumes biomass were predicted by a similar set of variables, but the total biomass model was able to explain almost a third of the variance, almost three times more than the proportion's model. None of the global climatic drivers of productivity were significant to explain any of the variables, but the AIC criterion retained temperature range as an informative variable for total biomass. The variable better explained was the total legume's biomass, when using Faith's phylogenetic distance

Response Total†			Proportion		
Predictor	Using S	Using PD	Using richness	Using PD	
(Intercept)	0.810604	0.265321	-0.00303	-0.009081	
Temperature range	-0.079506	-0.07996			
Anthropogenic	1.375766 ***	1.537756 ***	0.030968 *	0.033062 *	
Richness (S)	0.029419 ***		0.000614 *		
Faith's PD (PD)		0.000793 ***		0.000014.	
R ²	0.341	0.3453	0.1376	0.1189	

Table S9: Linear models explaining legumes' biomass and proportion of legumes' biomass.

Note: † Log+1 transformed

C.4.2 Legumes as a predictor variable

Legumes biomass was consistently a better predictor of total biomass than proportion of legumes. Further, in contrast with previous research, the interaction of precipitation and legumes was non-significant (p-value = 0.1636), indicating that in this global grasslands experiment legumes increased biomass independently of the effect of total precipitation.

Table S10: Total biomass as predicted by (a) legumes' biomass and (b) proportion of legumes' biomass with and without interaction with climatic variables (either temperature or precipitation).

a) Legumes biomass as predictor of total biomass

Predictor	Legumes	Legumes: Temperature	Legumes: Precipitation	Legumes: Precipitation ²
(Intercept)	4.996335 **	4.891114 **	4.271493 *	3.826916 *
Temperature	-0.068356 *	-0.061714.	-0.072713 *	-0.063571.
Abs(Latitude)	-0.040978 *	-0.039633 *	-0.041824 *	-0.033921.
Elevation	-0.000530 ***	-0.000522 **	-0.000555 ***	-0.000506 **
Precipitation	0.458931 **	0.455071 **	0.583355 ***	0.586930 ***
Precipitation ²				-0.869553
Legumes BM [†]	0.150338 **	0.195818.	0.172979 **	0.149878 *
Legumes BM [†] :Temperature		-0.004281		
Legumes BM [†] :Precipitation			-0.672958	-0.572446
Legumes BM [†] :Precipitation ²				0.035304
R^2	0.4691	0.4712	0.4895	0.5126
AIC	93.62	95.38	93.38	94.75
Shapiro (p.value)††	0.689	0.6125	0.941	0.937

Note: † Log+1 transformed. †† Residuals' normality test.

Prodictor	Logumos	Legumes:	Legumes:	Legumes:
Tredictor	Leguines	Temperature	Precipitation	Precipitation ²
(Intercept)	4.376173 *	4.395525 *	3.946231 *	3.286491.
Temperature	-0.063070.	-0.063886	-0.069071.	-0.058535.
Abs(Latitude)	-0.037602.	-0.037868.	-0.039599 *	-0.029064
Elevation	-0.000501 **	-0.000503 **	-0.000532 **	-0.000475 **
Precipitation	0.544424 ***	0.544599 ***	0.633340 ***	0.648190 ***
Precipitation ²				-1.358582 *
Legumes prop.	1.329392	1.13414	2.424481	2.130696
Legumes prop.: Temperature		0.020637		
Legumes prop.:Precipitation			-22.256797	-21.023972
Legumes prop.:Precipitation ²				7.643273
R^2	0.393	0.3931	0.4084	0.4632
AIC	101.2	103.2	101.8	100.2
Shapiro(p.value)†	0.4767	0.4799	0.7181	0.593

Note: † Residuals' normality test.

C.4.3 Conclusions

Overall, these analyses support the selection of legumes' biomass as an independent predictor of total biomass.

C.5 Other aspects of the structural equation model

Here we expand the interpretation of other aspects of the final model that are relevant to interpret it, but were not directly related to the relatedness disparity between dominant and non-dominant plants. To do this, we start with the final path model and assessed the paths present, but also the absent ones as the absence of a path between two variables suggest that those variables are not causally related. We also identified variables that can have opposite effects on a descendent variable, and for each pair of ancestor and descendant variables we estimated the partial residuals of the former onto the latter (which include both direct and indirect links) and estimated the overall effect of the ancestor onto the descendant using Pearson's correlation. More detailed analyses of these interactions are warranted, but are beyond the scope of this study.

C.5.1 Results

We found that Northern hemisphere sites have higher precipitation than the Southern hemisphere sites, while proportion of graminoids was higher in the South. Management (anthropogenic and grazing) and tree topology descriptors had mostly non-significant relations with the other variables of the model.

We found a counterintuitive negative effect of temperature on total biomass. We also found that biomass decreased in sites located farther from the equator (r = -0.31) and at higher elevations (r = -0.44), ant that the proportion of graminoids decreased with elevation (r = -0.42). Given that temperature is also affected by elevation and latitude, the negative effect of temperature on biomass is likely a confounding covariate effect. In fact, removing either elevation or latitude results in the temperature effect on biomass becomes positive, though not significant (p > 0.1 in both cases, results not shown).

C.5.2 Discussion

In the sites sampled we found that graminoids performance improved under some environmental conditions, in particular wider temperature ranges and lower elevations, consistent with a grassland biome or habitat (e.g. Olson et al. 2001). However, we did not find a correlation between success of graminoids (measured as proportion of graminoids biomass) and total site biomass, probably because woody (and heavier) species could outcompete grasses under more

productive climatic conditions and controlling by herbivory pressure (Figure 15) (Olson et al. 2001).

Other paths the model show that higher productivity occurs at sites that are at lower latitude, lower elevation and have greater precipitation, consistent with Grace et al. (2007, 2016). The hemisphere effect on different variables likely captures asymmetric distribution of some plant families at either side of the equator, or perhaps the more extensive historical deforestation in the Northern hemisphere which, in turn, may result in more NutNet sites established in areas that previously supported forests or another type of biome. The uneven representation of the samples along the latitudinal gradient, and the fact that it may actually capture very different pattern and processes (light availability, temperature seasonality, biogeography) deter any further interpretation of the mechanisms underlying of this relationship.

C.6 Are C3/C4 pathways relevant to characterize graminoids prevalence?

We did a brief test of trait relevance to distinguish among dominant and non-dominant graminoid species. Using the information already recorded by the Nutrient Network we compared the phylogenetic signal of the dominance in graminoids, with the phylogenetic signal of photosynthetic path as represented by the proportion of C3 species. A strong link between dominance and either C3 or C4 photosynthetic path should appear as dominant branches (green in the right panel) being either red or green in the left panel. A visual inspection of both phylogenetic trees does not suggest any correlation between them, as some C3 lineages are more likely to be dominant than to be in any other dominance category (green lineages in the left panel), and the same lack of pattern can be seen for the C4 lineages (red lineages in the left panel). Other traits available in the Nutrient Network dataset were either variable by location, and therefore more likely to be biased towards dominant species only, or have no internal variation in graminoids (i.e. life form and legumes).

D/(D+ID+ND)

C3/(C3+C4)



D Appendices of chapter 5

D.1 Supplementary tables

Table S11: Variables used to model the difference in the dispersion of dominants and nondominant plants and the sources of information used to measure them.

Variables	Definition	Source*
Functional g	roup suffix	
Gr	Graminoids, plants in the families Poaceae and Cyperaceae	
Lg	Legumes, plants in the family Fabaceae	
Fr	Forbs, any tracheophyte which is not in a legume or a graminoid	
Plant commu	inity biomass descriptors	
GM	Graminoids aboveground annual net production by m2 (log + 1	Based on NutNet biomass
	transformation)	
FM	Forbs aboveground annual net production by m2 (log + 1	Based on NutNet biomass
	transformation)	
LM	Legume aboveground annual net production by m2 (log + 1	Based on NutNet biomass
	transformation)	
Litter	Litter mass by m2 (log + 1 transformation)	Based on NutNet biomass
Plant commu	inity biodiversity metric	
SR	Species richness	Based on cover biomass
PD	Faith's PD (Faith 1992)	Based on NutNet biomass
MPD	Mean pairwise phylogenetic distance (Cadotte and Davies 2016)	Based on NutNet biomass
MNTD	Mean nearest taxonomic distance (Cadotte and Davies 2016)	Based on NutNet biomass
Climatic var	iables*	
Т	Mean annual temperature.	Worldclim 2
TR	Range of the annual temperature.	Worldclim 2
Р	Total annual precipitation (log transformed)	Worldclim 2
TW	Average temperature of the wettest quarter	Worldclim 2
Location var	iables	
Elev	Altitude in meters above see level.	Worldclim 2
Lat	Distance from the equator in degrees.	NutNet description
Hem	North/South hemisphere dummy	NutNet description
Management		
Gz	Is the site subject to grazing?	NutNet description
An	Is the site of anthropogenic origin? (e.g. restored)	NutNet description
Fr	Is the site subject to prescriptive fire?	NutNet description

Notes: * All Worldclim 2 (Fick and Hijmans 2017) variables were included in the model, but discarded because they (1) were correlated among each other or (2) they were poor predictors. Variables not present in the final model were discarded for similar reasons.

Table S12 Conceptual mechanisms describing the relation between productivity and

diversity.

Models are organized according the expected pattern created, the causal direction of the mechanism and if the mechanism relate to the potential or the realized biomass or diversity. Potential models focus on biomass/diversity relationships in the absence of species interactions, while realized models focus on those interactions.

Model #	Expected pattern	Mechanism	Causal direction	Potential/ Realized	References	Comment
1.	Positive	Productivity →Odds of survival →Richness	$P \rightarrow D$	Realized	Grime 1973 in Grace et al.	
2.		Productivity → Density → Richness	P → D	Potential	(Hubbell 2001a)	Can be interpreted in the frame of neutral theory.
3.		Richness → Phylogenetic information/functional groups → Odds of complete resource exploitation → Productivity	$D \rightarrow P$	Potential	(Loreau and Hector 2001)	Name: Complementarity effect
4.		Richness → Phylogenetic information/functional groups → Odds of high efficiency → Productivity	$D \rightarrow P$	Potential	(Loreau and Hector 2001)	Name: Selection effect
5.		Richness →Odds of facilitation →Productivity	$D \rightarrow P$	Realized	(Cardinale et al. 2002)	
6.		Productivity → Species pool → Richness	P → D	Potential	Gillman and Wright in 2006 in Grace et al.	"Importance of species pools typically predominates over local interactions in shaping observed patterns"
7.	Negative	Productivity → Competitive asymmetry → Richness	$P \rightarrow D$	Realized	(Huston 1994)	Name: Dynamic equilibrium model

Model	Expected	Mechanism	Causal	Potential/	References	Comment
#	pattern		direction	Realized		
8.		Productivity → spatial heterogeneity in the ratios of limiting resources → coexistence via exploitation of unique resource ratios → Richness	P → D	Realized	(Tilman 1977) Tilman 1982 in Grace et al.	Name: Resource ratio model
9.		(Negative correlation: Productive sites, Area for evolution) Area available for evolution → Species pool	P → D†	Potential	Taylor, Aarssen, and Loehle in Grace et al.	Name: Habitat templet model
10.	Hump-shaped	Productivity → Competitive asymmetry → Richness → Odds of survival → Richness <i>Together</i> : Richness is hump-shaped	P → D	Potential/ Realized	Grime 1973 in Grace et al.	Name: Humped- back model Combination of 1 and 7
11.		Productivity → Increase in average individual size → Increase in density <i>Condition</i> : Density x average individual size is capped. <i>Together</i> : Density is hump-shaped Density → Richness	P → D	Potential	(Oksanen 1996)	Based on 2

Notes: Mechanism descriptors: $A \rightarrow B$: A increases B. $A \rightarrow B$: A decreases B. Green arrows (\rightarrow) indicate a causal variable simultaneously causes more than one dependent variable (e.g. $A \rightarrow B \rightarrow C$: A increases B and increases C). If the dependent variables have themselves more dependent variables, the later ones are indented. Other operators are described in italics (i.e. *correlation*: correlation between variables; *condition*: extra condition imposed in the model; *together*: the combined result of parallel green arrows or of their dependent variables). Causal direction indicates the overall direction of the productivity-diversity relationship: P >> D, productivity drives diversity; D >> P, diversity drives productivity. † Given that a correlation between productive sites and area for evolution drives the model, there is no real causal link between productivity and diversity.





Figure S18. Average effect of each functional group on diversity for the species richness

(top three rows) and Faith's PD (bottom three rows) models.

Vertical dashed lines indicate the location of the maximum in the estimated fitted curve for the internal effect of a functional group biomass in its own diversity, when present in the observed mass range. A solid line indicates that linear and quadratic terms were significant in the model (p-value < 0.05). Dot-dashed line indicates that only the linear terms were significant while dotted lines indicate that none term was significant. Mass values were log(+1) transformed.



Figure S19 Effect of the spatial sampling process in each site on each functional group diversity estimated from the Faith's PD and the species richness models.



Figure S20 Effect of each functional group diversity and biomass on litter mass for the species richness (top row) and Faith's PD (bottom row) models.

A solid line indicate that linear and quadratic terms were significant in the model (p-value < 0.05). Dot-dashed line indicates that only the linear terms were significant while dotted lines indicate that none term was significant. Mass values were log(+1) transformed.

D.3 Detailed description of structural equation models: potential biodiversity-biomass correlation

We fitted several models to test for different directions of causal links, as well as to assess if other variables used to predict similar models could improve the predictions. Because different models included different response or predictor variables a single metric to compare among all them is not feasible. Instead, we compare each one with the final models used in the main document, which we present in detail first here.

D.3.1 Final model

Our metamodel (Figure S21) represents our general expectation that the constraints created by the environment (location, climate, management) will set the conditions where species pool relatedness will modify biomass production, which in turn controls litter accumulation. Because only 31 sites were available we used path analysis, fitting first independent models and

combining them later (Shipley 2016). To facilitate the interpretation, we pre-identified uncorrelated climatic variables and used them without combining them (Table S11), only precipitation was log transformed.



Figure S21 Scheme of the basic relation expected between the location, climate,

management, phylogenetic biodiversity, biomass and litter.

Litter depended on legume (LM), forb (FM) and graminoid (GM) biomass. Each biomass was expected to be a function of climate (described as precipitation, P, average temperature of the wettest quarter, TW, annual average temperature, T, and temperature range, TR), species phylogenetic relatedness and local management. Site location (described as a North-South hemisphere dummy, Hem, distance to equator, Lat, and elevation, Elev). In turn, site location controlled the climatic conditions. Except for the links between location and climate, all the other links could be dropped.

We used a step-backwards regression to identify the most important drivers that better describe the metamodel. To capture the sequential role of environmental drivers followed by diversity ones we first regressed biomass using climate and management descriptors, and identified the most important drivers using step-backwards regression. Then, we added the phylogenetic information and run a second step-backwards linear regression to refine the biomass models. We compared these models with models using a single step-backwards step. Then, we confirmed that no one other climatic variable from the Worldclim v2 dataset could improve the functional groups biomass models. First, we described litter as a function of biomass of each functional group only. Finally, site location variables were included as descriptors of climatic variables to capture the original way climatic variables have been interpolated (Fick and Hijmans 2017).

The independent models were placed together in a single structural equation model (SEM) using the package *piecewiseSEM* (Lefcheck 2016) and tested for missing links. We added missing links if significant and logical (Shipley 2016). Once we obtained the final model, we re-fitted the model using structural equation models in the package *lavaan* (Rosseel et al. 2017). This second approach fits all the parameters of each partial regression in a single step using maximum likelihood. In contrast to the piecewise approach it is robust to circularities in the model, but it requires many more observations.

The main missing links that we found were related to a negative impact of precipitation on temperature range and of the several functional group diversity descriptors on litter. Interestingly, temperature range was originally dropped as a predictor of legumes by the first step-backwards process, but we added later to the model as the d-separation tests suggested that there was a missing link between them.

The piecewise approach provided a good fit of the model (Fisher's C = 182.48, d.f. = 184, p-value = 0.518), but when fitting the same model with the maximum likelihood approach, the fit measures of the model were not as good as expected (robust p-value = 0.902, CFI = 0.722, RMSEA = 0.167, SRMR = 0.088). That problem was mostly caused by the low predictability associated to the climatological descriptors, because when the models that aimed to capture their correlation were dropped the model met the expected criteria (robust p-value = 0.989 > 0.05, CFI = 0.988 > 0.9, RMSEA = 0.037 < 0.08, SRMR = 0.035 < 0.08). These descriptors are not required in the maximum likelihood method used by *lavaan*, because the model is robust to correlations between independent (exogenous) variables.

The d-separation test identified some missing paths (p-value < 0.1) that we decide to not include in the model, because the actual mechanism was hard to explain (Table S13). Interestingly, five of the seven missing paths suggest correlations between climatic and phylogenetic variables. Further research is needed to validate the logic of those relations and incorporate them properly in a model.

Table S13 Missing paths identified by d-separation tests with p-value less than 0.1 in the final model.

Response	Predictor	Conditional	Estimate	SD	df	Critical value	p-value
lgram_mass	slatitude	mpd_fr + mpd_gr + anthropogenic + mpd_lg + mntd_lg + MAT_v2 + IMAP_v2 + TEMP_WET_Q_v2 + MAT_RANGE_v2	-0.2632	0.1317	20	-1.9986	0.0594
llitter_mass	elevation	grazed + mpd_gr + pd_gr + mpd_lg + mntd_lg + lforb_mass + lgram_mass	-0.0009	0.0004	22	-2.3002	0.0313
MAT_v2	mpd_gr	slatitude + alatitude + elevation	-0.0423	0.0106	26	-3.9828	0.0005
TEMP_WET_Q_v2	mpd_gr	slatitude + alatitude + elevation	0.0804	0.0384	26	2.0941	0.0461
TEMP_WET_Q_v2	pd_fr	slatitude + alatitude + elevation	0.0040	0.0021	26	1.9337	0.0641
TEMP_WET_Q_v2	pd_gr	slatitude + alatitude + elevation	0.0459	0.0158	26	2.9033	0.0074
IMAP_v2	mpd_lg	slatitude + alatitude + elevation	0.0037	0.0018	26	1.9890	0.0573

Table S14 Final model summary

lavaan (0.5-23.1097) converged normally after 197 iterations

Number of observations	31		
Estimator	ML		
Minimum Function Test Statistic	35.440		
Degrees of freedom	34		
P-value (Chi-square)	0.400		
P-value (Bollen-Stine Bootstrap)	0.989		
Parameter Estimates:			
Information	Observed		
Standard Errors	Bootstrap		
Number of requested bootstrap draws	1000		
Number of successful bootstrap draws	819		
Regressions:			
Estimate Std.Err	z-value P(> z)	Std.lv	Std.all
llegu_mass ~			

grazed	1.432	0.873	1.640	0.101	1.432	0.388
1MAP_v2	0.557	0.601	0.927	0.354	0.557	0.163
mpd_fr	0.025	0.007	3.586	0.000	0.025	0.599
mntd_fr	-0.019	0.006	-3.245	0.001	-0.019	-0.673
mpd_gr	-0.016	0.008	-1.932	0.053	-0.016	-0.437
mntd gr	0.040	0.016	2.434	0.015	0.040	0.443
MAT_RANGE_v2	-0.204	0.108	-1.889	0.059	-0.204	-0.356
lforb mass ~						
MAT RANGE v2	-0.117	0.048	-2.419	0.016	-0.117	-0.311
pd fr	0.001	0.000	3.451	0.001	0.001	0.532
pd gr	-0.007	0.002	-4.233	0.000	-0.007	-0.639
mntd gr	0.018	0.010	1.785	0.074	0.018	0.308
lgram mass ~						
anthropogenic	-0.951	0.261	-3.647	0.000	-0.951	-0.405
1MAP v2	0.854	0.232	3.687	0.000	0.854	0.330
MAT V2	-0.028	0.022	-1.266	0.205	-0.028	-0.154
MAT RANGE v2	-0.095	0.034	-2.807	0.005	-0.095	-0.221
TEMP WET O v^2	0.081	0.017	4.779	0.000	0.081	0.691
mpd fr	0.007	0.004	1.696	0.090	0.007	0.209
mpd lg	-0.012	0.005	-2.246	0.025	-0.012	-0.446
mntd la	0.013	0.006	2.285	0.022	0.013	0.483
mpd gr	-0.017	0.003	-5.290	0.000	-0.017	-0.584
llitter mass ~	-0.017	0.005	-5.290	0.000	-0.017	-0.504
lgram mass	1.040	0.473	2,198	0.028	1.040	0.568
lforb mass	_1 030	0 539	_1 910	0.056	_1 030	-0 490
mpd la	-1.030	0.000	1 474	0.030	-1.030	-0.400
mpa_ig	0.029	0.019	1.4/4	0.141	0.029	0.602
	-0.029	0.023	-1.209	0.204	-0.029	-0.567
pd_gr	-0.026	0.007	-3.750	0.000	-0.026	-1.150
mpd_gr	0.035	0.01/	1.990	0.047	0.035	0.672
grazed	-1.293	1.430	-0.904	0.366	-1.293	-0.253
A						
Covariances:		a. 1 5			a. 1 1	a. 1 11
11	Estimate	Std.Err	z-va⊥ue	P(> z)	Std.IV	Std.all
.llegu_mass ~~						
.llitter_mass	0.247	0.263	0.942	0.346	0.247	0.239
grazed ~~						
IMAP_v2	0.015	0.027	0.547	0.584	0.015	0.100
mpd_fr	-0.304	2.168	-0.140	0.888	-0.304	-0.026
mntd_fr	7.183	4.357	1.648	0.099	7.183	0.410
mpd_gr	-5.827	2.081	-2.800	0.005	-5.827	-0.436
mntd_gr	-1.534	0.658	-2.332	0.020	-1.534	-0.275
MAT_RANGE_v2	-0.211	0.134	-1.578	0.114	-0.211	-0.241
pd_fr	-69.141	35.998	-1.921	0.055	-69.141	-0.332
pd_gr	-12.172	4.512	-2.698	0.007	-12.172	-0.399
anthropogenic	0.120	0.039	3.081	0.002	0.120	0.744
MAT_v2	0.379	0.330	1.150	0.250	0.379	0.185
$TEMP_WET_Q_v2$	0.397	0.432	0.920	0.358	0.397	0.123
mpd_lg	-2.227	1.982	-1.123	0.261	-2.227	-0.154
mntd_lg	-1.459	1.990	-0.733	0.464	-1.459	-0.105
lmap_v2 ~~						
mpd_fr	1.909	2.141	0.892	0.373	1.909	0.150
mntd_fr	-0.155	3.185	-0.049	0.961	-0.155	-0.008
mpd_gr	-2.164	2.244	-0.964	0.335	-2.164	-0.150
mntd_gr	-1.272	1.218	-1.044	0.297	-1.272	-0.211
MAT_RANGE_v2	-0.319	0.157	-2.035	0.042	-0.319	-0.339
pd_fr —	14.011	31.951	0.439	0.661	14.011	0.062
pd gr	0.236	3.790	0.062	0.950	0.236	0.007
anthropogenic	0.038	0.029	1.323	0.186	0.038	0.222
MAT V2	0.152	0.400	0.380	0.704	0.152	0.069
$\overline{\text{TEMP}}$ WET O v2	0.868	0.707	1.227	0.220	0.868	0.249
mpd lg	5,502	2,762	1.992	0.046	5.502	0.353
mntd lg	4.807	2.492	1,929	0.054	4.807	0.321
mpd fr ~~						
mntd fr	639.723	367.586	1,740	0.082	639.723	0.419
mpd ar	136.036	155.003	0.878	0.380	136.036	0.117
			0,719	0,472	39.803	0.082
mntd ar	39,803	55.349	J • / ± J	5.112	57.005	0.002
MAT RANCE V2	39.803	55.349 9.479	0.626	0,531	5.937	0.078
mntd_gr MAT_RANGE_v2 pd_fr	39.803 5.937 2678 580	55.349 9.479 2945 597	0.626	0.531	5.937 2678 580	0.078
mntd_gr MAT_RANGE_v2 pd_fr pd_gr	39.803 5.937 2678.580 687.842	55.349 9.479 2945.597 404.619	0.626 0.909	0.531 0.363 0.089	5.937 2678.580 687.842	0.078 0.148 0.259
mntd_gr MAT_RANGE_v2 pd_fr pd_gr anthropogenic	39.803 5.937 2678.580 687.842	55.349 9.479 2945.597 404.619	0.626 0.909 1.700 0.170	0.531 0.363 0.089	5.937 2678.580 687.842	0.078 0.148 0.259
mntd_gr MAT_RANGE_v2 pd_fr pd_gr anthropogenic MAT_v2	39.803 5.937 2678.580 687.842 0.527 38.002	55.349 9.479 2945.597 404.619 3.098 23.402	0.626 0.909 1.700 0.170	0.531 0.363 0.089 0.865	5.937 2678.580 687.842 0.527	0.078 0.148 0.259 0.038
mntd_gr MAT_RANGE_v2 pd_fr pd_gr anthropogenic MAT_v2 TEMP_WET_02	39.803 5.937 2678.580 687.842 0.527 38.908	55.349 9.479 2945.597 404.619 3.098 23.402	0.626 0.909 1.700 0.170 1.663	0.531 0.363 0.089 0.865 0.096 0.842	5.937 2678.580 687.842 0.527 38.908	0.078 0.148 0.259 0.038 0.218
mntd_gr MAT_RANGE_v2 pd_fr pd_gr anthropogenic MAT_v2 TEMP_WET_Q_v2 mpd_lc	39.803 5.937 2678.580 687.842 0.527 38.908 8.249	55.349 9.479 2945.597 404.619 3.098 23.402 41.410 217 568	0.626 0.909 1.700 0.170 1.663 0.199	0.531 0.363 0.089 0.865 0.096 0.842	5.937 2678.580 687.842 0.527 38.908 8.249	0.078 0.148 0.259 0.038 0.218 0.029
mntd_gr MAT_RANGE_v2 pd_fr pd_gr anthropogenic MAT_v2 TEMP_WET_Q_v2 mpd_lg mntd_lc	39.803 5.937 2678.580 687.842 0.527 38.908 8.249 -149.293 -309 155	55.349 9.479 2945.597 404.619 3.098 23.402 41.410 217.568 202 432	0.626 0.909 1.700 0.170 1.663 0.199 -0.686 -1.527	0.531 0.363 0.089 0.865 0.096 0.842 0.493 0.137	5.937 2678.580 687.842 0.527 38.908 8.249 -149.293 -309 155	0.078 0.148 0.259 0.038 0.218 0.029 -0.118
<pre>mntd_gr MAT_RANGE_v2 pd_fr pd_gr anthropogenic MAT_v2 TEMP_WET_Q_v2 mpd_lg mntd_lg mntd_fr_cc</pre>	39.803 5.937 2678.580 687.842 0.527 38.908 8.249 -149.293 -309.155	55.349 9.479 2945.597 404.619 3.098 23.402 41.410 217.568 202.438	0.626 0.909 1.700 0.170 1.663 0.199 -0.686 -1.527	0.531 0.363 0.089 0.865 0.096 0.842 0.493 0.127	5.937 2678.580 687.842 0.527 38.908 8.249 -149.293 -309.155	0.078 0.148 0.259 0.038 0.218 0.029 -0.118 -0.255
mpd gr	-262.540	293.059	-0.896	0.370	-262.540	-0.152
---	--	---	---	--	--	--
mntd gr	19.651	102.837	0.191	0.848	19,651	0.027
MAT DANCE 172	-26 138	18 581	_1 407	0 160	-26 138	_0 231
nd fr	15000 000	ECAE AAO	-1.407	0.100	15000 000	-0.231
pd_II	-15888.008	5645.448	-2.814	0.005	-15888.008	-0.590
pd_gr	-1188.106	657.754	-1.806	0.071	-1188.106	-0.301
anthropogenic	6.925	4.548	1.522	0.128	6.925	0.332
MAT v2	12.699	39.679	0.320	0.749	12.699	0.048
$\overline{\text{TEMP}}$ WET O v2	-7.928	59.899	-0.132	0.895	-7.928	-0.019
mpd la	-551 138	278 319	_1 980	0 048	-551 138	_0 294
mpa_ig	220 055	270.011	-1.500	0.040	220 055	0 122
mnta_ig	-239.955	2/0.411	-0.88/	0.3/5	-239.955	-0.133
mpd_gr ~~						
mntd gr	373.805	151.505	2.467	0.014	373.805	0.678
MAT RANGE v2	-0.297	15.722	-0.019	0.985	-0.297	-0.003
nd fr	4607.583	3249.016	1.418	0.156	4607.583	0.224
pa_11	2007.005	265 220	I.410	0.150	2007.005	0.224
pa_gr	2069.005	305.329	5.003	0.000	2069.005	0.687
anthropogenic	-5.210	2.407	-2.165	0.030	-5.210	-0.328
MAT_v2	-102.813	39.603	-2.596	0.009	-102.813	-0.508
TEMP WET O v2	-33.699	78.556	-0.429	0.668	-33.699	-0.106
mpd la	-71.885	224,209	-0.321	0.749	-71.885	-0.050
mptd la	12 000	199 054	0 060	0 0/5	12 000	0 000
	-12.900	109.054	-0.009	0.945	-12.900	-0.009
mntd_gr ~~						
MAT_RANGE_v2	-1.856	6.048	-0.307	0.759	-1.856	-0.052
pd fr	-160.883	1426.237	-0.113	0.910	-160.883	-0.019
nd gr	417,410	150.315	2.777	0.005	417,410	0.332
anthronogonic	_1 770	0 010	_2 176	0 030	_1 770	-0 267
anchiopogenic	-1.770	0.013	-2.1/0	0.030	-1.770	-0.207
MAT_v2	-35.702	23.661	-1.509	0.131	-35.702	-0.422
$TEMP_WET_Q_v2$	-44.309	46.814	-0.946	0.344	-44.309	-0.332
mpd la	-92.949	98.429	-0.944	0.345	-92.949	-0.156
mntd la	-60.426	82.353	-0.734	0.463	-60.426	-0.105
MAD DANCE #2	00.120	02.000	0.,01	0.105	001120	0.105
MAT_RANGE_VZ ~~	~~ ~~~				~~ ~~~	
pd_fr	92.583	253.384	0.365	0./15	92.583	0.069
pd_gr	-4.298	28.524	-0.151	0.880	-4.298	-0.022
anthropogenic	-0.219	0.158	-1.382	0.167	-0.219	-0.211
MAT v2	2.652	2,164	1,225	0.220	2.652	0.201
	1 270	3 206	0 300	0 697	1 270	0 061
IEMF_WEI_Q_V2	-1.270	12 075	-0.309	0.097	-1.270	-0.001
mpa_ig	-15.043	13.8/5	-1.084	0.278	-15.043	-0.161
mntd_lg	-18.487	11.996	-1.541	0.123	-18.487	-0.206
pd fr ~~						
pd gr	18894.317	7667.203	2.464	0.014	18894.317	0.403
anthronogenic	-67 867	37 784	_1 796	0 072	-67 867	_0 274
		607 174	-1.750	0.072		-0.274
MAT_V2	242.857	08/.1/4	0.353	0.724	242.85/	0.077
$TEMP_WET_Q_v2$	1295.473	913.367	1.418	0.156	1295.473	0.261
mpd_lg	6270.238	3202.960	1.958	0.050	6270.238	0.282
mntd lg	-170.792	3109.931	-0.055	0.956	-170.792	-0.008
nd ar ~~						
pa_gr	12 224	4 000	2 456	0 014	10 004	0 227
anthropogenic	-12.234	4.982	-2.450	0.014	-12.234	-0.337
MAT_v2	13.155	72.106	0.182	0.855	13.155	0.028
$TEMP_WET_Q_v2$	177.145	86.734	2.042	0.041	177.145	0.243
mpd lg	976.365	545.595	1.790	0.074	976.365	0.299
mntd la	648.826	408,453	1,588	0.112	648.826	0.207
anthronogonia az	0100020	1000100	10000		0100020	01207
anchiopogenice and	0 104	0 0 0 0	0 5 2 5	0 500	0 104	
MAT_V2	0.194	0.363	0.535	0.592	0.194	0.080
$TEMP_WET_Q_v2$	1.068	0.494	2.164	0.030	1.068	0.278
mpd_lg	-1.947	2.787	-0.699	0.485	-1.947	-0.113
mntd lg	-0.309	2.725	-0.114	0.910	-0.309	-0.019
MAT v2 ~~		5				
	26 717	12 016	1 0 2 /	0 052	26 717	0 546
TEMP_WET_Q_V2	20./1/	13.810	1.934	0.053	20./1/	0.546
mpd_lg	69.697	36.249	1.923	0.055	69.697	0.318
mntd_lg	46.035	34.312	1.342	0.180	46.035	0.219
TEMP WET O v2 ~~						
mpd la	123,989	64.217	1,931	0.054	123,989	0.359
mptd la	117 003	59 170	2 014	0 044	117 003	0 355
minicu_ry	11/.003	50.4/9	2.014	0.044	111.003	0.305
mpa_ig ~~						
mntd_lg	1343.070	269.567	4.982	0.000	1343.070	0.903
Intercepts:						
L	Estimate	Std.Err	z-value	P(> z)	Std.lv	Std.all
			0 720	- (- 4)	2 2 4 1 4	2 50.011
	2 1 1 1	/ ///0	-0./20	0.400	-3.444	-2.009
.llegu_mass	-3.444	4.728	6 050	0 000	C 000	7 010
<pre>.llegu_mass .lforb_mass</pre>	-3.444 6.990	4.728	6.859	0.000	6.990	7.812
.llegu_mass .lforb_mass .lgram_mass	-3.444 6.990 -0.614	4.728 1.019 1.679	6.859 -0.366	0.000 0.714	6.990 -0.614	7.812 -0.599
.llegu_mass .lforb_mass .lgram_mass .llitter mass	-3.444 6.990 -0.614 13.434	4.728 1.019 1.679 4.022	6.859 -0.366 3.340	0.000 0.714 0.001	6.990 -0.614 13.434	7.812 -0.599 7.148
.llegu_mass .lforb_mass .lgram_mass .llitter_mass grazed	-3.444 6.990 -0.614 13.434 0.161	4.728 1.019 1.679 4.022 0.064	6.859 -0.366 3.340 2.514	0.000 0.714 0.001 0.012	6.990 -0.614 13.434 0.161	7.812 -0.599 7.148 0.439
.llegu_mass .lforb_mass .lgram_mass .llitter_mass grazed lMAP_v2	$ \begin{array}{r} -3.444 \\ 6.990 \\ -0.614 \\ 13.434 \\ 0.161 \\ 6.678 \end{array} $	4.728 1.019 1.679 4.022 0.064	6.859 -0.366 3.340 2.514	0.000 0.714 0.001 0.012	6.990 -0.614 13.434 0.161	7.812 -0.599 7.148 0.439
.llegu_mass .lforb_mass .lgram_mass .llitter_mass grazed 1MAP_v2	-3.444 6.990 -0.614 13.434 0.161 6.678	4.728 1.019 1.679 4.022 0.064 0.070	6.859 -0.366 3.340 2.514 95.488	0.000 0.714 0.001 0.012 0.000	6.990 -0.614 13.434 0.161 6.678	7.812 -0.599 7.148 0.439 16.849
.llegu_mass .lforb_mass .lgram_mass .llitter_mass grazed 1MAP_v2 mpd_fr	-3.444 6.990 -0.614 13.434 0.161 6.678 230.283	4.728 1.019 1.679 4.022 0.064 0.070 5.693	6.859 -0.366 3.340 2.514 95.488 40.450	0.000 0.714 0.001 0.012 0.000 0.000	6.990 -0.614 13.434 0.161 6.678 230.283	7.812 -0.599 7.148 0.439 16.849 7.179

mpd_gr	43.525	6.567	6.628	0.000	43.525	1.198
mntd gr	12.940	2.795	4.629	0.000	12.940	0.853
MAT_RANGE_v2	12.112	0.397	30.527	0.000	12.112	5.102
pd_fr	1645.574	95.946	17.151	0.000	1645.574	2.909
pd_gr	480.095	14.625	32.827	0.000	480.095	5.789
anthropogenic	0.258	0.075	3.458	0.001	0.258	0.590
MAT_v2	12.110	0.988	12.262	0.000	12.110	2.175
TEMP WET Q v2	14.006	1.502	9.325	0.000	14.006	1.594
mpd_lg	47.463	7.232	6.563	0.000	47.463	1.206
mntd_lg	36.666	6.865	5.341	0.000	36.666	0.970
Variances:						
	Estimate	Std.Err	z-value	P(> z)	Std.lv	Std.all
.llegu_mass	0.701	0.197	3.561	0.000	0.701	0.381
.lforb_mass	0.439	0.089	4.939	0.000	0.439	0.549
.lgram_mass	0.103	0.026	3.935	0.000	0.103	0.097
.llitter_mass	1.525	0.477	3.197	0.001	1.525	0.432
grazed	0.135	0.043	3.132	0.002	0.135	1.000
lMAP_v2	0.157	0.034	4.559	0.000	0.157	1.000
mpd_fr	1029.053	240.848	4.273	0.000	1029.053	1.000
mntd_fr	2269.742	785.744	2.889	0.004	2269.742	1.000
mpd_gr	1319.625	221.073	5.969	0.000	1319.625	1.000
mntd_gr	230.241	102.426	2.248	0.025	230.241	1.000
MAT_RANGE_v2	5.635	1.168	4.825	0.000	5.635	1.000
pd_fr	320027.535	71533.406	4.474	0.000	320027.535	1.000
pd_gr	6878.124	2268.107	3.033	0.002	6878.124	1.000
anthropogenic	0.191	0.036	5.331	0.000	0.191	1.000
MAT_v2	31.014	7.550	4.108	0.000	31.014	1.000
$TEMP_WET_Q_v2$	77.201	26.492	2.914	0.004	77.201	1.000
mpd_lg	1548.598	231.155	6.699	0.000	1548.598	1.000
mntd_lg	1427.598	304.298	4.691	0.000	1427.598	1.000
R-Square:						
	Estimate					
llegu_mass	0.619					
lforb_mass	0.451					
lgram_mass	0.903					
llitter mass	0.568					

D.3.2 Building a full community model

We followed a similar approach but using full community biomass and biodiversity descriptors. First, we fitted the total of forbs, legumes and graminoids biomass as a function of environmental site conditions using a step-backwards approach. Then we created three models and added site richness to the first; site level Faith's PD to the second; and site level Faith's PD, mean phylogenetic distance (MPD) and mean nearest taxonomic distance (MNTD) to the third. We run a second step-backwards model with each of them and compared the three models using AIC and choose the one with lowest AIC (Table S15). The models with richness and PD only dropped those variables, so we added those variables again to explore the information that they can provide. Separately, we fitted the litter model as a function of biomass and combined with

the PD-MPD-MNTD model, which was the best in the previous step. We used path analysis to assess the model quality and identify potential missing links.

Table S15 Model selection process to represent the effect of diversity on biomass.

Model	Initial	Final	R ²	AIC	dAIC
Basic	T, TR, P, TW, Elev, Lat, Hem, Gz, An, Fr	T, TR, P, TW, Fr	0.741	38.9	1.39
Richness	T, TR, P, TW, Fr, SR	T, TR, P, TW, Fr, SR †	0.746	40.3	2.74
PD	T, TR, P, TW, Fr, PD	T, TR, P, TW, Fr, PD^{\dagger}	0.745	40.4	2.84
PD, MPD, MNTD	T, TR, P, TW, Fr, PD, MPD, MNTD	T, TR, P, TW, Fr, MPD	0.767	37.5	0
NT / [†] N7 · 11 1	1 1 11 1 '				

Only MPD provided relevant information to improve the expected biomass model.

Note: [†] Variable dropped and added again

The combined model (PD, MPD and MNTD) was a good descriptor of the causal links (Fisher's C = 39.4, df = 40, p-value = 0.497). We included missing links identified by the d-separation test that were useful to improve litter mass prediction, which decreased AIC but increased AICc. Only one link had p-value < 0.1 and was not included because of unclear interpretation (Table S16).

Table S16 Missing paths identified by d-separation tests with p-value less than 0.1 in the model with total estimates of biomass and biodiversity.

Response	Predictor	Conditional	Estimate	SD	df	Critical value	p-value
TEMP_WET_Q_v2	burned	slatitude + alatitude + elevation	6.466	3.5632	26	1.81	0.081

In comparison with this model, the final model that partitioned the community in functional groups was able (1) to explain a larger proportion of litter accumulation and standing biomass and (2) provide more insight in the mechanisms that could be driving biomass and litter accumulation.

D.3.3 Single step step-backwards regression

This approach was used to validate if splitting the variables in two groups when running the stepbackwards regressions generated any improvement to the biomass models compared with a single step-backwards regression (Table S17).

Step-backwards	Legumes		Graminoids		Forbs	
approach	df	AIC	df	AIC	df	AIC
Single step	10	102.4	11	53.3	13	100.8
Double step	9	103.4	11	53.3	6	78.0

Table S17 Comparing single two double step step-backwards approach.

Resulting graminoid models were identical, forb model was better with the double step and legume model with a single step. The only difference between the legume model with a single step and the double step, was that the single step included a negative effect of temperature on legumes. Because that negative trend was hard to explain in the absence of interactions with other species, we opted to keep the double step model in all the regressions.

D.3.4 Other climatic variables

Once the biomass models were fitted we compared the models with models fitted using climatic variables only, and tested in particular if adding average insolation of the site improved the biomass estimates of the functional groups. We used the Insolation Incident On A Horizontal Surface (kWh/m²/day) estimated and made available by NASA Surface meteorology and Solar Energy (SSE) Release 6.0 Data Set (Jan 2008) which average monthly data over a 22-year period. (http://eosweb.larc.nasa.gov/sse/).

Purely climatic variables always had higher AICc values (Table S18) and models including insolation did not improve the biomass predictions (Table S19).

Type of predictor variables		gumes	Graminoids Forb			orbs
		AICc	df	AICc	df	AICc
Climate only	20	189.3	15	86.9	16	112.0
Climate+Management+Diversity	9	103.4	11	53.3	6	78.0

 Table S18 Comparing single two double step step-backwards approach

Table S19 Log-likelihood ratio test of the	final model against the model	including insolation
--	-------------------------------	----------------------

Model	Res.Df	RSS	Df	Sum of Sq	F	Pr(>F)
Legumes	23	21.7				
	22	20.3	1	1.339	1.450	0.241
Graminoids	26	13.6				
	25	13.5	1	0.111	0.205	0.655
Forbs	21	3.2				
	20	3.2	1	0.001	0.005	0.942

D.3.5 Including phylogenetic descriptors and species richness together

We re-did the step-backwards selection process for the biomass models to test if adding species richness (besides PD, MPD and MNTD) could improve the models. Models that included species richness were not better than the models with phylogenetic information only (Table S20).

Table S20 Comparing the use of species richness besides phylogenetic information.

Type of predictor variables		Legumes		Graminoids		orbs
		AICc	df	AICc	df	AICc
Including SR, PD, MPD and MNTD	8	105.1	11	91.5	11	53.2
Final model (including PD, MPD and MNTD)	9	103.4	6	78.0	11	53.2

D.3.6 Using PD but not MPD or MNTD

We re-fitted the model to test the relevance of the information provided by phylogenetic similarity (MPD and MNTD). We followed similar steps, using the double step approach but without MPD or MNTD, adding missing climatic links and diversity-litter links. The resultant model was good (Fisher's C=117.72, df=124, p-value=0.642). However, the d-separation test detected a potential missing link between forbs and legumes biomass (Table S21). We fitted models adding the link in both directions (forbs \leftarrow legumes and legumes \leftarrow forbs) and found that the AIC and the AICc were lower when forbs affected legumes than the opposite direction (Table S22). That direct link is lost in the final model, which includes MPD and MNTD. Another interesting result of this model is that graminoid diversity has a negative impact on graminoid biomass, but the trend is positive in forbs and legumes (Table S23).

Response	Predictor	Conditional	Estimate	SD	df	Critical value	p-value
MAT_v2	pd_lg	slatitude + alatitude + elevation	0.0166	0.0075	26	2.1985	0.037
TEMP_WET_Q_v2	pd_fr	slatitude + alatitude + elevation	0.004	0.0021	26	1.9337	0.0641
TEMP_WET_Q_v2	pd_gr	slatitude + alatitude + elevation	0.0459	0.0158	26	2.9033	0.0074
llitter_mass	MAT_v2	slatitude + alatitude + elevation + grazed + pd_gr + lforb_mass + lgram mass	-0.2937	0.1218	22	-2.4119	0.0247
lforb_mass	llegu_mass	grazed + pd_lg + IMAP_v2 + pd_fr + pd_gr + MAT_RANGE_v2	0.2588	0.1239	23	2.0888	0.048

Table S21 Missing paths identified by d-separation tests with p-value less than 0.1 in the model with diversity described by PD only.

Table S22	Comparing	PD-only mode	els with and	without	direct link	between	forb	and

legume biomass.

Model	AIC	AICc	К	n
PD-only model	209.72	-60.53	46	31
PD-only model + (forbs ← legumes)	208.74	-56.67	47	31
PD-only model + (legumes ← forbs)	204.61	-60.80	47	31

Table S23 Structural equation model using only Faith's PD as a descriptor of biodiversity.

The model represents the effect of (1) climate, management and phylogenetic relatedness on functional group biomass production; and (2) management, phylogenetic relatedness and functional group biomass production on litter accumulation. The part of the model that captures the effect of location on climate is identical to the climate part in the final model. Values represent coefficients; squared bracketed values represent standardized coefficients.

Predictor	Legumes (gr)†	Forbs (gr) †	Graminoi ds (gr) †	Litter (gr)†
(Intercept)	-9.657 *	6.768 ***	0.345	11.583 ***
Temp. (°C)			0.034 . [0.185]	
Temp. range (°C)		-0.121 * [-0.321]	-0.111 * [-0.256]	
Precip. (mm)†	0.831 [0.246]	[•••=-]	1.087 *** [0.420]	
Temp. wet quarter			[0.663]	
Anthropogenic			-0.841 ** [-0.359]	
Burned				
Grazed	0.961 [0.264]			-1.886 * [-0.366]
Forbs (gr)†	0.476 * [0.318]			-0.747 * [-0.353]
Graminoids (gr)†				0.523 . [0.283]
Legumes (gr)†				[0.205]
Graminoids: PD		-0.006 ** [-0.515]	-0.004 ** [-0.315]	-0.015 *** [-0.652]
Forbs: PD		0.001 ** [0.477]	-0.000 [-0.150]	
Legumes: PD	0.009 * [0.331]	[]	L · · · J	
SD	1.1204	0.7613	0.4782	1.5120
R ²	0.4125	0.3695	0.839	0.4663
Normality test	0.6123	0.5983	0.704	0.1614

Notes: \dagger variable were log-transformed or log+1 transformed if zeros present. Cells' colors: Red: Variable not added. White: Included in the model, if no numbers, the variable was dropped by the step backwards algorithm. Orange: Variables fixed (not affected by step backwards algorithm). Green: Variables added after the piecewise SEM detected high correlation for the conditional variables. Model fit: Fisher's C = 182.48, d.f. = 184, p-value = 0.518.

D.3.7 Alternative litter models

We compared the final litter model with models without partitions using species phylogenetic descriptors (Faith's PD, MPD and MNTD) and richness only. The last two models were built using a similar process by doing first a step-backwards model to reduce the set of variables in the models to those that provide relevant information. We compared the models using AIC and AICc (Table S24).

Table S24 Comparing litter models using partitions and without partitions with differentdiversity descriptors

Type of predictor variables	Litter (log+1 transformed)					
Type of predictor variables	n	R ²	AIC	AICc		
Final model with partitions	31	0.579	118.8	0		
Phylogenetic model without partitions	31	0.250	130.7	11.89		
Species richness model without partitions	31	0.252	130.6	11.81		

D.3.8 Building the step-backwards models without diversity info

We tested fitting the models without diversity information, using a single step-backwards regression. The combined model was good (Fisher's C=96.25, df=90, p-value=0.307) but the biomass models were poorer predictors than the models including diversity information (Table S25).

Table S25 Comparing the use of species richness besides phylogenetic information

	L	itter.	Le	gumes	Grar	ninoids	Fe	orbs
Type of predictor variables		AICc	df	AICc	df	AICc	df	AICc
Models without diversity	4	132.0	4	108.2	7	67.6	3	85.5
Final models	9	127.4	9	103.4	11	53.3	6	78.0

D.3.9 Productivity as a driver of diversity

We tested the directionality of the causal link by reversing the metamodel and repeating the algorithm that we used to build the final model: Running first step-backwards models for biomass using climate and management, then for phylogenetic diversity (relatedness) descriptors using biomass and for litter using biomass only. Then we combined the individual models and do a path analysis test, and added the missing links that were significant and logic. As before, we did not add direct links between phylogenetic diversity descriptors and climate or site location predictors, and we only added predictors if the sign was consistent with some biological expectation. The final reversed model was unlikely to be a good model to explain the observed data (Fisher's C=318.2, df=264, p-value=0.012), and it was worst than the final model (Δ AICc = 168.5, Table S26).



Figure S22 Scheme of the basic relation between the location, climate, management,

phylogenetic biodiversity, biomass and litter when assuming that biomass drive diversity.

Litter depended on legume (LM), forb (FM) and graminoid (GM) biomass. Each biomass was expected to be a function of climate (described as precipitation, P, average temperature of the wettest quarter, TW, annual average temperature, T, and temperature range, TR) and local management. Species relatedness was caused by biomass. Site location (described as a North-South dummy, Hem, distance to equator, Lat, and elevation, Elev). In turn, site location controlled the climatic conditions. Except for the links between location and climate, all the other links could be dropped. The red arrows represent the arrows reversed compared to Figure S21.

Matamadal	Fi	sher's	AIC				
Metamodel	Fisher.C	df	p.value	AIC	AICc	К	n
Diversity is caused by biomass	318.2	264	0.012	474.2	217.4	78	31
Biomass is caused by diversity	182.5	184	0.518	294.5	48.9	56	31

Table S26 Comparing the use of species richness besides phylogenetic information.

D.4 Details of the realized biomass-biodiversity models

We represented functional group diversity as a function of functional groups biomass, litter mass as a function of each functional group diversity and biomass. We used equation 1 to represent functional group diversity, and equation 2 to represent litter mass and linked them using the packages *lavaan* (Rosseel et al. 2017) and *piecewiseSEM* (Lefcheck 2016). Biomass was represented using quadratic terms. As before, we used a step-backwards approach using AIC to drop variables that were not needed. Linear terms were always kept if the quadratic terms were retained by the step-backwards approach. We started fitting the model with *piecewiseSEM* and testing for missing links. If a link was needed (p-value < 0.1), we added if biologically logical. We found in several cases that a direct link between functional groups diversity metrics was required, but we did not include those links in the absence of a conceptual model that support them (for instance, graminoid PD significantly decreased legume PD even after controlling for the effect of above ground biomass, which was non-significant p-value > 0.15 for the linear and the quadratic term).

D.4.1 Final Faith's PD model

lavaan (0.5-23.1097) converged normally after 1191 iterations

Number of obser	vations		1697					
Estimator				ML				
Minimum Functio	on Test Stati	stic		7.190				
Degrees of free	edom			3				
P-value (Chi-sq	[uare)			0.066				
P-value (Bollen	-Stine Boots	trap)		1.000				
Parameter Estimat	es:							
Information			C	bserved				
Standard Errors	5		Bo	otstrap				
Number of reque	sted bootstr	ap draws		1000				
Number of succe	essful bootst	rap draws		1000				
Regressions:								
	Estimate	Std.Err	z-value	P(> z)	Std.lv	Std.all		
pd_lg ~								
arch.us	-305.208	18.553	-16.451	0.000	-305.208	-0.196		
azi.cn	-5.048	14.566	-0.347	0.729	-5.048	-0.003		
barta.us	-189.969	34.285	-5.541	0.000	-189.969	-0.122		
bnch.us	-100.077	32.520	-3.077	0.002	-100.077	-0.064		
bogong.au	-303.214	17.940	-16.901	0.000	-303.214	-0.194		
bttr.us	-106.608	30.402	-3.507	0.000	-106.608	-0.068		
bunya.au	-180.438	37.590	-4.800	0.000	-180.438	-0.116		
burrawan.au	-302.907	18.085	-16.749	0.000	-302.907	-0.194		
burren.ie	-5.027	13.928	-0.361	0.718	-5.027	-0.003		
cbgb.us	-212.007	27.732	-7.645	0.000	-212.007	-0.181		
cdcr.us	-236.828	27.111	-8.735	0.000	-236.828	-0.195		
cdpt.us	-201.511	29.686	-6.788	0.000	-201.511	-0.181		
chilcas.ar	-173.516	36.942	-4.697	0.000	-173.516	-0.111		
cowi.ca	130.672	17.091	7.646	0.000	130.672	0.084		
doane.us	-151.863	30.894	-4.916	0.000	-151.863	-0.136		
gall.it	-139.300	35.239	-3.953	0.000	-139.300	-0.092		
gilb.za	19.744	11.262	1.753	0.080	19.744	0.014		
glac.us	-106.129	32.320	-3.284	0.001	-106.129	-0.068		
glcr.us	-179.097	35.436	-5.054	0.000	-179.097	-0.115		
hall.us	-293.756	22.876	-12.841	0.000	-293.756	-0.188		
hast.us	-32.676	32.407	-1.008	0.313	-32.676	-0.021		
hnvr.us	-35.527	19.082	-1.862	0.063	-35.527	-0.023		
kibber.in	-182.124	46.805	-3.891	0.000	-182.124	-0.117		
kidman.au	-76.634	28.811	-2.660	0.008	-76.634	-0.063		
kirik.ee	-0.943	12.453	-0.076	0.940	-0.943	-0.001		

koffler.ca	-29.994	19.554	-1.534	0.125	-29.994	-0.021
konz.us	12.367	38.358	0.322	0.747	12.367	0.008
lead.us	-322.895	21.739	-14.853	0.000	-322.895	-0.207
look.us	-272.335	31.183	-8.733	0.000	-272.335	-0.175
marc.ar	-200.210	44.354	-4.514	0.000	-200.210	-0.128
mclaus	-279.749	19.212	-14.562	0.000	-279.749	-0.179
mtca.au	-234.875	31.773	-7.392	0.000	-234.875	-0.173
ning au	-26 647	15 896	-1 676	0 094	-26 647	-0 017
ping.au	-116 921	39 015	-2 997	0.003	-116 921	
pinj.au	140 270	11 535	2.557	0.003	140.270	-0.075
potrok.ar	-140.270	30 672	-3.130	0.002	19 512	-0.090
saye.us	212 472	10 621	16 026	0.040	212 472	0.012
sedg.us	-313.4/2	10.031	-10.020	0.000	-313.472	-0.201
sereng.tz	-81.449	36.144	-2.253	0.024	-81.449	-0.052
sgs.us	-300.238	20.071	-14.959	0.000	-300.238	-0.192
sier.us	-32.595	19.6/4	-1.65/	0.098	-32.595	-0.021
smith.us	-5.798	11.546	-0.502	0.616	-5./98	-0.004
spin.us	-29.564	23.915	-1.236	0.216	-29.564	-0.019
spv.ar	-281.364	28.890	-9.739	0.000	-281.364	-0.180
sva⊥.no	-328.181	23.420	-14.013	0.000	-328.181	-0.210
tyso.us	51.448	20.623	2.495	0.013	51.448	0.038
ukul.za	-7.752	30.396	-0.255	0.799	-7.752	-0.005
unc.us	-76.002	31.007	-2.451	0.014	-76.002	-0.049
uwo.ca	-24.238	19.977	-1.213	0.225	-24.238	-0.016
valm.ch	-6.354	16.735	-0.380	0.704	-6.354	-0.004
yarra.au	-278.781	25.139	-11.089	0.000	-278.781	-0.203
llegu_mass	87.495	12.545	6.975	0.000	87.495	0.575
llegu_mass2	-13.635	2.723	-5.007	0.000	-13.635	-0.328
lgram_mass	46.267	19.356	2.390	0.017	46.267	0.259
lgram_mass2	-6.049	2.256	-2.682	0.007	-6.049	-0.292
lforb_mass	2.695	8.153	0.331	0.741	2.695	0.023
lforb_mass2	0.734	1.482	0.495	0.620	0.734	0.035
pd_gr ~						
arch.us	126.265	7.261	17.388	0.000	126.265	0.255
azi.cn	103.595	3.415	30.339	0.000	103.595	0.209
barta.us	84.251	7.450	11.308	0.000	84.251	0.170
bnch.us	90.642	5.537	16.369	0.000	90.642	0.183
bogong.au	94.987	4.785	19.852	0.000	94.987	0.192
bttr.us	65.027	15.481	4.200	0.000	65.027	0.131
bunya.au	17.074	7.374	2.315	0.021	17.074	0.035
burrawan.au	74.050	7.175	10.321	0.000	74.050	0.150
burren.ie	93.729	3.628	25.836	0.000	93.729	0.189
cbgb.us	-2.851	3.980	-0.716	0.474	-2.851	-0.008
cdcr.us	58.341	7.221	8.080	0.000	58.341	0.151
cdpt.us	97.962	4.102	23.883	0.000	97.962	0.277
chilcastar	96.958	7.609	12.743	0.000	96.958	0,196
cowi ca	-8 0UE	3 120	_2 550	0 010	-8 00E	_0 018
doane us	20.900	5.40U 6 070	6 /11	0.010	30 061	0 110
abane.us	31 052	0.070	2 606	0.000	31 053	0.110
yail.it	31.053	0.4UL 0.111	3.090 25 600	0.000	31.033	0.005
giib.za	233.808	9.111	23.009	0.000	233.808	0.51/
giac.us	46.714	/.930	5.890	0.000	46.714	0.094
gicr.us	11.637	3.270	3.558	0.000	11.637	0.024
hall.us	-1.820	6.465	-0.282	0.778	-1.820	-0.004
hast.us	-0.987	4.392	-0.225	0.822	-0.987	-0.002
hnvr.us	15.321	7.448	2.057	0.040	15.321	0.031
kibber.in	69.491	32.627	2.130	0.033	69.491	0.140
kidman.au	16.387	5.563	2.946	0.003	16.387	0.042
kirik.ee	86.382	3.101	27.853	0.000	86.382	0.175
koffler.ca	-1.239	7.029	-0.176	0.860	-1.239	-0.003
konz.us	93.292	3.874	24.079	0.000	93.292	0.189
lead.us	-12.122	6.229	-1.946	0.052	-12.122	-0.024
look.us	84.104	6.788	12.391	0.000	84.104	0.170
marc.ar	45.519	10.127	4.495	0.000	45.519	0.092
mcla.us	-6.364	3.722	-1.710	0.087	-6.364	-0.013
mtca.au	58.037	4.776	12.151	0.000	58.037	0.135
ping.au	-4.910	3.945	-1.245	0.213	-4.910	-0.010
pini.au	7.660	3,979	1.925	0.054	7.660	0.015
potrokar	98.880	5.353	18,473	0.000	98.880	0.200
POLLOX. al	_4 579	25 592	_0 170	0 858	_4 579	_0 000
seda us	-4.210 -4.210	4 103	-2 288	0.000	-4.210	_0 019
seuy.us	130 500	4.1US	-2.200	0.022	139 500	0 202
sereng.tz	139.300	0.040	20.300	0.000	139.300	0.202
sgs.us	/0.555	5.2/0	13.388	0.000	/0.555	0.143
sier.us	-4.796	2.453	-1.955	0.051	-4.796	-0.010
smith.us	0.982	2.419	0.406	0.685	0.982	0.002
spin.us	0.041	3.600	0.011	0.991	0.041	0.000
spv.ar	29.834	5.836	5.112	0.000	29.834	0.060

sval.no	9.443	8.507	1.110	0.267	9.443	0.019	
tyso.us	16.759	3.053	5.489	0.000	16.759	0.039	
ukul.za	25.976	5.493	4.729	0.000	25.976	0.052	
unc.us	54.150	3.353	16.149	0.000	54.150	0.109	
uwo.ca	71.494	5.882	12.154	0.000	71.494	0.144	
valm.ch	65.894	8.077	8.158	0.000	65.894	0.133	
yarra.au	27.816	7.580	3.670	0.000	27.816	0.064	
llegu_mass	1.448	2.863	0.506	0.613	1.448	0.030	
llegu_mass2	-0.711	0.670	-1.062	0.288	-0.711	-0.054	
lgram_mass	60.370	16.221	3.722	0.000	60.370	1.064	
lgram_mass2	-5.999	1.625	-3.691	0.000	-5.999	-0.912	
lforb_mass	-1.311	1.926	-0.681	0.496	-1.311	-0.035	
lforb_mass2	0.455	0.368	1.237	0.216	0.455	0.069	
pd_fr_~							
arch.us	127.521	54.333	2.347	0.019	127.521	0.043	
azı.cn	918.200	62.714	14.641	0.000	918.200	0.310	
barta.us	-202.805	50.516	-4.015	0.000	-202.805	-0.068	
bnch.us	-315.646	53.835	-5.863	0.000	-315.646	-0.107	
bogong.au	362.926	57.345	6.329	0.000	362.926	0.122	
bttr.us	1/2.306	/0.9/0	2.428	0.015	1/2.306	0.058	
bunya.au	-288.9/3	50.419	-5./31	0.000	-288.9/3	-0.098	
burrawan.au	-/4.968	61.663	-1.216	0.224	-/4.968	-0.025	
burren.le	681.302	69.9/4	9./36	0.000	681.302	0.230	
cbgb.us	-30/.515	44.943	-6.842	0.000	-307.515	-0.138	
cdcr.us	-254.072	55.103	-4.611	0.000	-254.072	-0.110	
capt.us	4.627	51.923	0.089	0.929	4.627	0.002	
chilcas.ar	-127.715	58.442	-2.185	0.029	-127.715	-0.043	
cowi.ca	-625.668	60.962	-10.263	0.000	-625.668	-0.211	
doane.us	-430.319	45.488	-9.460	0.000	-430.319	-0.204	
gall.it	143.470	64.852	2.212	0.027	143.470	0.050	
gilb.za	302.551	50.035	6.047	0.000	302.551	0.112	
glac.us	-294.911	58.852	-5.011	0.000	-294.911	-0.100	
glcr.us	-426.728	55.381	-7.705	0.000	-426.728	-0.144	
hall.us	-277.841	54.830	-5.067	0.000	-277.841	-0.094	
hast.us	-291.048	56.756	-5.128	0.000	-291.048	-0.098	
hnvr.us	-47.572	51.830	-0.918	0.359	-47.572	-0.016	
kibber.in	-449.213	62.642	-7.171	0.000	-449.213	-0.152	
kidman.au	-294.409	45.831	-6.424	0.000	-294.409	-0.128	
kirik.ee	616.929	55.247	11.167	0.000	616.929	0.208	
koffler.ca	-285.460	51.414	-5.552	0.000	-285.460	-0.105	
konz.us	-129.964	47.085	-2.760	0.006	-129.964	-0.044	
lead.us	-262.759	55.161	-4.764	0.000	-262.759	-0.089	
look.us	-261.636	54.244	-4.823	0.000	-261.636	-0.088	
marc.ar	-394.407	54.082	-7.293	0.000	-394.407	-0.133	
mcla.us	-51/.366	61.96/	-8.349	0.000	-51/.366	-0.1/5	
mtca.au	-2.199	52.771	-0.042	0.967	-2.199	-0.001	
ping.au	-292.183	48.8//	-5.9/8	0.000	-292.183	-0.099	
pinj.au	-364.769	56.976	-6.402	0.000	-364.769	-0.123	
potrok.ar	128./42	54.985	2.341	0.019	128./42	0.043	
sage.us	139.045	40 202	2.090	0.037	139.045	0.04/	
sedg.us	-339.543	48.202	-/.044	0.000	-339.543	-0.115	
sereng.tz	-230.109	40.401 E0 266	-4.914	0.000	-230.109	-0.080	
sys.us	-2/5.010	50.300	-3.472	0.000	-2/3.010	-0.093	
sith us	120 121	50 947	2 550	0.000	120 121	-0.050	
spin us	_307 133	41 642	-7 376	0.010	_307 133	_0 104	
spin.us	-386 775	41.042	-8 0/8	0.000	-386 775	-0.131	
spv.ar	-300.773	56 003	1 644	0.000	92 0/1	-0.131	
	-284 182	54 761	-5 190	0.100	-284 182	-0.031	
	406.380	57.857	7.024	0.000	406.380	0.137	
	_142 191	57 944	-2 454	0.000	_142 191	_0 048	
	-362 321	53 545	-6 767	0 000	-362 321	_0 122	
valm.ch	858.360	67.527	12.711	0.000	858.360	0.290	
varra, au	-238.978	47.710	-5.009	0.000	-238.978	-0.092	
llegu mass	15.023	16.687	0.900	0.368	15.023	0.052	
llegu mass2	-4.301	4.016	-1.071	0.284	-4.301	-0.054	
lgram mass	-4.608	23.052	-0.200	0.842	-4.608	-0.014	
lgram mass2	-3.705	2.696	-1.374	0.169	-3.705	-0.094	
lforb mass	42.396	11.848	3.578	0.000	42.396	0.190	
lforb mass2	-4.741	2.202	-2.153	0.031	-4.741	-0.120	
llitter mass ~							
arch.us	-1.988	0.173	-11.500	0.000	-1.988	-0.130	
azi.cn	-1.323	0.155	-8.527	0.000	-1.323	-0.086	
barta.us	-0.484	0.157	-3.092	0.002	-0.484	-0.032	
bnch.us	-1.236	0.151	-8.206	0.000	-1.236	-0.081	

bogong.au	-2.119	0.176	-12.021	0.000	-2.119	-0.138	
bttr us	-0.620	0 169	-3 672	0 000	-0 620	_0 040	
bunya au	_0 272	0 286	_0 951	0 342	-0.272	_0 018	
burrawan au	2 720	0.200	0.775	0.042	2 720	-0.010	
bullawallau	-2.720	0.270	-9.775	0.000	-2.720	-0.178	
burren.1e	-0.459	0.152	-3.030	0.002	-0.459	-0.030	
cbgb.us	0.669	0.134	5.009	0.000	0.669	0.058	
cdcr.us	-2.021	0.154	-13.090	0.000	-2.021	-0.169	
cdpt.us	-5.490	0.169	-32.584	0.000	-5.490	-0.502	
chilcas.ar	0.467	0.129	3.627	0.000	0.467	0.031	
cowi ca	-1.870	0.176	-10.635	0.000	-1.870	-0.122	
doane us	_0 282	0 139	_2 022	0 043	_0 282	_0 026	
	-0.202	0.155	-2.022	0.043	-0.202	-0.020	
gall.it	-2.161	0.287	-/.521	0.000	-2.161	-0.146	
gilb.za	-5.829	0.181	-32.271	0.000	-5.829	-0.416	
glac.us	-0.653	0.132	-4.953	0.000	-0.653	-0.043	
glcr.us	0.449	0.125	3.602	0.000	0.449	0.029	
hall.us	0.115	0.167	0.689	0.491	0.115	0.008	
hastus	-0.643	0.170	-3.771	0.000	-0.643	-0.042	
have us	1 102	0.100	5 5 7 1	0.000	1 102	0.072	
hibben in	-1.102	0.199	-3.534	0.000	-1.102	-0.072	
kibber.in	-4.418	0.253	-1/.434	0.000	-4.418	-0.288	
kidman.au	-1.784	0.262	-6.810	0.000	-1.784	-0.149	
kirik.ee	-0.698	0.191	-3.654	0.000	-0.698	-0.046	
koffler.ca	-0.146	0.148	-0.987	0.324	-0.146	-0.010	
konz.us	-5.825	0.223	-26.176	0.000	-5.825	-0.380	
lead.us	0.394	0.158	2,491	0.013	0.394	0,026	
lock	1 600	0.10	_7 040	0.010	1 600	_0 110	
	-1.009	0.212	-1.949	0.000	-1.009	-0.110	
marc.ar	0.196	0.198	0.990	0.322	0.196	0.013	
mcla.us	0.118	0.137	0.865	0.387	0.118	0.008	
mtca.au	-5.310	0.153	-34.739	0.000	-5.310	-0.399	
ping.au	-2.318	0.195	-11.874	0.000	-2.318	-0.151	
pini.au	-0.569	0.145	-3.926	0.000	-0.569	-0.037	
potrok ar	-1.624	0.163	-9.962	0.000	-1.624	-0.106	
	1 000	0.105	= J • J 0 Z	0.000	1 000	0 072	
sage.us	-1.099	0.215	-5.120	0.000	-1.099	-0.072	
sedg.us	-0.076	0.150	-0.503	0.615	-0.076	-0.005	
sereng.tz	-3.777	0.153	-24.766	0.000	-3.777	-0.247	
sgs.us	-5.669	0.201	-28.209	0.000	-5.669	-0.370	
sier.us	-1.138	0.216	-5.263	0.000	-1.138	-0.074	
smithuus	-1.202	0.128	-9.366	0.000	-1.202	-0.078	
spin us	_0 571	0 118	_1 816	0 000	_0 571	_0 037	
spin.us	-0.571	0.110	-1.010	0.000	-0.571	-0.037	
spv.ar	-0.697	0.189	-3.0/9	0.000	-0.697	-0.045	
sval.no	-1.305	0.191	-6.834	0.000	-1.305	-0.085	
tyso.us	-2.937	0.188	-15.642	0.000	-2.937	-0.221	
ukul.za	-5.220	0.174	-30.017	0.000	-5.220	-0.341	
unc.us	-1.241	0.216	-5.754	0.000	-1.241	-0.081	
uwo.ca	-1.336	0.146	-9.169	0.000	-1.336	-0.087	
valm ch	_3 332	0 245	_13 613	0 000	_3 332	_0 218	
	-5.552	0.245	-13.015	0.000	-5.552	-0.210	
yarra.au	-0.524	0.144	-3.020	0.000	-0.524	-0.039	
pd_fr	-0.000	0.000	-2./1/	0.007	-0.000	-0.052	
pd_gr	0.003	0.000	6.191	0.000	0.003	0.096	
pd_lg	-0.000	0.000	-2.007	0.045	-0.000	-0.031	
llegu mass	-0.141	0.070	-2.001	0.045	-0.141	-0.094	
llegu mass2	0.024	0.016	1.508	0.132	0.024	0.059	
laram mass	0.248	0.125	1.978	0.048	0.248	0.141	
lgram magg2	-0.001	0 015	_0 097	0 021	-0.001	_0 006	
	-0.001	0.013	-0.087	0.931	-0.001	-0.000	
liorb_mass	-0.098	0.043	-2.290	0.022	-0.098	-0.085	
lforb_mass2	0.027	0.008	3.312	0.001	0.027	0.130	
Intercepts:							
	Estimate	Std.Err	z-value	P(> z)	Std.lv	Std.all	
.pd la	212.688	49.512	4.296	ົວ່ວວ່ວ	212.688	1.034	
nd ar	254 527	39 560	6 434	0 000	254 527	3,903	
. ba ta	007 500	71 115	10 040	0.000	007 500	2.203	
•pd_11	927.532	/1.115	13.043	0.000	927.532	2.3/0	
.llitter_mass	3.564	0.319	11.158	0.000	3.564	1.765	
Variances:							
	Estimate	Std.Err	z-value	P(> z)	Std.lv	Std.all	
.pd lg	16158.872	683.122	23,654	0,000	16158-872	0.382	
nd ar	1452 190	220 964	6.572	0 000	1452 190	0.342	
. 50 ⁻²¹	33655 664	1406 007	22 212	0.000	1-1-1-1-1-1-1-1-1-1-1-1-1-1-1-1-1-1-1-	0.342	
•pu_11	52055.004	1400.93/	23.210	0.000	52055.004	0.214	
.111tter_mass	0.492	0.043	11.433	0.000	0.492	0.121	
R-Square:							

Estimate
0.618
0.658
0.786

Fit measures: cfi rmsea srmr 1.000 0.029 0.000

D.4.2 Final species richness model

lavaan (0.5-23.1097) converged normally after 460 iterations

Number of obser	rvations					
Estimator Minimum Functic Degrees of free D walue (Chi an	on Test Stat edom	istic		ML 28.114 3		
P-value (Boller	n-Stine Boot	strap)		1.000		
Parameter Estimat	ces:					
Information				Observed		
Standard Errors	3		В	ootstrap		
Number of reque	ested bootst	rap draws		1000		
Number of succe	essful boots	trap draw	S	1000		
Regressions:						
Regressions.	Estimate	Std.Err	z-value	P(> z)	Std.lv	Std.all
sr_lg ~						
arch.us	-1.554	0.082	-18.866	0.000	-1.554	-0.165
azi.cn	0.957	0.141	6.793	0.000	0.957	0.102
barta.us	-1.313	0.099	-13.276	0.000	-1.313	-0.140
bnch.us	-0.970	0.117	-8.307	0.000	-0.970	-0.103
bogong.au	-1.574	0.083	-18.876	0.000	-1.574	-0.167
bttr.us	-1.176	0.106	-11.131	0.000	-1.176	-0.125
bunya.au	-1.181	0.120	-9.822	0.000	-1.181	-0.126
burrawan.au	-1.557	0.082	-18.933	0.000	-1.557	-0.166
burren.ie	1.817	0.204	8.895	0.000	1.817	0.193
cbgb.us	-1.291	0.108	-11.931	0.000	-1.291	-0.183
cdcr.us	-1.346	0.099	-13.601	0.000	-1.346	-0.184
cdpt.us	-1.285	0.106	-12.088	0.000	-1.285	-0.191
chilcas.ar	-1.241	0.112	-11.090	0.000	-1.241	-0.132
dooro ug	1.082	0.10/	0.4/0	0.000	1.082	0.115
apil it	-1.100	0.112	-10.370	0.000	-1.100	-0.173
gall.ll gilb ga	-1.122	0.131	-0.373	0.000	-1.122	-0.123
glac us	_1 182	0.122	_10 944	0.000	_1 182	-0.126
glac.us	-1.182	0.108	-10.944	0.000	-1.182	-0.120
ball ug	-1.109	0.118	17 220	0.000	-1.109	-0.120
hast us	-0.552	0.050	-3 737	0.000	-0 552	-0.059
hnyr us	-0.545	0 137	-3 975	0.000	-0.545	-0.059
kibber in	-1.167	0.149	-7.854	0.000	-1.167	-0.124
kidman.au	-0.900	0.120	-7.499	0.000	-0.900	-0.123
kirik.ee	2.161	0.204	10.607	0.000	2.161	0.230
koffler.ca	-1.053	0.094	-11.138	0.000	-1.053	-0.122
konz.us	-0.539	0.162	-3.336	0.001	-0.539	-0.057
lead.us	-1.676	0.100	-16.732	0.000	-1.676	-0.178
look.us	-1.486	0.106	-13.992	0.000	-1.486	-0.158
marc.ar	-1.376	0.147	-9.356	0.000	-1.376	-0.146
mcla.us	-1.429	0.083	-17.114	0.000	-1.429	-0.152
mtca.au	-1.389	0.108	-12.868	0.000	-1.389	-0.170
ping.au	-0.276	0.119	-2.317	0.020	-0.276	-0.029
pinj.au	-0.785	0.191	-4.116	0.000	-0.785	-0.083
potrok.ar	-1.041	0.147	-7.098	0.000	-1.041	-0.111
sage.us	-0.747	0.106	-7.035	0.000	-0.747	-0.079
sedg.us	-1.604	0.084	-19.054	0.000	-1.604	-0.171
sereng.tz	-0.922	0.123	-7.505	0.000	-0.922	-0.098
sgs.us	-1.519	0.085	-17.835	0.000	-1.519	-0.161
sier.us	-0.578	0.126	-4.597	0.000	-0.578	-0.061

smith.us	1.495	0.196	7.623	0.000	1.495	0.159	
spin.us	-0.560	0.118	-4.730	0.000	-0.560	-0.060	
spv.ar	-1.486	0.103	-14.389	0.000	-1.486	-0.158	
sval.no	-1.689	0.102	-16.609	0.000	-1.689	-0.180	
tyso.us	-0.599	0.096	-6.260	0.000	-0.599	-0.073	
ukul.za	0.576	0.251	2.296	0.022	0.576	0.061	
unc.us	-1.034	0.095	-10.852	0.000	-1.034	-0.110	
uwo.ca	-0.399	0.160	-2.489	0.013	-0.399	-0.042	
valm.ch	0.593	0.188	3.160	0.002	0.593	0.063	
varra.au	-1.458	0.094	-15.543	0.000	-1.458	-0.176	
llegu mass	0.315	0.061	5.166	0.000	0.315	0.343	
llegu mass2	-0.037	0.016	-2.364	0.018	-0.037	-0.147	
lgram mass	0.189	0.057	3,317	0.001	0.189	0.175	
lgram mass2	-0.025	0.007	-3.511	0.000	-0.025	-0.196	
lforb mass	0.017	0.034	0 / 97	0 619	0 017	0 024	
lforb mass?	0.01/	0.007	0.457	0.506	0.017	0.024	
er ar ~	0.001	0.007	0.000	0.500	0.004	0.035	
si_gi	2 200	0 166	5 1 2 7	0 000	2 200	0 127	
	2.390	0.400	10 020	0.000	2.390	0.127	
	3.935	0.300	10.920	0.000	3.935	0.209	
barta.us	2.418	0.365	0.030	0.000	2.418	0.129	
bncn.us	0.106	0.354	0.299	0.765	0.106	0.006	
bogong.au	0.969	0.306	3.165	0.002	0.969	0.052	
bttr.us	-0.770	0.377	-2.043	0.041	-0.770	-0.041	
bunya.au	-2.189	0.321	-6.825	0.000	-2.189	-0.116	
burrawan.au	0.940	0.349	2.693	0.007	0.940	0.050	
burren.ie	4.262	0.369	11.553	0.000	4.262	0.227	
cbgb.us	-2.173	0.280	-7.765	0.000	-2.173	-0.154	
cdcr.us	0.510	0.311	1.641	0.101	0.510	0.035	
cdpt.us	0.361	0.299	1.205	0.228	0.361	0.027	
chilcas.ar	3.429	0.589	5.821	0.000	3.429	0.182	
cowi.ca	-3.175	0.256	-12.400	0.000	-3.175	-0.169	
doane.us	-1.527	0.263	-5.803	0.000	-1.527	-0.114	
gall.it	0.142	0.331	0.429	0.668	0.142	0.008	
gilb.za	5.242	0.337	15.542	0.000	5.242	0.305	
glac.us	-2.019	0.289	-6.982	0.000	-2.019	-0.107	
alcr.us	-1.655	0.269	-6.150	0.000	-1.655	-0.088	
hall.us	-1.770	0.287	-6.170	0.000	-1.770	-0.094	
hast us	_1 329	0 309	-4 295	0 000	_1 329	-0 071	
hase us	-2 019	0 285	-7 073	0.000	-2 019	-0 107	
kibber in	_1 900	0 387	-4 907	0.000	_1 900	-0.101	
kidman au	1 990	0.307		0.000	1 990	0 120	
kirik oo	-1.000	0.275	5 734	0.000	2 412	0 120	
kilik.ee	2.412	0.421	10 210	0.000	2.412	0.120	
kong ug	-2.000	0.260	-10.210	0.000	-2.000	-0.100	
konz.us	2 2 2 2 1	0.351	10.309	0.000	2 2 2 2 1	0.194	
lead.us	-3.231	0.308	-10.504	0.000	-3.231	-0.172	
look.us	-1.0/9	0.324	-3.329	0.001	-1.0/9	-0.057	
marc.ar	2.345	0.3//	6.223	0.000	2.345	0.125	
mcla.us	-1.//9	0.310	-5./35	0.000	-1.//9	-0.095	
mtca.au	-0.5/1	0.288	-1.983	0.04/	-0.5/1	-0.035	
ping.au	-1.11/	0.328	-3.411	0.001	-1.11/	-0.059	
pinj.au	-1.612	0.298	-5.404	0.000	-1.612	-0.086	
potrok.ar	1.122	0.339	3.312	0.001	1.122	0.060	
sage.us	-2.158	0.333	-6.485	0.000	-2.158	-0.115	
sedg.us	-1.153	0.298	-3.875	0.000	-1.153	-0.061	
sereng.tz	4.731	0.327	14.473	0.000	4.731	0.252	
sgs.us	-1.344	0.315	-4.262	0.000	-1.344	-0.072	
sier.us	-0.744	0.323	-2.305	0.021	-0.744	-0.040	
smith.us	1.480	0.275	5.387	0.000	1.480	0.079	
spin.us	-1.492	0.224	-6.668	0.000	-1.492	-0.079	
spv.ar	0.676	0.319	2.117	0.034	0.676	0.036	
sval.no	-2.737	0.314	-8.730	0.000	-2.737	-0.146	
tyso.us	-0.301	0.274	-1.100	0.272	-0.301	-0.018	
ukul.za	2.041	0.371	5.496	0.000	2.041	0.109	
unc.us	0.751	0.345	2.174	0.030	0.751	0.040	
uwo.ca	-0.074	0.320	-0.231	0.817	-0.074	-0.004	
valm.ch	0.804	0.345	2.328	0.020	0.804	0.043	
yarra.au	-0.604	0.312	-1.934	0.053	-0.604	-0.037	
llegu mass	0.218	0.121	1.790	0.073	0.218	0.119	
llegu mass2	-0.043	0.029	-1.480	0.139	-0.043	-0.086	
lgram mass	0.669	0.163	4.112	0.000	0.669	0.311	
lgram mass2	-0.068	0.018	-3.708	0.000	-0.068	-0.271	
lforb mass	0.125	0.077	1.636	0.102	0.125	0.089	
lforb mass?	-0.007	0.014	-0.483	0.629	-0.007	-0.027	
sr fr ~	3.007	0.011	3.105	0.025	0.007		
arch.us	1.500	0.523	2,871	0.004	1.500	0.040	

azı.cn	16.873	0.661	25.521	0.000	16.873	0.453	
barta.us	-1.173	0.494	-2.376	0.017	-1.173	-0.032	
bnch us	-1.885	0.472	-3,991	0.000	-1.885	-0.051	
bogong av	E 260	0 577	0 202	0.000	E 260	0 144	
bogong.au	5.309	0.5//	9.302	0.000	5.309	0.144	
bttr.us	1.804	0.725	2.489	0.013	1.804	0.048	
bunva.au	-2.825	0.487	-5.806	0.000	-2.825	-0.076	
hurrawan au	-2 015	0 437	-4 611	0 000	-2 015	-0 054	
bullawan.au	-2.015	0.457	-4.011	0.000	-2.015	-0.034	
burren.1e	6.008	0./01	8.5/4	0.000	6.008	0.161	
cbgb.us	-0.614	0.481	-1.277	0.202	-0.614	-0.022	
cdcr.us	-1.056	0.482	-2.190	0.029	-1.056	-0.036	
cdpt us	1 119	0 503	2 225	0 026	1 119	0 042	
cape.us	1.050	0.505	1 000	0.020	1.050	0.042	
chilcas.ar	-1.052	0.554	-1.899	0.058	-1.052	-0.028	
cowi.ca	-4.239	0.452	-9.380	0.000	-4.239	-0.114	
doane.us	-3.385	0.415	-8.148	0.000	-3.385	-0.127	
all it	1 265	0 609	2 078	0 038	1 265	0 035	
	1.205	0.009	2.070	0.000	1.205	0.055	
gilb.za	4.668	0.594	/.860	0.000	4.668	0.13/	
glac.us	-3.133	0.451	-6.950	0.000	-3.133	-0.084	
glcr.us	-3.272	0.444	-7.373	0.000	-3.272	-0.088	
hall ug	_0 920	0 525	_1 752	0 080	_0 920	-0 025	
hast us	-0.520	0.525	-1./JZ	0.000	2 0 6 0	-0.025	
nast.us	-2.960	0.502	-5.899	0.000	-2.960	-0.080	
hnvr.us	-0.219	0.512	-0.427	0.669	-0.219	-0.006	
kibber.in	-4.047	0.573	-7.068	0.000	-4.047	-0.109	
kidman au	-2 834	0 413	-6 862	0 000	-2 834	_0 098	
himil oo	14 010	0.413		0.000	14 010	0.077	
kirik.ee	14.019	0.64/	21.0/3	0.000	14.019	0.377	
koffler.ca	-1.109	0.563	-1.968	0.049	-1.109	-0.033	
konz.us	1.317	0.476	2.769	0.006	1.317	0.035	
lead.us	-3.211	0.507	-6.327	0.000	-3.211	-0.086	
look ya	1 7/1	0 507	3 31E	0 001	1 7/1	0.047	
look.us	-1./41	0.525	-3.315	0.001	-1./41	-0.04/	
marc.ar	-3.224	0.548	-5.884	0.000	-3.224	-0.087	
mcla.us	-3.679	0.397	-9.263	0.000	-3.679	-0.099	
mtca.au	3.314	0.504	6.578	0.000	3.314	0.103	
ning ou	2 7 4 4	0 110	0 260	0 000	2 7 / /	0 101	
ping.au	-3./44	0.440	-0.300	0.000	-3./44	-0.101	
pinj.au	-3.049	0.420	-/.256	0.000	-3.049	-0.082	
potrok.ar	2.222	0.534	4.165	0.000	2.222	0.060	
sage.us	1.183	0.603	1.964	0.050	1.183	0.032	
sedg.us	-3.081	0.460	-6.693	0.000	-3.081	-0.083	
sereng tz	-2 285	0 430	-5 314	0 000	-2 285	-0 061	
Sereng. cz	2.205	0.440	-J.JI4	0.000	2.205	-0.001	
sgs.us	-2.250	0.442	-5.103	0.000	-2.250	-0.061	
sier.us	-2.924	0.460	-6.361	0.000	-2.924	-0.079	
smith.us	4.955	0.564	8.791	0.000	4.955	0.133	
spin.us	-2.721	0.392	-6.940	0.000	-2.721	-0.073	
sny ar	_3 221	0 434	-7 430	0 000	_3 221	_0 087	
	1 750	0 5 2 1	2 272	0.000	1 750	0.047	
sval.no	-1.759	0.521	-3.3/3	0.001	-1.759	-0.047	
tyso.us	-2.38/	0.432	-5.521	0.000	-2.38/	-0.0/4	
ukul.za	6.487	0.671	9.671	0.000	6.487	0.174	
unc.us	-1.198	0.507	-2.365	0.018	-1.198	-0.032	
uwo, ca	-2.046	0.557	-3.676	0.000	-2.046	-0.055	
walm ch	10 400	0 645	16 127	0 000	10 100	0 200	
vaim.cn	10.409	0.045	10.127	0.000	10.409	0.280	
yarra.au	-2.215	0.424	-3.224	0.000	-2.215	-0.068	
llegu_mass	0.261	0.166	1.573	0.116	0.261	0.072	
llegu mass2	-0.043	0.041	-1.061	0.289	-0.043	-0.043	
lgram mass	-0.025	0.205	-0.123	0.902	-0.025	-0.006	
lgram mass2	-0.029	0.024	-1.195	0.232	-0.029	-0.058	
lforb mage	0 433	0 100	2 963	0 000	0 133	0 154	
	0.455	0.109	3.905	0.000	0.435	0.134	
liorb_mass2	-0.026	0.023	-1.160	0.246	-0.026	-0.053	
litter_mass ~							
arch.us	-1.587	0.166	-9.574	0.000	-1.587	-0.104	
azi.cn	-0.888	0.194	-4.585	0.000	-0.888	-0.058	
barta ug	0 251	0 157	1 602	0 100	0 251	0 016	
Darca.us	-0.231	0.137	-1.002	0.109	-0.231	-0.010	
bnch.us	-0.902	0.14/	-6.154	0.000	-0.902	-0.059	
bogong.au	-1.675	0.170	-9.874	0.000	-1.675	-0.109	
bttr.us	-0.320	0.180	-1.777	0.076	-0.320	-0.021	
bunya au	-0 061	0 282	_0 215	0 830	-0 061	-0 004	
burration of	2 171	0.202	0 501	0.000	2 171	0 1/1	
burrawall.au	-2.4/1	0.288	-0.001	0.000	-2.4/1	-0.101	
burren.ie	-0.392	0.163	-2.405	0.016	-0.392	-0.026	
cbgb.us	0.913	0.142	6.449	0.000	0.913	0.079	
cdcr.us	-1.742	0.149	-11.688	0.000	-1.742	-0.146	
cdpt.us	-5.094	0.165	-30.818	0.000	-5.094	-0.466	
childas ar	0 669	0 1/1	4 751	0 000	0 669	0 0//	
	1 704	0 100	/JI	0.000	1 704	0.116	
cow1.ca	-1./84	0.182	-9.805	0.000	-1./84	-0.110	
doane.us	-0.026	0.146	-0.180	0.857	-0.026	-0.002	
gall.it	-2.007	0.312	-6.439	0.000	-2.007	-0.135	
gilb.za	-5.322	0.166	-31.971	0.000	-5.322	-0.380	
glac.us	-0.392	0.139	-2.817	0.005	-0.392	-0.026	
alor ye	0 6/1	0 120	1 020	0 000	0 6/1	0 042	
grer . us	0.041	0.120	7.740	0.000	0.041	0.042	

]

hall.us	0.355	0.173	2.047	0.041	0.355	0.023	
hagt ug	-0 585	0 171	-3 414	0 001	-0 585	-0.038	
howr ug	-0.935	0 211	-4 422	0.001	-0.935	-0.061	
kibbor in	-0.555	0.211	1/ 020	0.000	4 063	-0.001	
kidman au	1 624	0.272	-14.000	0.000	1 624	-0.205	
kiuman.au	-1.024	0.200	-0.251	0.000	-1.024	-0.130	
kirik.ee	-0.283	0.222	-1.2/4	0.203	-0.283	-0.018	
koiiler.ca	0.052	0.155	0.33/	0.736	0.052	0.004	
konz.us	-5.633	0.231	-24.429	0.000	-5.633	-0.368	
lead.us	0.600	0.170	3.528	0.000	0.600	0.039	
look.us	-1.267	0.206	-6.138	0.000	-1.267	-0.083	
marc.ar	0.303	0.207	1.462	0.144	0.303	0.020	
mcla.us	0.308	0.138	2.236	0.025	0.308	0.020	
mtca.au	-4.900	0.157	-31.278	0.000	-4.900	-0.368	
ping.au	-2.316	0.204	-11.373	0.000	-2.316	-0.151	
pinj.au	-0.426	0.145	-2.931	0.003	-0.426	-0.028	
potrok.ar	-1.280	0.165	-7.772	0.000	-1.280	-0.084	
sage.us	-0.999	0.230	-4.344	0.000	-0.999	-0.065	
sedg.us	0.062	0.155	0.400	0.689	0.062	0.004	
sereng.tz	-3.557	0.160	-22.240	0.000	-3.557	-0.232	
sgs.us	-5.280	0.198	-26.610	0.000	-5.280	-0.345	
sier.us	-1.118	0.214	-5.236	0.000	-1.118	-0.073	
smith.us	-1.162	0.149	-7.793	0.000	-1.162	-0.076	
spin.us	-0.491	0.125	-3.915	0.000	-0.491	-0.032	
spv.ar	-0.533	0.193	-2.763	0.006	-0.533	-0.035	
sval.no	-1.052	0.209	-5.037	0.000	-1.052	-0.069	
tyso.us	-2.882	0.185	-15.543	0.000	-2.882	-0.217	
ukul.za	-5.135	0.185	-27.761	0.000	-5.135	-0.335	
unc.us	-1.075	0.203	-5.304	0.000	-1.075	-0.070	
uwo.ca	-1.078	0.131	-8.205	0.000	-1.078	-0.070	
valm.ch	-3.058	0.235	-13.036	0.000	-3.058	-0.200	
varra, au	-0.312	0.153	-2.036	0.042	-0.312	-0.023	
sr fr	-0.034	0.009	-3.788	0.000	-0.034	-0.083	
sr ar	0.047	0.014	3,293	0.001	0.047	0.058	
sr la	0.019	0.032	0.599	0.549	0.019	0.012	
llegu mass	-0 174	0 071	-2 / 39	0 015	-0.174	-0 116	
llegu mass?	0 029	0.016	1 742	0.081	0 029	0 070	
laram magg	0.378	0 126	3 013	0.001	0.025	0.215	
lgram magg2	0.014	0.120	0 9/3	0.005	0.014	0.215	
lforb mage	-0.014	0.014	-0.943	0.340	-0.014	-0.007	
lforb maga2	-0.105	0.044	-2.307	0.010	-0.105	-0.091	
1101D_massz	0.028	0.008	3.411	0.001	0.028	0.139	
Tatomaanta							
intercepts:	Datimata				C+-1]	C+4 all	
an 1 a	ESTIMATE	Stu.Err	z-value	P(z)	5tu.IV		
.sr_ig	1.148	0.158	7.249	0.000	1.148	0.926	
.sr_gr	2.520	0.462	5.45/	0.000	2.520	1.018	
•sr_tr	5.550	0.630	8.809	0.000	5.550	1.132	
.llitter_mass	4.056	0.346	11./22	0.000	4.056	2.009	
variances:		a. 1 5	7		a	a. 1 11	
	Estimate	Std.Err	z-value	P(> z)	Std.lv	Std.all	
.sr_lg	0.324	0.01/	18.803	0.000	0.324	0.211	
.sr_gr	1.344	0.066	20.293	0.000	1.344	0.219	
.sr_fr	3.431	0.149	22.977	0.000	3.431	0.143	
.llitter_mass	0.503	0.042	12.032	0.000	0.503	0.123	
D. G							
k-square:	Datimat						
1.	Estimate						
sr_ig	0./89						
sr_gr	0./81						
sr_ir	0.857						
llitter_mass	0.877						

Fit measures: cfi rmsea srmr 0.998 0.070 0.001

D.4.3 Using the same set of sites as for the global model

These models were intended to confirm that the difference in the directionality was not related to different sites included in each dataset. The models confirm that using the 31 sites included in the dataset #1 directionality patterns do not change: at the local scale, biomass is more likely to explain diversity than the opposite (Table S27).

Table S27 Effect of the direction in the relation between diversity and biomass descriptors on basic structural equation models of diversity, biomass and litter globally and locally using only sites used for the global analysis.

Diversity is described as either species richness (S) or Faith's PD (PD). Models were tested using path analysis with Fisher's C (with degrees of freedom, d.f. and p-value testing that the structural model fits the model). For all the models, the number of parameters K=27 and the number of observations n=876.

Madal	Path analyses							
WIOdel	AIC	AICc	Fisher's C	d.f.	p-value			
Acceptance criteria	Low	Low	Low		≥ 0.05			
Local scale								
S ← Biomass	68.2	69.9	14.2	6	0.028			
PD ← Biomass	76.7	78.5	22.7	6	0.001			
Biomass ← S	94.9	96.7	40.9	6	0.000			
Biomass ← PD	98.7	100.4	44.7	6	0.000			