Diversification patterns of *Solanum* L. (Solanaceae), plant macroecology and responses to land-use change

a thesis submitted for the degree of
Doctor of Philosophy in Life Science
by

Susy Echeverría-Londoño

Department of Life Science
Imperial College London
London SW7 2BZ, United Kingdom

2018

Declaration of Originality

I certify that this thesis, and the research to which it refers, are the product of my own work, unless mentioned here or in the text. I collected none of the ecological data. Data from Chapter 4 and 5 came from a number of researchers who have generously shared their data to the PREDICTS project. Chapter 5 has been published in the Journal "Diversity and Distributions" and has benefited from comments by co-authors and anonymous reviewers. Parts of the manuscript have been commented and proofread by my supervisors, Andy Purvis and Sandra Knapp, whereas others have been commented and proofread by the postdoctoral researchers Isabel Fenton and Adriana De Palma.

Susy Echeverría-Londoño

Copyright Declaration

Diversification patterns of *Solanum* L. (Solanaceae), plant macroecology and responses to land-use change

Abstract

Current patterns of biodiversity reflect, to a certain degree, the legacy effects of species adaptations to past environmental and geological settings. A more in-depth understanding of this history and the traits that shape it, therefore, will help us to improve our predictions of how species will respond to current environmental change. In this thesis, I apply different analytical approaches to a range of case studies at a variety of taxonomic and geographic scales to test the importance of this fundamental hypothesis.

To examine the consequences of evolutionary and biogeographic history on the evolution of global biodiversity, in Chapter 2, I focus on the hyperdiverse plant genus *Solanum* (ca. 1300 species). This genus is an ideal case study since it combines a complete high-level phylogeny with global species distribution data, and covers an ecological spectrum from endemics to weeds. Chapter 2 shows that the vast diversity of *Solanum* is the result of at least two radiations. The majority of the lineages distributed in the Old World represent the most significant recent radiation, diversifying nearly twice as rapidly as any other group of solanums. This chapter also provides a brief comparison of the current approaches for modelling multi-rate diversification. In Chapter 3, I explore how the evolutionary legacy of colonisation, dispersal and climatic history have affected patterns of diversity in the genus. In this chapter, I show how successful colonisation of new areas and environmental changes can — but does not always — drive explosive diversification.

In Chapter 3, I show how arid-adapted lineages within *Solanum* have benefited from widespread habitat drying over the last few million years. This successful expansion reveals the potential evolutionary capacity of this group to expand and colonise currently disturbed and open areas, which is supported by the spread of some species considered as weeds such as *S. elaeagnifolium, S. torvum, S. nigrum.* In Chapter 4, therefore, I undertake a global analysis to assess whether the climatic preferences that have shaped the macroevolution of *Solanum* are now also shaping plant macroecology worldwide. For this analysis, I broaden my taxonomic scope to consider all plants, analysing an extensive global database that I helped to compile on how terrestrial assemblages respond to land use change, using a simple and very coarse classification of land use. In this

chapter, I demonstrate that species adapted to mesic conditions are highly sensitive to habitat conversion, compared with widespread arid and warmer-adapted species. These results show how land-use and climate change may favour similar species, thus potentially increasing the rate of homogenisation caused in the Anthropocene.

As shown in Chapter 4 species' responses to current environmental changes vary widely, depending on their ecological traits and climatic adaptations. This heterogeneity of responses could drive significant rearrangements in the composition of ecosystems, especially in the tropics where most of the species with narrow geographic and climatic ranges are found. In chapter 5, therefore, I quantify the impacts of land-use change on the composition of tropical assemblages using Colombia as a case study. In this chapter, I statistically analyse plant and animal data from 285 sites in Colombia to model how terrestrial assemblages are responding to land use change, using a much more finely-resolved land-use classification than that used in Chapter 4. I combine these models with four projections of land use to investigate how diversity is expected to change under future scenarios associated with different climate change policies. Here I demonstrate that land-use change has driven an increasing change in the composition of the ecological assemblages in Colombia and that depending on the policies implemented in the future, this negative effect could continue, risking the quality of ecosystems unless the impacts are mitigated.

To my husband, my aunt Vicky and my uncle Hermes.

ACKNOWLEDGMENTS

I would like to thank my PhD advisor, Professor Andy Purvis who gave me the opportunity to join his team in Silwood Park as a master student and at the Natural History Museum as a PhD student. Thanks Andy for supporting me during the past five years and for helping me during this very challenging period of my life. I am also very grateful with my co-supervisor Dr Sandra Knapp for all her advice and trust during my PhD. Thanks Sandy for your enthusiasm and for sharing with me your knowledge and adventures. I also have to thank Tiina Särkinen and Natalie Cooper for their scientific advice and insightful discussions and suggestions. Natalie thanks a lot for all your help not only during my PhD but also in finding a Postdoc.

I will forever be thankful to Isabel for all her advice, discussions, suggestions and proofreading. Isabel your support was vital for me. I very much appreciate all the academic and personal conversations, which truly made a difference during my PhD. I also want to thank Adriana for all her statistical advice, proofreading and for all the hugs that she gave me when I most needed them. Thanks Adriana for all your good vibe and for buying me coffee and cake for my engagement and birthdays!

I also want to thank my bachelor supervisor Daniel Rafael Miranda-Esquivel for all the training and advice. Everything I learned back in Colombia made a significant contribution to my performance as a Masters and PhD student.

I thank all the members of the Purvis lab: Katia, Rogier, Sara, Victoria, Dan, Sam, Helen, Adriana, Isabel and Connor for all the fun and for listening to my complaints every day (also for trying to understand what I was saying). We share a lot of fun moments but also some upsetting ones, 2016 and 2017, not the greatest years for the human race and the world in general. I also want to thank Dom who became an exceptional friend since the masters. Thanks Dom for all the movies, the coffee and our career crisis conversations.

My latin girls! Katia, Belen, Blanca and Nadia. You made my time in the museum unforgettable. Thanks for all your support, the fun, the tears and all the coffee! I will miss you all! Belen, I will miss you so much. You became one of my best friends, and I am grateful for all what you did for me during my time in London. I will never forget all the good and also the sad moments we share together.

I would like to thank my aunt Vicky and uncle Hermes for all their support with my university studies. There is no doubt that without them I would never have achieved anything. There are no words to express how grateful I am for your trust, love and support.

The best outcome for the last five years is finding my soul-mate and now husband, Francesco. Thank you for all your love, patience and support. Thanks for believing in me more than anyone else and helping me when sometimes I was sad and depressed. Thanks for choosing me as your life partner and for being as you are.

I gratefully acknowledge the funding sources that made my PhD possible. I was partially funded by the Administrative Department of Science, Technology and Innovation of Colombia (COLCIENCIAS) for my masters and two years of my PhD. I also want to thank my husband and father-in-law for their financial support during the last two years of my PhD.

List of figures

List of tables

CONTENTS

CHAPTER 1

General introduction

Ongoing anthropogenic pressures are now challenging biodiversity patterns modelled over millions of years of evolutionary history. The future of biodiversity will depend on how species are able to respond to these environmental challenges. To understand these species' responses, we first need to address how historical factors have left their imprint in the form, function and distribution of the modern biota. Most importantly, understanding the context in which current patterns of biodiversity evolved provides an insight into how species will respond to future environmental changes. In this review, I address how macroevolutionary tools can help us to identify the processes that have modelled the diversification of the current biota. I also examine how these historical contingencies are particularly crucial in understanding the species response to disturbance and local change. Here, I focus primarily on the effect of land-use change, the major current pressure to biodiversity, and the consequences of this ongoing disturbance on plant species and local assemblages in vulnerable regions such as the neotropics.

1.1 Dynamics of diversification

Although there are still millions of species yet to be discovered and new places to explore, the information gathered during several decades from different researchers across the world has shown that biodiversity is not even. There are some groups and regions with a disproportional number of species, for example, angiosperms have 260,000 species while their sister clade is *Amborella trichopoda* (Soltis & Soltis, 2004; Moore *et al.*, 2007); monocots (*ca.* 60,000 species) compared with Acorales (2-4 species, Hertweck *et al.* 2015); or therian mammals (*ca.* 5,000 species) compare with monotremes (6 species, Luo 2007). Understanding the patterns and processes of this uneven distribution of diversity is a fundamental question in ecology and evolutionary biology, and attempting to answer this question could reveal important mechanisms about the origin and maintenance of biodiversity.

A “hollow-curve” of biodiversity is created by the preponderance of species-poor groups and relatively few species-rich groups (Willis & Yule, 1922). Many potential explanations have been proposed for this pattern, including taxonomic artefacts (Scotland & Sanderson, 2004) and the stochastic nature of a birth-death process (Raup *et al.*, 1973; Purvis, 1996a). One explanation, in particular, has been prominent in the macroevolutionary literature: that differences in rates of speciation and extinction are the cause of unequal diversity patterns (Strathmann & Slatkin, 1983; Purvis, 1996a; Barraclough & Nee, 2001; Wiens, 2011). This hypothesis is usually tested by departures from null models of lineage equivalence, which lead to both clades with unexpectedly high diversity (e.g., radiations) and clades with unexpectedly low diversity (e.g., depauperons) (Purvis, 1996a; Donoghue & Sanderson, 2015).

Heterogeneity in diversification rate can be driven by a number of factors, including the influence of biotic factors such as species interactions or key innovations; and the impact of abiotic factors such as environmental and geographical changes (Vrba, 1992; Benton, 2009; Badgley & Finarelli, 2013). Evolutionary innovations can promote changes in diversification rates through ecological opportunities such as increases in fitness (e.g., reproductive success, ecological specialisation) or allowing the invasion to new adaptive zones (Heard & Hauser, 1995; Yoder *et al.*, 2010). Classic examples of key innovations promoting diversification include nectar spurs of columbines (Hodges & Arnold, 1995), hypocone in mammals (Hunter & Jernvall, 1995), and the perennial habit in *Lupinus* (Drummond *et al.*, 2012), among others. Differences in diversification among taxa can also occur as movement into new areas presents new ecological opportunities. Colonisations of new habitats can trigger an increase in the diversification of a group through the ecological release by escaping antagonists (e.g., parasites or competitors), the colonisation of depopulated niche spaces and/or the availability of abundant and underutilised resources (Moore & Donoghue, 2007; Losos, 2010; Schluter, 2000). Examples of diversification changes promoted by the colonization to new areas include the dispersion to new mountainous regions (e.g., Dipsicales, Moore & Donoghue 2007; *Lupinus*, Hughes & Eastwood 2006) or the dispersion to new emerged islands (e.g., Hawaiian spiders, Gillespie 2004; Hawaiian silversword alliance, Baldwin 1997; and Californian shrubby tarweeds, Baldwin 2007).

1.1.1 Diversification approaches

The first step to uncover the influence of ecological and geographical factors to current patterns of diversity is to quantify the differences in diversification rates among groups and regions. Phylogenetic Comparative Methods (PCM) provide a series of approaches to study diversification using reconstructed phylogenies along with various stochastic models (see Stadler 2013 and Morlon 2014 for a detailed review of diversification approaches). Since the first PCM approach in diversification was first proposed (Nee *et al.*, 1992, 1994b; Harvey *et al.*, 1994) a number of models and methods have been developed based on the premise that phylogenies contain information about past diversification events. Following Moore & May (2014), the study of diversification can be divided into four fundamental questions (1) What are the speciation and extinction rates of particular taxa? (2) Are diversification rates constant through time? (3) Are the diversification dynamics heterogeneous along the branches of a phylogeny? (4) Is there a correlation between changes in lineage diversification and a character state? Here, I focus on the third question since identifying the location and magnitude of diversification rate variation is considered an essential step to potentially uncover the mechanisms that have shaped current patterns of diversity (see Appendix A for an introduction of the approaches usually implemented to answer the other diversification questions).

Are the diversification dynamics heterogeneous along the branches of a phylogeny? Not accounting for rate heterogeneity can produce biases in the estimation of net diversification rates due to the influence of various evolutionary dynamics (Morlon *et al.*, 2011; Rabosky, 2010; Chira & Thomas, 2016). Detecting significant variation in speciation and extinction rates among clades in phylogenies can help us to understand patterns of accelerated or decelerated diversification, and a posteriori to identify the mechanisms behind these (*e.g.,* adaptive radiation or key innovations). The most commonly used methods to detect significant rates shifts along branches of a phylogeny are MEDUSA (Modelling Evolutionary Diversification Using Stepwise Akaike Information Criterion, Alfaro *et al.* 2009); and BAMM (Bayesian Analysis of Macroevolutionary mixtures, Rabosky 2014; Rabosky *et al.* 2014b). Other approaches include Morlon *et al.* (2011) and Etienne & Haegeman (2012).

MEDUSA is usually used to detect shifts in diversification rates among predefined clades whose diversity is known. This approach assumes constant rates of speciation and extinction within each clade and uses extant species richness information to account for incomplete taxon sampling. MEDUSA applies either a constant-rate pure-birth model or a constant rate birth-death model to an ultrametric phylogeny and then fits a range of different values of speciation and extinction to the branches. It then adds a diversification rate to a random subclade — creating a two-rate model which includes a separate set of speciation-extinction rate parameters — and recalculates the likelihood. If the one-rate model is significantly worse than the two-rate model (i.e., the difference in the AIC score exceeds a particular threshold), then the two-rate model is accepted. This procedure is repeated with additional diversification-rates along different subclades until the AIC score is no longer improved.

BAMM detects heterogeneity in evolutionary rates assuming that rates can vary through time. This method divides the branches in the phylogeny into several intervals or segments where the rates are assumed to be constant but may change from one branch segment to another (Rabosky, 2014). The underlying branching process in BAMM is more complex as the speciation and extinction parameters in the model include the effect of time-dependence in diversification (i.e., the age of a lineage can affect its diversification rate). This method uses reversible-jump Markov Chain Monte Carlo method to explore a larger space of parameters and candidate models of diversification than in likelihood approaches. BAMM identifies several sets of rate shifts that are sampled together and computes their relative probabilities, similar to the inference of maximum clade credibility tree in phylogenetic reconstructions (Rabosky, 2015a). Since BAMM follows a Bayesian statistical framework (rather than the maximum likelihood framework like MEDUSA), this method accounts for the uncertainty in parameter estimates by providing a distribution of marginal posterior probabilities instead of point estimates.

While recent papers have called into question the performance and reliability of patterns inferred by these two widely used approaches (see May & Moore 2016; Moore *et al.* 2016), others have not found any methodological issues besides the idiosyncrasies that are evident in any diversification approach (Rabosky *et al.*, 2017; Mitchell & Rabosky, 2016). Diversification modelling is a field with rapidly growing data, and there is a lot of scope for improving existing methods and also the developing of new ones. While the debate about the reliability of these approaches is still ongoing, their use — with the careful consideration of their assumptions, caveats and biases — can help us reveal fundamental insights into the processes that generate and maintain biodiversity.

1.1.2 Incomplete phylogenetic sampling

Despite the increasing number of molecular phylogenetic studies, we are still far from having complete phylogenies even for major groups of organisms (Reddy, 2014). Limited taxon sampling is one of the primary sources of uncertainty in comparative analysis and its effects are exacerbated by non-random or uniform sampling among clades and regions (Höhna *et al.*, 2011). Achieving DNA sequence-based phylogenies for species-diverse groups with wide distributional extents will be a great achievement to understand the dynamics that have shaped biodiversity. However, obtaining the DNA of species from groups with great species richness is non-trivial, sometimes requiring expensive and logistically demanding biological expeditions to largely inaccessible regions. Adding taxa to incomplete phylogenies could buffer the misleading impacts of limited taxon sampling in diversification approaches such as the inaccurate estimation of long branches (Wiens & Tiu, 2012). By using birth-death polytomy resolvers such as those developed by Kuhn *et al.* (2011) and Thomas *et al.* (2013), it is now possible to include these missing taxa into phylogenetic trees using information from detailed taxonomic studies. These methods enable researchers to generate distributions of complete species-level trees despite the lack of genetic data of many species. By placing missing species into the phylogeny using taxonomic constraints and a stochastic birth-death model, these methods represent the most conservative option with respect to divergence times and diversification rate estimation (Kuhn *et al.*, 2011), at least in the absence of genetic samples. Even though Rabosky (2015b) demonstrated that polytomy resolvers are not appropriate for studies that involve trait data, he also provided evidence that any inference in diversification dynamics is expected to be conservative under birth-death polytomy resolvers methods, providing a significant advance for diversification studies.

1.1.3 Diversification of megadiverse taxa, *Solanum* L (Solanaceae) as a case study

Angiosperms, or flowering plants, constitute one of the most significant radiations of terrestrial organisms (Lidgard & Crane, 1988; Niklas *et al.*, 1983). In a brief period after the Cretaceous, angiosperms experienced a dramatic increase in diversity and the resulting heterogeneity in morphological features and adaptive strategies allowed them to become ecologically widespread (Lidgard & Crane, 1988; Niklas *et al.*, 1983; Silvestro *et al.*, 2015; Magallón *et al.*, 2015; Crepet & Niklas, 2009; Friis *et al.*, 2011). The vast diversity of angiosperms has been shaped by a heterogeneity of evolutionary dynamics and nested radiations (Magallón & Sanderson, 2001; Beaulieu *et al.*, 2015), making these organisms very useful for understanding the evolution of biodiversity.

The diversity in angiospersms is not distributed evenly, with some groups supporting dramatically larger numbers of species than others (Takhtajan, 2009; Ricklefs & Renner, 1994). Magallón & Sanderson 2001 and Tank *et al.* 2015 showed that this disproportional diversity reflects differences in evolutionary processes among taxa and regions (i.e., differences in speciation and extinctions rates). Understanding the macroevolutionary processes that have shaped the diversity of clades with great diversity provides an excellent opportunity for the better understanding of origin and maintenance of spatial and temporal diversity. Due to the size and distribution of groups with large species richness, it has been very difficult to study these megadiverse clades in their entirety. However, thanks to the rapid and collective taxonomic and systematic studies in the last decade, we now have a better understanding of the lineage resolution and distribution of megadiverse groups.

One of the most diverse angiosperm groups is the genus *Solanum* L (Solanaceae). *Solanum* is considered as one of the top ten species rich genera in angiosperms (Frodin, 2004), and is the biggest genus of the family Solanaceae with *ca.* 49% of its species (more than 1200 species, Särkinen *et al.* 2013). In the last decade, there has been a significant increase in the number of species-level taxonomic and systematic studies in *Solanum* in both the Old and the New World (Tepe & Bohs, 2011; Stern *et al.*, 2013; Knapp, 2013; Clark *et al.*, 2015; Wahlert *et al.*, 2015; Aubriot *et al.*, 2016; Knapp & Vorontsova, 2016; Särkinen *et al.*, 2015), thanks to the collaborative efforts of researchers around the world. An electronic monograph of the entire genus is now available in the web resource Solanaceae Source (http://www.solanaceaesource.org).

Despite the great morphological variation in habit, all species of *Solanum* share a similar underlying flower structure, with a sympetalous corolla and anthers opening by terminal pores (Bohs, 2005; Weese & Bohs, 2007) (Figure 1.1). One of the principal divisions of *Solanum* based on morphology is that between the spiny (prickly species) and non-spiny solanums. This classification has been long standing since Dunal (1813, 1816) and the monophyly of the spiny solanums known as the "subgenus Leptostemonum" has been well-supported by the majority, if not all, the phylogenetic studies of *Solanum*. Non-spiny solanums do not form a monophyletic group according to the most recent phylogenetic studies (Bohs, 2005; Weese & Bohs, 2007; Särkinen *et al.*, 2015), however, this group can be divided into several monophyletic clades, of which the Geminata, the Potato clade and the Brevantherum clade are the most diverse. The majority of non-spiny solanums (with the exception of Normania, Archaeosolanum and the African Non-Spiny subclades) are distributed in the New World. In contrast, besides the great diversity of spiny solanums in the New World, this group also supports significant diversity centres in the Old World (e.g., Australia ca. 130 species, and Africa ca. 79 species).

Because of the great diversity of *Solanum*, phylogenetic studies of this group have been concentrated mainly in its subgroups/sections and few studies have examined the phylogeny of the entire genus (see Table 1.1). Thanks to these systematics studies, *Solanum* has experienced important taxonomic and systematic rearrangements, including the addition of genera such as *Cyphomandra*, *Lycopersicon*, *Normania* and *Triguera* (Spooner

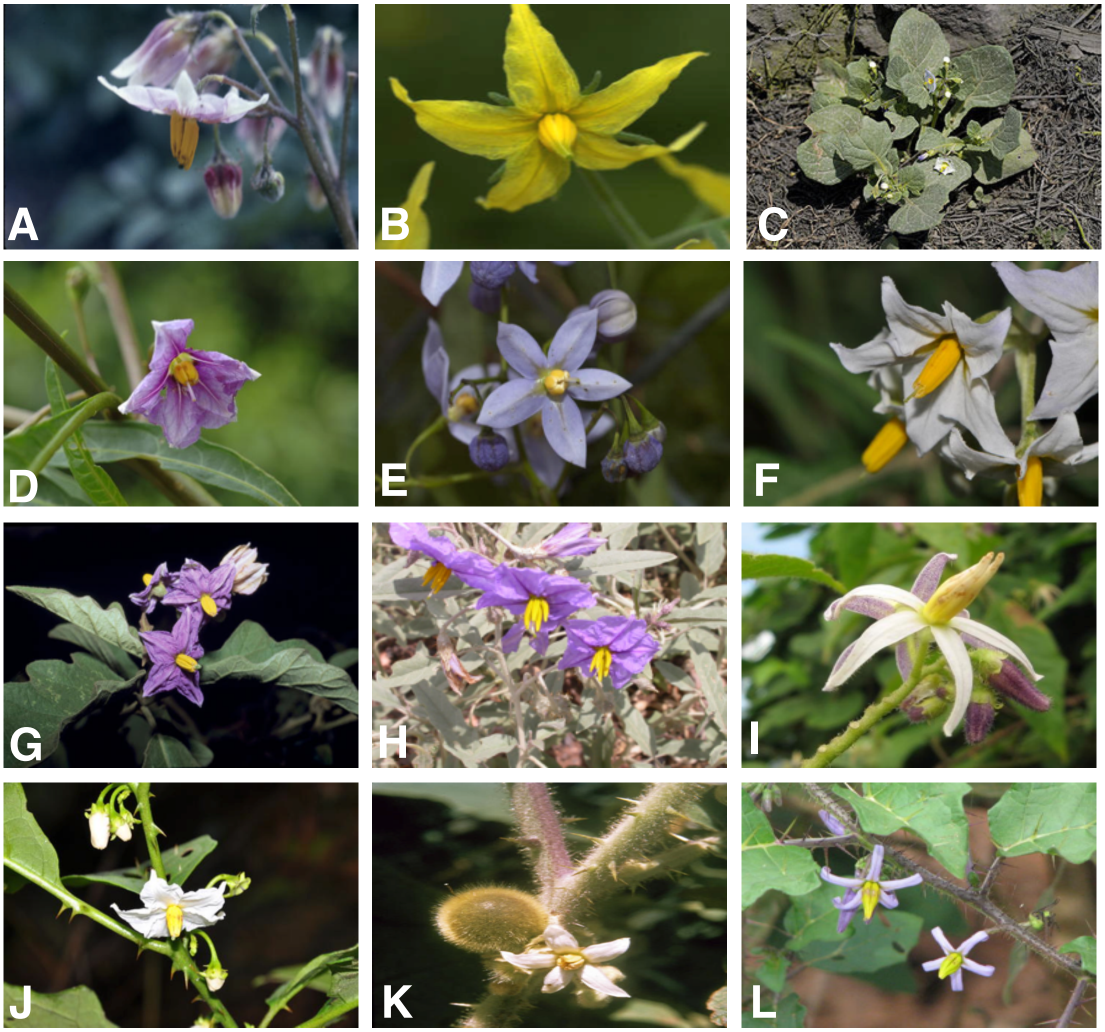


Figure 1.1: Example of the floral morphology in *Solanum* species from different subclades or sections. Figures A-F are examples of non-spiny solanums and Figures G-L of spiny solanums. (A) Petota, *S. tuberosum* L. (©Sandy Knapp) (B) Tomato, *S. lycopersicum* L. (©Sandy Knapp) (C) Regmandra, *S. montanum* L. (©Jonathan Bennett) (D) Archaesolanum, *S. aviculare* G.Forst. (©Lynn Bohs) (E) Dulcamaroids, *S. seaforthianum* Andrews (©Sandy Knapp) (E) Morelloids, *S. salicifolium* Phil. (©Sandy Knapp) (F) Old World clade, *S. campylacanthum* Hochst. ex A.Rich. (©Sandy Knapp) (G) Elaeagnifolium, *S. elaeagnifolium* Cav. (©Sandy Knapp) (H) Micracantha, *S. volubile* Sw. (©Stephen Stern) (J) Asterophorium, *S. asterophorum* Mart. (©Sandy Knapp) (K) Lasiocarpa, *S. stramoniifolium* Jacq. (©Dennis Adams) (L) Acanthophora, *S. palinacanthum* Dunal (©John R.I. Wood). All images were taken from `http://solanaceaesource.org` .

et al., 1993; Bohs & Olmstead, 2001). The most recent and more exhaustive phylogenetic study of *Solanum* was made by Särkinen *et al.* (2013). This study represents a time-calibrated species-level phylogeny for the family Solanaceae estimated using two nuclear and six plastid loci from 1075 species. This study included 450 (*ca.* 34%) species of *Solanum* so has the highest taxon sampling and resolution published so far. Based on this phylogenetic reconstruction, *Solanum* began to diversify roughly 13-18 million years ago (Mya) most likely in the Neotropics (Dupin *et al.*, 2016; Olmstead, 2013). Since then a great expansion appears to have occurred throughout the New World, and several dispersal events help this genus colonise the Old World (Dupin *et al.*, 2016; Olmstead, 2013).

In the first two empirical chapters of this thesis, I aim to understand the evolutionary processes that have shaped the diversity of the megadiverse genus *Solanum*. I specifically explore whether clade-specific differences in diversification rates have driven the apparent imbalance in species richness in *Solanum* (Chapter 2); and investigate how historical biogeographic events and climatic affiliations have influenced changes in the diversification of this megadiverse genus (Chapter 3). With these chapters, I attempt to contribute to the understanding of the evolutionary dynamics that has shaped the diversity of groups with high species richness using an integrative approach.

1.2 Biotic responses to land-use change

Global biodiversity patterns shaped over millions of years of evolutionary history are currently facing significant rearrangements by human impacts (Pereira *et al.*, 2010; Sala *et al.*, 2000; Newbold *et al.*, 2015, 2016a). Among these human activities, land-use change is considered the major pressure for terrestrial biodiversity (Foley *et al.*, 2005; Sala *et al.*, 2000). Almost one-third of the forests in the planet has been permanently cleared, and the majority of remaining forests are fragmented (Haddad *et al.*, 2015; Hurtt *et al.*, 2011). Agricultural activities are the major causes of forest clearing, representing almost the 30% of the Earth's surface (Hurtt *et al.*, 2011; Ellis *et al.*, 2010). The expansion of these human-dominated habitats is responsible for significant declines in species diversity and abundance at global and local scales (Millennium Ecosystem Assessment, 2005; Newbold *et al.*, 2015; Pimm *et al.*, 2014; Tittensor *et al.*, 2014). As global human population and gross domestic product are likely to continue rising, the expansion of agricultural lands for food production is also projected to increase (Harfoot *et al.*, 2014; Foley *et al.*, 2005; Van Vuuren *et al.*, 2011). Thus, further declines in biodiversity are expected to continue (Newbold *et al.*, 2015; Tittensor *et al.*, 2014).

A recent global model of the biotic consequences of land-use change (Newbold *et al.*, 2015) suggested that land-use pressures have already reduced numbers of species found in local assemblages by an average of almost 14% worldwide, based on an analysis of the PREDICTS database (Hudson *et al.*, 2014). These land-use pressures influence biodiversity directly by habitat loss and fragmentation or indirectly through losses of resources such as nesting, foraging habitats and biotic interactions (Owens & Bennett, 2000). However, not all species are affected by anthropologic changes in the same way. Land-use change can also open up new habitats for other species, increasing their abundance and expanding their ranges (McKinney & Lockwood, 1999; Wardle *et al.*, 2011).

Species' responses to land-use change vary widely, and recent studies reveal that these responses might depend on species' ecological traits. For instance, traits such as body

Source	Group or section	Number of ***Solanum*** species	Genes
Olmstead *et al.* (2008)	Solanaceae	11	ndhF, trnLF.
Särkinen *et al.* (2013)	Solanaceae	451	ITS, matK, ndhF, psbA-trnH, trnS-G, trnL-F, waxy
Bohs & Olmstead (1997)	*Solanum*	18	ndhF
Olmstead & Palmer (1997)	*Solanum*	36	cpDNA restriction site
Bohs (2005)	*Solanum*	120	ndhF
Weese & Bohs (2007)	*Solanum*	102	ndhF, waxy, trnS-trnG
Palmer & Zamir (1982)	Lycopersicon	15	cpDNA restriction site
Spooner *et al.* (1993)	Potato, Cyphomandra, Lycopersicon	24	cpDNA restriction site
Bruneau *et al.* (1995)	Lasiocarpa	11	cpDNA restriction site, morphology, isozyme
Castillo & Spooner (1997)	Petota	181	cpDNA restriction site, morphology
Bohs & Olmstead (2001)	Normania and Triguera	40	ITS, ndhF
Peralta & Spooner (2001)	Lycopersicon	69 accesions of 10 species	GBSSI
Spooner *et al.* (2005)	Lycopersicon	65 accessions of 10 species	AFPL, cpDNA, GBSSI, ITS, morphology
Levin *et al.* (2005)	Acanthophora	14	ITS, waxy, trnT-trnF, trnS-trnG
Martine *et al.* (2006)	Melongena	250 accessions	ITS
Levin *et al.* (2006)	Leptostemonum	134	ITS, waxy, trnS-trnG
Bohs (2007)	Cyphomandra	60	ITS
Tepe & Bohs (2010)	Pteroidea	28 accessions of 16 species	ITS, GBSSI, trnT-trnF, 3 COSII makers
Stern *et al.* (2011)	Leptostemonum	160	ITS, waxy, trnT
Vorontsova *et al.* (2013)	African spiny Leptostemonum	128 accessions of 93 Solanum	ITS, waxy, trnT
Wahlert *et al.* (2014)	Carolinense	27 accessions of 10 species	ITS, waxy, trnT-trnF

Table 1.1: Phylogenetic studies in the genus and subclades of *Solanum*

size, generation time, level of specialisation, or migratory status have shown to influence the responses to land-use change in bird species (Newbold *et al.*, 2014). In bees and butterflies, for example, mobility features such as flight duration or foraging range, appear to influence most of their responses to the expansion of human-dominated habitats (De Palma *et al.*, 2015; Dennis *et al.*, 2017; Essens *et al.*, 2017). The effect of these responses usually differs among taxa and type of habitat (Frishkoff *et al.*, 2015).

Despite the importance of using ecological traits to predict species' vulnerability to habitat conversion, there is still a limitation in the number of species with available trait information, which constrain the research to specific taxa and regions. In this context, the use of niche-based traits could provide a beneficial approach to consider more generalised patterns of tolerance to habitat change. In fact, a niche-based approach could be even more informative about the sensitivity of species to anthropological changes since it describes the set of environmental conditions in which an organism develops and survives. Besides, in some studies such as Essens *et al.* (2017), climatic niche traits predicted more the variation in responses to habitat conversion than ecological traits.

Niche-based traits are usually implemented to predict the responses of species to climate change (Thuiller *et al.*, 2004; Buckley, 2007; Kearney & Porter, 2009). However, despite the significant impact of habitat conversion in biodiversity loss (Pereira *et al.*, 2010; Newbold *et al.*, 2015), environmental traits have rarely been used to assess biodiversity responses to land-use change (but see Frishkoff *et al.* 2016, 2015; Köster *et al.* 2013).

Land-use change affects significantly vegetation cover, and therefore, it can dramatically change the climatic conditions at a local scale. For example, several studies have demonstrated that habitat conversion in the tropics tends to increase local temperature (Senior *et al.*, 2017; Findell *et al.*, 2007; Lawrence & Vandecar, 2015; Pielke *et al.*, 2011), which can even surpass the temperatures projected by the maximum warming scenario (Senior *et al.*, 2017). As a consequence, habitat conversion tends to produce drier and warmer habitats compared with natural ones (Frishkoff *et al.*, 2015). This could result directly because human-dominated habitats usually lack canopy cover which buffers microsites climate changes (Scheffers *et al.*, 2014). Species' responses to these microclimatic changes are expected to vary among taxa and regions. For example, species adapted to warm and dry areas such as shrubland could benefit from the creation of human-dominated habitat with similar climate regimes. In this context, Frishkoff *et al.* (2016) showed that, in a set of over 300 Neotropical forest birds, the species adapted to drier climates are better able to colonise agricultural land. These results appear to show that climate and land-use change benefit and affect the same species, producing even more dramatic declines in biodiversity. However, the study of species' responses to the microclimatic changes produced by habitat conversion has been restricted to few studies in particular taxa and regions, and a broader test would help to clarify whether the predicted patterns under these pressures apply more generally.

Identifying why some species are more vulnerable or tolerant to environmental changes would increase our predictive power to identify taxa at risk and forecasting changes in biodiversity. Given the importance of plant communities in the function and stability of ecosystems (Díaz *et al.*, 2013), Chapter 4 of this thesis attempts to identify which plant species are most vulnerable and tolerant to land-use change and assess whether there is a link between the species responses and their climatic adaptations.

The heterogeneity of species' responses to habitat conversion can lead to significant changes in community composition. As a consequence, natural habitats are evolving into new emerging environments usually dominated by generalists species, whose expansion will eventually lead to the homogenization of the biodiversity (McKinney & Lockwood, 1999; Gámez-Virués *et al.*, 2015; Rooney *et al.*, 2007; Hobbs & Mooney, 1998; McCune & Vellend, 2013). As the simplest and most commonly used metrics of biodiversity (e.g., local species richness) do not always capture the changes under anthropological pressures (Dornelas *et al.*, 2014; Magurran & Henderson, 2010; McCune & Vellend, 2013; Thomas, 2013), indicators of species composition may provide a more sensitive analysis and early diagnostic of these changes in disturbed habitats (Collins *et al.*, 2008; Mac Nally, 2007). The use of species composition in assessing the biodiversity responses to habitat conversion is even more significant in tropical regions, where the recovery of site-level richness can be more rapid (e.g., 20-40 years, Dunn 2004) than the restoration of species composition seen in undisturbed habitats (Martin *et al.*, 2013).

Understanding changes in species composition under habitat conversion are especially relevant in the tropics because these regions support the most biodiverse terrestrial habitats (Dirzo & Raven, 2003; Orme *et al.*, 2005) and are considered the most vulnerable areas by land-use change (Sala *et al.*, 2000; van Vuuren *et al.*, 2006; Jetz *et al.*, 2007; Alkemade *et al.*, 2009; Pereira *et al.*, 2010). Given the high heterogeneity of landscapes and differences in socio-economic practices among regions, Chapter 5 assess the impact of land-use change on the biodiversity of Colombia — one of the most biodiverse countries in the Neotropics with a detailed historical revision of land-use change. Using a space-for-time substitution approach, this chapter also estimates future changes of biodiversity under different models of land-use change from the climate change policies of the four representative concentration pathways (RCPs).

Overall, the primary aim of this thesis is to understand the processes that have modelled today's biodiversity using one of the most species-rich taxa on the globe — the plant genus *Solanum* (Chapter 2 and 3); and how these patterns are being affected by current environmental changes at multiple levels of biological organisation: species-level (Chapter 4) and community-level (Chapter 5).

CHAPTER 2

Dynamics of diversification in *Solanum* L (Solanaceae)

2.1 Abstract

Solanaceae represents one of the most diverse and economically important plant groups with around 3000 species distributed across the worldwide. Almost half of the diversity of this family belongs to the mega-diverse genus *Solanum* L, which has species that occupy a variety of niches and great heterogeneity of richness among its subclades. In this chapter, I aim to understand the role of diversification rate heterogeneity in shaping the diversity of this mega-diverse genus. Using taxonomy and phylogenetics coupled with a wide range of tools of macroevolution, I show that the diversity of *Solanum* has been shaped by a heterogeneity of rates among clades and regions, with some contrasting evolutionary dynamics. The highest net rates of diversification were found in a clade formed exclusively by spiny solanums distributed in the Old World, suggesting that geographic factors such as the colonisation to new areas, could have contributed to the lineage diversification of this group.

2.2 Introduction

The uneven distribution of taxonomic diversity among the branches of the tree of life and among geographic regions is one of the most intriguing puzzles in biology. The increasing availability of field studies and large phylogenies continues to reveal how some groups are highly diverse compared with their depauperate sister groups. We expect the stochastic nature of the diversification process to produce asymmetric patterns of diversity which do not necessarily require deterministic explanations (Raup *et al.*, 1973; Slowinski & Guyer, 1993; Purvis, 1996b). Nevertheless, identifying significant patterns of unequal diversity (i.e., those that depart from the null expectation) throughout the tree of life could allow us to uncover the mechanistic drivers that have shaped current patterns of biodiversity.

One of the main hypotheses in macroevolution is that unequal diversity can arise from significant differences in diversification rates, both among lineages and through time and space (Wiens, 2011; Stanley *et al.*, 1981; Alfaro *et al.*, 2009; Wiens, 2015). These differences are usually controlled by biotic interactions and/or by the external influences of environment and geography (Van Valen, 1973; Vrba, 1992; Benton, 2009; Purvis, 2008; Ezard *et al.*, 2011). More specifically, asymmetry in evolutionary rates is often interpreted as a result of a spectrum of ecological opportunities such as the opening of niche space by the development of new key traits, the extinction of antagonists and/or the colonisation to new open areas (Simpson, 1955; Schluter, 2000; Yoder *et al.*, 2010; Moore & Donoghue, 2007; Losos, 2010; Rabosky *et al.*, 2013). Quantifying this variation in the speciation and extinction rates along lineages of a phylogeny is the first step to uncovering the drivers that have shaped the diversity of taxa.

In this study, I investigate the macroevolutionary patterns and diversification dynamics of the extant species of the megadiverse plant genus *Solanum* (Solanaceae). With more than 1200 species, *Solanum* is the biggest genus of the family Solanaceae with *ca.* 49% of its species, (Särkinen *et al.*, 2013) and is considered one of top ten species rich genera of angiosperms (Frodin, 2004). This genus has a cosmopolitan distribution but it is most diverse in the southern hemisphere, especially in the Americas (Olmstead & Palmer, 1997).

As one of the most diverse genera in angiosperms, *Solanum* as a whole is expected to have high diversification rates. Moreover, the great imbalanced of diversity among its main groups, especially in groups with high species richness such as Petota (potatoes and their relatives), the Old World spiny and Geminata clades, suggests there is a great heterogeneity of rates among *Solanum*. The Neotropics, with almost 70% of species diversity of the genus, represents the centre of diversity and endemism for *Solanum*. This high diversity of lineages in the Neotropics could suggest higher diversification rates or longer accumulation of species compared with other regions in the world (Stadler, 2014)

Understanding the evolutionary processes that have promoted the diversification of radiations in older clades is likely to be challenging, due to the difficulty of accounting for extinct species that are not included in molecular phylogenies (Liow *et al.*, 2010; Quental & Marshall, 2010; Rabosky, 2010). Therefore, focussing on recently emerged clades with widespread distribution, that are species-rich and ecologically diverse, such as *Solanum*, will provide us with a more accurate picture of diversity dynamics when using well-sampled high-level molecular phylogenies. Another challenge in quantifying the differences in diversification rates among groups is the non-random distribution of species lacking genetic data along phylogenies (Thomas *et al.*, 2013; Höhna *et al.*, 2011). However, in groups with detail taxonomic information such as *Solanum*, complete species phylogenies can be inferred incorporating prior taxonomic information for the placement of missing data and account for non-random species absence.

Here, I present the first study that quantifies diversification rates in *Solanum* to test whether the variation in species richness is the result of diversification rate heterogeneity or whether it reflects other factors such as age. I assemble the first set of complete time-calibrated and species-level phylogenies of extant *Solanum* species (1169 species) compiled in a Bayesian framework. I then compare the dynamics of diversification across lineages, space and time, and identify regions with significant shifts in evolutionary rates using multiple comparative phylogenetics approaches.

2.3 Methods

Phylogenetic relationships and divergence times among the main clades of *Solanum* were obtained from the Särkinen *et al.* (2013) Solanaceae phylogeny. This tree was built using two nuclear and six plastid loci from 454 *Solanum* species (ca. 34% of the expected number). This phylogeny identifies a total of 37 clades and subclades within the genus, with a vast heterogeneity of species richness within these (see Figure 2.1).

2.3.1 Correcting for non-random taxon sampling

The incomplete nature of the Särkinen *et al.* (2013) phylogeny could be problematic. Non-random incomplete sampling in phylogenetic studies could significantly affect the estimate of macroevolutionary rates (Höhna *et al.*, 2011). Despite the claims that current speciation-extinction models account for incomplete species sampling (e.g., MEDUSA, BAMM, TreePar, TESS, RevBayes, etc.), these models often assume a random and even sampling of the species across the phylogeny (i.e., each terminal in the phylogeny is assumed to have the same probability of being sampled) (FitzJohn *et al.*, 2009). However, this assumption is very often violated in empirical examples due to geographic, temporal and taxonomic sampling biases. If we assume that species in the *Solanum* phylogeny have been sampled randomly and uniformly, we would expect roughly the same percentage of species to be sampled within each section of the genus. As Table 2.1 shows, the distribution of species sampling among clades in the phylogeny of *Solanum* varies widely with under-sampled clades as Geminata that have $\approx 7\%$ of species included in the phylogeny and over-sampled clades such as Tomato with $\approx 83\%$ of the species represented in the

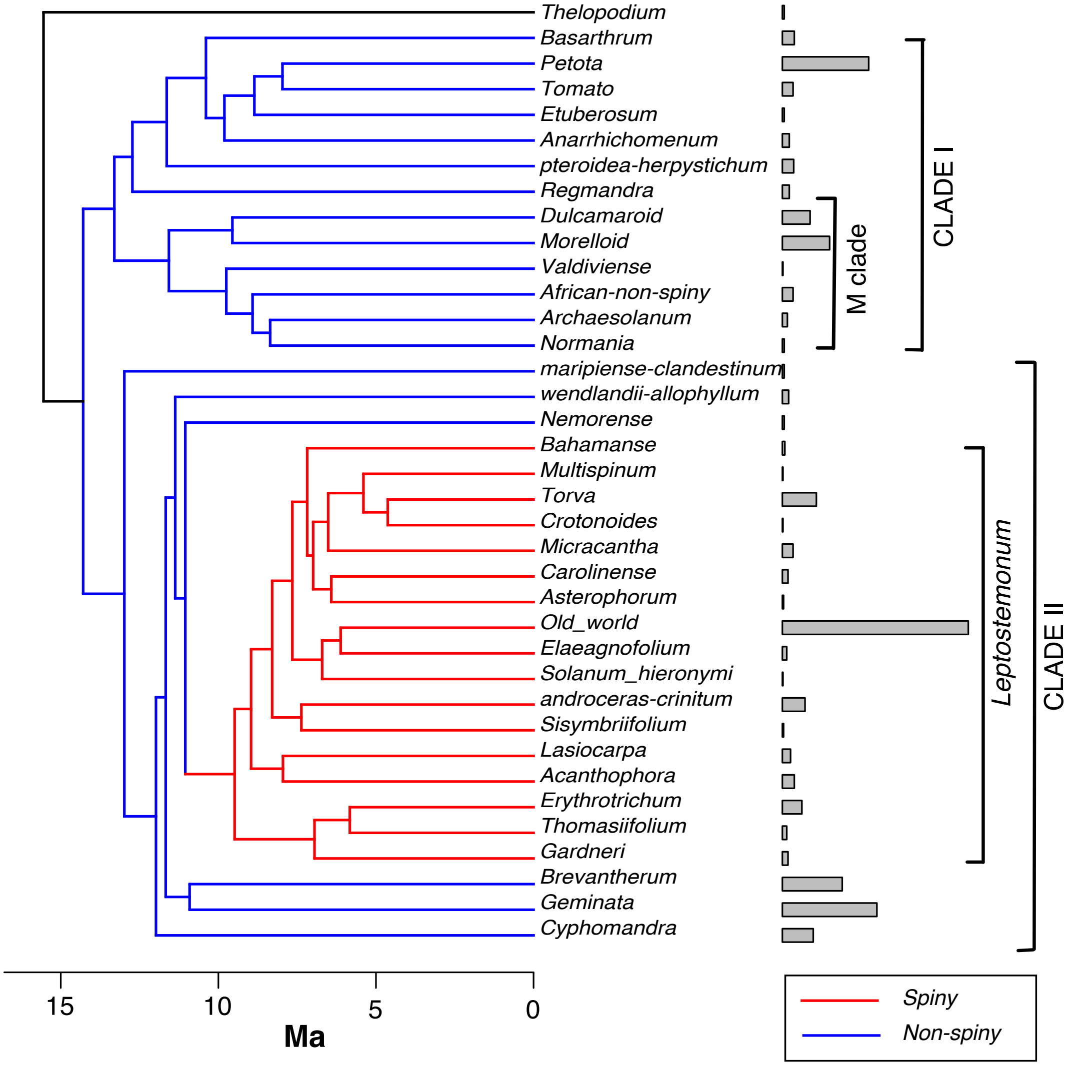


Figure 2.1: Backbone tree of *Solanum* from Särkinen *et al.* (2015). The tree shows the relationships among the main monophyletic sections of *Solanum*. The bars represent the relative species richness among sections. The largest bar corresponds to the richness in the Old World clade with 295 species. The time scale under the backbone is given in million years units.

phylogeny. To account for non-random incomplete taxon sampling in the diversification analysis, I used the polytomy resolver PASTIS (Thomas *et al.*, 2013), to place missing species in the phylogeny using taxonomic constraints following a birth-death model. This method is a conservative approach for increasing the sampling in phylogenies since it infers the timing of missing splits under a constant rates birth-death model. Therefore, any inference of rate-heterogeneity in the final analysis will have a strong support since this approach is expected to bias towards the detection of constant-rate models of diversification (Thomas *et al.*, 2013).

Sections	**Crown age (Mya)**	**PP**	**Expected number**	**Species in phylogeny**	**% in phylogeny**	**Undersampled**
Acanthophora	6.0 - 3.6	1	19	11	57.89	
African non spiny	8.6 - 3.1	0.86	17	2	11.76	*
Anarrhichomenum	4.0 - 0.8	1	11	2	18.18	*
Androceras-Crinitum	7.7 - 5.2	1	36	17	47.22	
Archaesolanum	6.1 - 2.7	1	8	8	100.00	
Asterophorum		1	2	1	50.00	
Bahamense	3.8 - 0.8	1	4	1	25.00	*
Basarthrum	6.4 - 2.2	1	19	4	21.05	*
Brevantherum	10.1 - 5.9	1	95	10	10.53	*
Carolinense	6.7 - 4.1	1	9	4	44.44	
Crotonoides		1	1	1	100.00	
Cyphomandra	7.8 - 4.5	1	49	31	63.27	
Dulcamaroid	9.5 - 5.7	1	44	11	25.00	*
Elaeagnifolium	5.6 - 2.3	1	7	3	42.86	
Erythrotrichum	5.2 - 2.9	0.8	31	11	35.48	
Etuberosum	9.3 - 4.4	1	3	2	66.67	
Gardneri	5.2 - 2.6	1	9	7	77.78	
Geminata	10.1 - 5.9	1	150	10	6.67	*
Lasiocarpa	5.1 - 2.5	1	13	12	92.31	
Mapiriense-Clandestinum	12.2 - 5.8	1	3	2	66.67	
Micracantha	5.5 - 2.6	1	17	7	41.18	
Morelloid	11.5 - 7.8	1	75	17	22.67	*
Multispinum		1	1	1	100.00	
Nemorense	10.3 - 2.4	1	3	3	100.00	
Normania	5.1 - 1.9	1	3	2	66.67	
Old World	6.4 - 4.4	1	296	119	40.20	
Petota	8.5 - 5.9	1	137	45	32.85	*
Pteroidea-Herpystichum	9.0 - 5	1	18	10	55.56	
Regmandra	6.6 - 1.9	1	11	3	27.27	*
Sisymbriifolium	4.6 - 1.3	1	2	2	100.00	
Solanum hieronymi		1	1	1	100.00	
Thelopodium	10.6 - 3	1	3	2	66.67	
Thomasiifolium	6.5 - 3.3	1	7	4	57.14	
Tomato	7.4 - 4.5	1	17	14	82.35	
Torva	4.1 - 2.6	0.84	54	22	40.74	
Valdiviense		1	1	1	100.00	
Wendlandii-Allophyllum	11.2 - 6.2	1	10	4	40.00	

Table 2.1: Sampling proportions for the main subclades of the *Solanum* phylogeny defined by Särkinen *et al.* (2013). The proportions were calculated dividing the number of species included in the phylogeny by the expected number of extant species within each section or clade. PP = Posterior probability. Undersampled groups are those that have less than 30% of the expected number of species.

I used the maximum clade credibility of the Särkinen *et al.* (2013) phylogeny as the "backbone" to define the general relationships among the main sections of the genus. Monophyletic constraints were used only for well-supported nodes (i.e., posterior probability $>$ 0.95%, see Table 2.1). Clades with poorly-supported nodes were left unconstrained (i.e., the Torva, African Non-spiny and Erythrotrichum clades). One of the nodes in the Särkinen *et al.* (2013) phylogeny within the Dulcamaroid clade had negative branch lengths; this can occur when a clade is poorly sampled across the posterior distribution (i.e., when there is a low support of a direct ancestor-descendant relationship). Since setting negative branch lengths to zero would create a non-ultrametric tree, I decided to drop all the tips from this weakly supported node from the backbone and include them in the PASTIS analysis as unconstrained. I then assigned each accepted name of *Solanum* to one of the clades or sections which are shown in Table 2.1, following the taxonomic treatment of *Solanum* species in Solanaceae Source and expert opinion (S. Knapp, pers. comm.).

For each clade or subclade, PASTIS creates an output file in a nexus format, which contains the full set of tree constraints ready to be executed in MrBayes. Therefore, posterior distributions of phylogenies for each clade were then inferred in MrBayes 3.2.3 (Ronquist & Huelsenbeck, 2003) using a relaxed clock model (independent branch rates - igr prior), with the default (exponential) prior on the distribution of branching rates. Two species from the closest sister clade were constrained as outgroups in each *Solanum* clade. Four chains, four independent runs and 400 million generations were run for clades with more than 150 species and 100 million generations for the rest of the clades. MrBayes was run using the Cipres gateway (Miller *et al.*, 2010, `http://www.phylo.org`). An example of the nexus block for a single clade run is:

```
prset brlenspr=clock:birthdeath;
prset Extinctionpr = Fixed(0);
prset Speciationpr=exponential(1);
prset clockvarpr=igr;
prset Igrvarpr=exponential(10);
```

I assessed the convergence, mixing and burn-in of all the parameters for the posterior distribution of each clade by visual examination using Tracer v 1.6.0 (Rambaut *et al.*, 2009). To create the complete species phylogeny of *Solanum*, a sample from the posterior distribution of the phylogeny or subtree of each group of *Solanum* was selected randomly and then grafted to the backbone following the methodology in Jetz *et al.* (2012). To do that, I first rescaled the depth of each subtree to 1. Then I computed the proportional depth of the crown group of each subtree, and grafted the subtree into the backbone at its original position keeping its ultrametricity (see Figure B.1 for a graphic explanation). This procedure was repeated 100 times for all subtrees (each representing a group of *Solanum*) until a distribution of 100 complete species phylogenies (i.e., containing nearly all described species, 1169 species in total) was produced.

2.3.2 Quantifying the diversification dynamics of *Solanum*

Several approaches have been developed to detect significant shifts in diversification across the branches of a phylogeny. In this analysis, I mainly used two approaches commonly used in the literature — MEDUSA (Alfaro *et al.*, 2009)(v1.41) and BAMM (Rabosky, 2014). However, other approaches such as TESS (Höhna *et al.*, 2016b) and RevBayes (Höhna *et al.*, 2016a) were also implemented to assess the robustness of the results obtained in BAMM.

Modelling Evolutionary Diversification Using Stepwise Akaike Information Criterion (MEDUSA)

Traditionally, the approach MEDUSA (Alfaro *et al.*, 2009) has been used to model changes in diversification rates at broad scales. This method integrates phylogenetic information with richness data to estimate differences in rates of speciation and extinction along branches. MEDUSA applies either a constant-rate pure-birth model or a constant rate birth-death model to an ultrametric phylogeny and then fits a range of different values of speciation and extinction to the branches. It then adds a diversification rate to a random subclade — creating a two rate model which includes a separate set of speciation-extinction rate parameters — and recalculates the likelihood. If the one-rate model is significantly worse than the two-rate model (i.e., the difference in the AIC score exceeds a specific threshold) then the two-rate model is accepted. This procedure is repeated with additional diversification-rates along different subclades until the AIC score is no longer improved. In this analysis, I implemented this approach using the function MEDUSA in the R package Geiger (v2.0.3) (Harmon *et al.*, 2008). As this was an exploratory analysis, this analysis was performed using on the maximum clade credibility (MCC) phylogeny of *Solanum* published by Särkinen *et al.* (2013). Both the constant-rate pure-birth model and the constant birth-death model were fitted to the phylogeny. The AIC threshold used was automatically calculated by the MEDUSA function based on the number of terminal lineages.

Bayesian Analysis of Macroevolutionary Mixtures (BAMM)

The BAMM approach (Rabosky, 2014) aims to provide an improvement on methods that identify heterogeneity in rates only across specific subclades, such as MEDUSA. The underlying branching process in BAMM is more complex as the speciation and extinction parameters in the model include the effect of time-dependence in diversification (i.e., the age of a lineage can affect its diversification rate). It also implements a diversity-dependent diversification model where the number of lineages in a clade may affect its diversification rates. This approach uses a reversible-jump Markov Chain Monte Carlo approach to explore a larger space of parameters and candidate models of diversification. Since it follows a Bayesian statistical framework (rather than the maximum likelihood framework like MEDUSA), BAMM implicitly accounts for the uncertainty in parameter estimates by providing a distribution of marginal posterior probabilities instead of point estimates. To implement this approach in the *Solanum* phylogeny, I used the distribution of 100 complete species trees produced in the PASTIS analysis to consider the phylogenetic uncertainty in the estimates of diversification rates. I set the priors of speciation (λ), extinction (λ) and the expected number of diversification shifts (γ=1) using the R package BAMMtools v 2.0.5 (Rabosky *et al.*, 2014b), which identifies the priors of the diversification parameters based on the distribution of divergence times of the phylogeny. For each of the 100 trees, an MCMC analysis was performed with four separate runs of 20 million generations. All the analyses were run using the C++ BAMM command line program v 2.5.0 on the Imperial College London's High-Performance computing cluster (`http://www.imperial.ac.uk/admin-services/ict/self-service/research-support/hpc/`). I then checked for convergence of the MCMC samples making sure the effect sample size was at least 200 for both the number of evolutionary shifts and likelihood using the CODA R package V 0.16-1 (Plummer *et al.*, 2006). The first 25% of the samples were discarded as burn-in.

For each of the 100 trees run in BAMM, a distribution of 1000 samples of the posterior probabilities of diversification were created. Each sample from the posterior includes either a single event (i.e., the diversification is described by a single time-varying process — no shifts in diversification) or a mixture of two or more shifts and associated parameters (see Figures B.2–B.5). For each set of posterior probabilities, I extracted the list of nodes associated with "core" rate shifts (i.e., rate shift with a marginal probability significantly higher than the probability expected from the prior alone) and calculated the frequency with which these nodes are associated with significant rate shifts across the 100 trees run in BAMM. Finally, I computed the mean of diversification rates for each of the species present in the Särkinen *et al.* (2013) phylogeny across the pooled distribution of posterior samples produced in the BAMM analysis.

A macroevolutionary cohort analysis sensu Rabosky *et al.* (2014a) and Shi & Rabosky (2015) was performed across all the 100 BAMM runs to visualise complex mixtures of dynamics of diversification. This approach calculates the pairwise probability that any two lineages share a common macroevolutionary regime. For a given sample from the posterior distributions, a pair of lineages is assigned a value of 1 when the species inherit a common rate regime and a value of 0 when the rate dynamics are completely decoupled.

The mean of these values is then calculated over all the posterior distribution samples in each of the 100 runs of BAMM. In this case, an average value of 0.1 between a pair of lineages implies that in 10% of the samples across the posterior distributions these two lineages support a common macroevolutionary rate.

BAMM sensitivity

A recent study by Moore *et al.* (2016) raised concerns about several methodological issues in BAMM including a potentially incorrect likelihood function, the strong influence of priors on posterior estimates, and some theoretical errors (e.g., the incorrect use of the Poisson distribution as the error distribution of the prior number of shifts of diversification). To assess the effects of these potential limitations on the results found in this study, I performed several sensitivity analyses focusing on the two most important issues stated by Moore *et al.* (2016): 1) the prior sensitivity of posterior distribution on the number of rate shifts, and 2) the reliability of the diversification rate estimates.

(1) Effects of the model prior: BAMM assumes that the expected number of shifts of diversification follows a Poisson prior distribution with an exponentially distributed hyperprior. This distribution can be simplified as a geometric distribution with mean γ (i.e., the expected number of shifts). According to Mitchell & Rabosky (2016) the prior distribution automatically set in BAMM is a conservative way to define the number-of-shifts prior since the zero-shift model is the most likely outcome. To test the sensitivity of the number of shifts found in BAMM to the prior distribution, I ran the analysis using five different prior expectations of $\gamma = 0.5, 1, 2, 10, 100$ using the original Särkinen *et al.* (2013) tree (due to computational limitations). The value of γ defines the shape of the geometric distribution with small values (e.g., 0.5) defining a strong prior with a skewed probability distribution towards a zero-shift model. In contrast, large values of γ (e.g., 100) represent liberal priors with a relatively flat probability distribution. Each γ treatment was run for 4 million generations and the first 20% of the samples were discarded as burn-in. I then plotted the marginal posterior distribution of the number of shifts obtained in each γ treatment and the prior distribution for each tree. A significant change in the marginal posterior distribution with different values of γ would demonstrate a strong sensitivity of the BAMM results to the prior distribution.

(2) Reliability of diversification rate estimates: Using an incorrect likelihood function could produce unreliable estimates of speciation and extinction rates and therefore, an unreliable number and location of shifts. To demonstrate the accuracy of the estimation and location of rates inferred from BAMM, it is necessary to perform an extensive series of simulations under different diversification models, something which is outside the scope of this analysis. However, to corroborate the results drawn from the BAMM analysis, I implemented two alternative approaches to estimate rates of speciation and extinction and identify shifts of diversification through time and across branches. With the first approach, I tested whether there is evidence of an episodic tree-wide increase in diversification rates through time using the R package "TESS" (Höhna *et al.*, 2016b). This package implements an approach which fits a birth-death model with episodically varying rates (i.e., identifies discrete tree-wide changes in speciation and extinction rates) assuming that the diversification among lineages at any point in time is constant (Höhna, 2015). As in other approaches that model diversification-rate through time such as TreePar (Stadler, 2011), TESS divides the tree into equal time intervals and tests whether there are significant changes in speciation and extinction rates among these intervals. For comparison purposes, I first performed this analysis using the Särkinen *et al.* (2013) phylogeny and then using the 100 complete species trees of *Solanum* obtained in PASTIS. For both datasets, I ran the function "tess.analysis" which uses a reversible-jump MCMC algorithm to estimate the number and magnitude of rate shifts. The number of events is assumed to follow a Poisson distribution and the estimates of speciation and extinction are assumed to follow a lognormal distribution with a fixed mean calculated empirically by an initial MCMC analysis under a constant birth-death model. All the analyses were run until the effective sample size reached 500 or the time reached a maximum of 24 hours.

The second approach I tested was the RevBayes program (Höhna *et al.*, 2016a), which estimates species-specific rates of speciation and extinction and locates branches with significant shifts in diversification. Unlike BAMM, RevBayes does not model the rates of diversification from a continuous distribution directly, but instead it divides the probability distribution into discrete rate categories. Then within each of the N quantiles of the distribution, in this case a lognormal, Revbayes integrates over all possible rates of speciation. The mean of this lognormal distribution is fixed to represent the mean of the expected diversification rate given each tree, which is equal to ln(ln(N taxa)/age)), with a fixed standard deviation of $(0.587405 * 2)$ to represent two orders of magnitude of variance in the rates. In this analysis, extinction rates were assumed to be equal in all the rate categories and the number of rate categories was set to 10. Without any prior information about the rate-shift events, I used a conservative prior distribution of the number of shifts defined by an exponential distribution with mean equal to one, giving the highest prior probability to a zero-shift model. I ran a pre-burn-in analysis using 10,000 generations to obtain starting values from the posterior distribution and improve the mixing of the MCMC analysis. I then ran the analysis using 100,000 generations or more until convergence was reached. Every 200 generations a tree was printed with the average parameters values of speciation and extinction at the branches and nodes. At the end of the analysis, all the trees were used to calculate the posterior distribution of the rates at the branches through a maximum a posteriori tree. Due to computational resource availability, I ran this analysis using only the first 20 trees from the pool of complete species phylogenies. The results from these two approaches were contrasted with those obtained using the BAMM analysis.

2.4 Results

In the MEDUSA analysis, four diversification shifts across the *Solanum* phylogeny were supported under a pure-birth model — three significant increases (within the Old World spiny, the Petota subclade and Torva clades) and one significant decrease (in the Thelopodium clade) (Figure 2.2). Among these shifts, the Old world clade showed the highest increase in diversification with 0.82 *lineages/Ma* compared with a background rate of diversification of 0.35 *lineages/Ma*.

In the BAMM analysis, in average across all the 100 complete species trees, the Old World clade was inferred to have radiated more than twice as rapidly — *ca.* 0.68 *lineages/Ma* — as the background rate 0.28 *lineages/Ma* (Figure 2.3). The node supporting the crown group of the Old World clade was the node with the highest number of shifts found across the 100 BAMM runs. Other groups such as the Petota subclade (from the Potato clade, Petota + Tomato) and the node supporting the Leptostemonum group (i.e., the spiny solanums) also show some signal of shifts in diversification but with a weaker support compared with the Old World clade. No shifts in diversification were found in either the Torva clade or Thelopodium clade, however, in contrast to the results from the MEDUSA analysis.

The macroevolutionary cohort analysis supports the decoupling of all the lineages of the Old World clade from all the rest of *Solanum* (Figure 2.4). Interestingly, this analysis also reveals, but with weaker support, a vast much heterogeneity in macroevolutionary regimes in other groups that did not show significant shifts in diversification rates in the previous analysis. The Petota subclade is highlighted as a group with around 50% probability of being decoupled from other solanums.

The number of diversification shifts found in the BAMM analysis is largely robust to the defined prior distribution. Figure B.2–B.5 show that across all 100 BAMM runs, the distribution of estimated shifts (posterior samples) differs from the distribution of the expected number of shifts (prior). Although the prior distribution (γ=1) applied in this analysis was strong and conservative (i.e., the zero-shift model was set to be the most

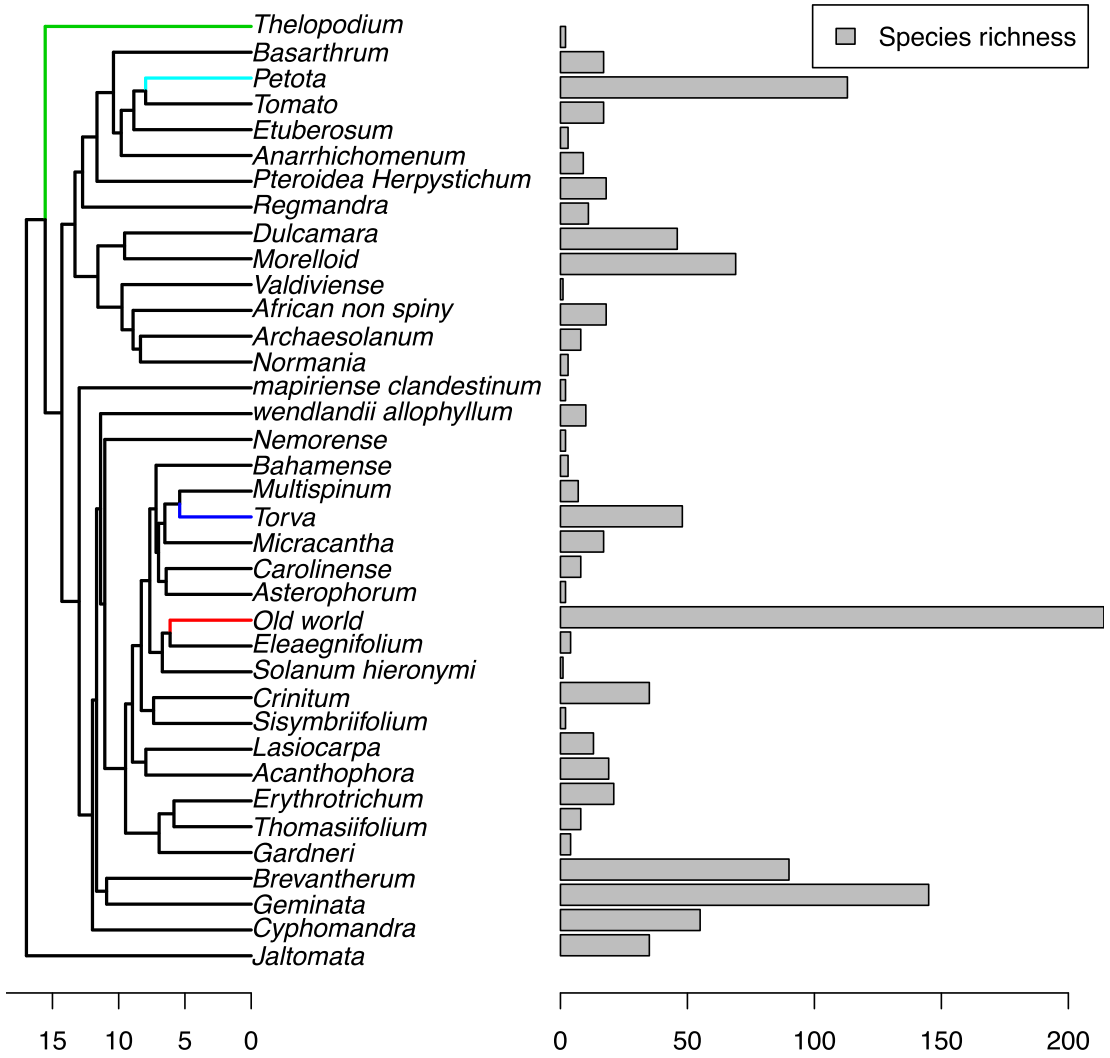


Figure 2.2: Diversification shifts identified in the phylogeny of *Solanum* using MEDUSA. The colours represent the different shifts in net diversification under a pure-birth model. The axes represent the divergence time in million years (left) and and number of species (right).

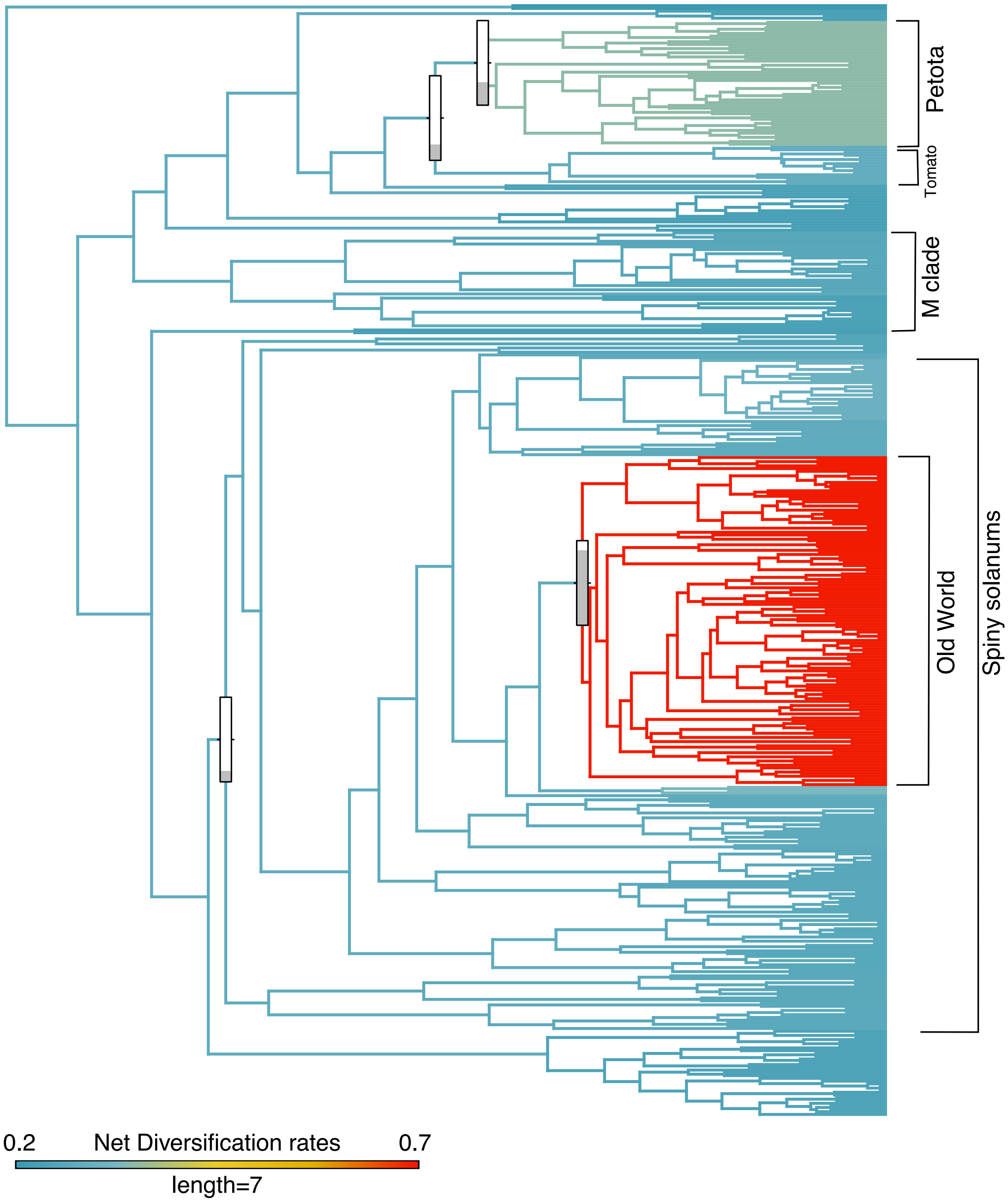


Figure 2.3: Overall mean net diversification rates of *Solanum* lineages obtained in BAMM and mapped onto the Särkinen *et al.* (2013) phylogeny Thermometers represent the frequency in which the nodes were associated with shift configurations across the 100 trees obtained by the polytomy resolver PASTIS (nodes with less than 10% frequency were omitted)

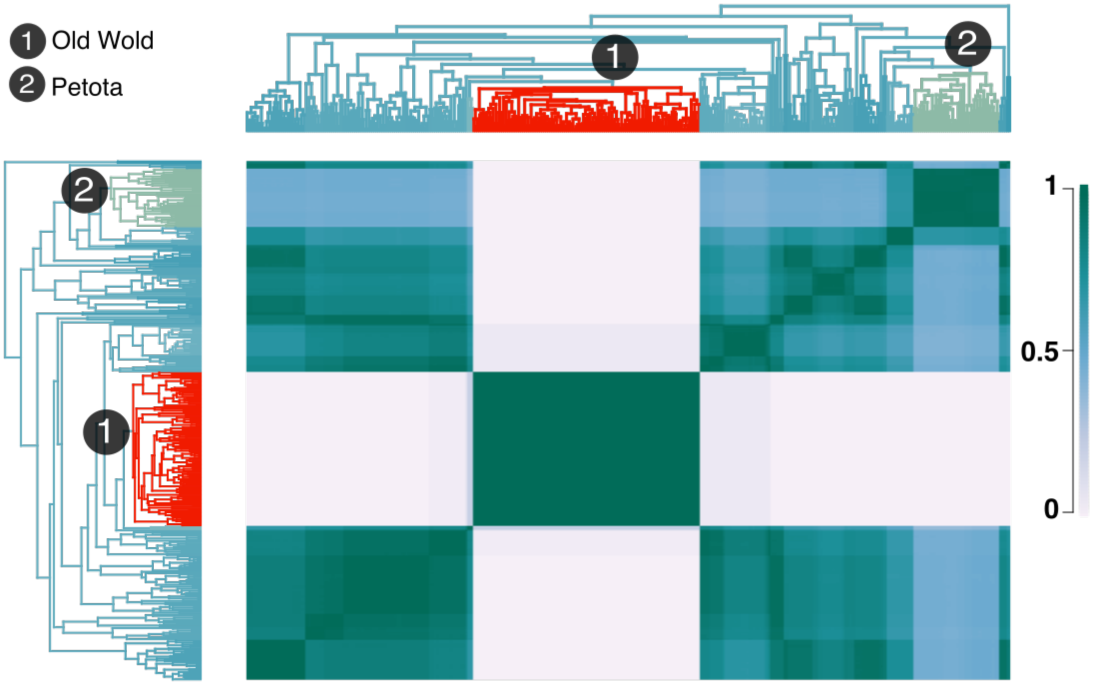


Figure 2.4: Average macroevolutionary cohort across the 100 trees obtained by PASTIS using the *Solanum* species in the Särkinen *et al.* (2013) phylogeny. Macroevolutionary cohort is the distance of diversification rate parameters across the taxa. Lineages are considered to be part of the same macroevolutionary cohort when there is an elevated pairwise probability (> 0.5) and completely decoupled when the probability is 0. For reference, the phylogeny is shown to the right and the top of the matrix.

likely outcome), the zero-shift model was never sampled in the posterior for any of the BAMM runs showing an overwhelming evidence of the heterogeneity of shifts found in this analysis. Moreover, the number of shifts found in the Särkinen *et al.* (2013) phylogeny was not sensitive to different priors of diversification rates (i.e., different values of γ, 0.5,1,2,10,100) as shown in Figure B.6.

By using the *Solanum* phylogeny from Särkinen *et al.* (2013) and assuming an even sampling of species throughout the phylogeny, the TESS approach identifies significant shifts in diversification (Figure B.7). However, when the distribution of complete species phylogenies of *Solanum* produced in PASTIS was used, the signal of diversification shifts through time is no longer supported (Figure B.8).

Overall, the analysis of diversification rates in RevBayes showed similar results to those from the BAMM analysis (i.e., a strong signal of a diversification shift in the node supporting the crown group of the Old world clade, see Figure B.9 and B.10). However, the RevBayes analysis revealed an overall greater heterogeneity of rates compared with BAMM and identified other signal of diversification in groups such as Torva which were not found in BAMM but were found in MEDUSA (see Figure B.9 and B.10)

2.5 Discussion

2.5.1 Main radiations of *Solanum*

This study represents the first attempt to quantify the variation in evolutionary rates across the branches of the phylogeny of the megadiverse genus *Solanum*. The results of most of the diversification analyses used in this study showed agreement, with strong support, for the heterogeneity in evolutionary rates along the phylogeny of Solanum. The great diversity of this genus is likely to be supported by at least two recent radiations — the Old World spiny clade and the Petota subclade. Previous studies such as Whalen (1979) and Whalen & Caruso (1983) were unable to identify unusual rates of speciation within *Solanum*, concluding that the great taxonomic diversity of this genus may reflect a gradual accumulation of species in a relatively old clade (which at the time was believed to have a Cretaceous origin, Hawkes & Smith 1965). However, the results of these previous studies were based on species-poor but well-known sections of *Solanum* such as Lasiocarpa (13 species) and Androceras (12 species) which ignored the great heterogeneity of evolutionary regimes present in the genus. These contrasting results demonstrate the importance of accounting for among-clade rate heterogeneity to understand the complex dynamics of diversification of megadiverse groups such as *Solanum* (Ricklefs, 2007; Morlon, 2014).

2.5.2 Radiation within the non-spiny solanums

In both the MEDUSA and the BAMM analyses, the Petota subclade was identified to have a significant increase in diversification rates (Figure 2.2 and 2.3). With around 104 species, this group represents one of most diverse monophyletic groups within the non-spiny solanums. Faster speciation rates as compared with its sister clade (tomatoes) have been proposed before (Rodriguez *et al.*, 2009), however, this is the first analysis to have quantified and identified this clade as a significant radiation within *Solanum*. Distinctive traits in the potatoes such as polyploidy and self-incompatibility may have played an important role in their range expansion and environmental differentiation which eventually could lead to increases in diversification (Hijmans *et al.*, 2007; Goldberg *et al.*, 2010). The species of potatoes occur in a wide variety of habitats and are predominantly distributed in the middle and high elevation of the Andes (Spooner *et al.*, 2004, 2014). Like other plant groups distributed in high elevation of the Andes, the diversification of the potatoes could have also be driven by ecological opportunities created by the uplifting of the Andes (Hughes & Eastwood, 2006; Drummond *et al.*, 2012; Bell & Donoghue, 2005; von Hagen & Kadereit, 2001). These hypotheses can be tested in the future by explicitly assessing the influence of distinctive traits and niche distribution in the diversification of this very important and intriguing group of *Solanum*.

In contrast to original expectations, I did not find evidence for diversification shifts in the largest major clade of non-spiny *Solanum* — the Geminata clade. The majority of species of this group are found in mesic, rainforest environments, so this result could be an indication of an environmental pattern. However with only the 7% of the species with genetic information in the *Solanum* phylogeny, it is probably not possible to detect any signal in diversification rate heterogeneity within this group, even if polytomy resolvers such as PASTIS are used to place missing species in the tree. Future phylogenetic research in this diverse group of *Solanum* should be pursued, especially since this is the group with the lowest proportion of species sampled in the phylogeny (see Table 2.1).

2.5.3 Old world clade radiation

The Old World clade, within the spiny solanums, represents the most rapid radiation of the genus *Solanum*, diversifying nearly as twice as rapidly as any other group (Figures 2.3 and B.9). This clade is the most diverse monophyletic group within the Leptostemonum clade, also known as spiny solanums (Figure 2.1), with approximately 250 species and has been traditionally defined exclusively by spiny solanums distributed out of the Americas (although Aubriot *et al.* 2016 has recently showed that six species thought to belong to the Old World clade, based on their distribution, now belong to the Torva clade). The Old

World clade shows a significant shift in diversification throughout all the analyses used in this study and has macroevolutionary dynamics that are decoupled from those occurring in all other *Solanum* lineages (see Figure 2.4). This suggests that diversity within the Old World spiny clade has been shaped by different rate dynamics of speciation and extinction than those in the rest of the genus. Although previous phylogenetic studies (e.g., Levin *et al.* 2006; Weese & Bohs 2007, 2010) have considered the Old World spiny clade as an important radiation within *Solanum*, this is the first study to have quantified and therefore demonstrate that the Old World spiny clade is indeed the most rapid and largest ongoing radiation of extant species of *Solanum*.

The Old World spiny clade is not defined by any single distinctive morphological traits or combination of morphological characters (Stern *et al.*, 2011) but instead is defined by its geographical distribution and strong monophyly (Levin *et al.*, 2006; Weese & Bohs, 2007, 2010; Stern *et al.*, 2011). This clade comprises nearly all species of spiny *Solanum* with their native distributions in Africa, Australia, Asia and the Pacific except for few derived species from each of the Lasiocarpa, Acanthophora and Torva clades which are believed to be recent introductions into the Old World, via long-distance dispersal with or without human intervention (Levin *et al.*, 2005; Bohs, 2004; Levin *et al.*, 2006; Aubriot *et al.*, 2016). Other native representatives of *Solanum* in the Old World belong to the non-spiny solanums, more specifically to lineages of the M clade (see Figure 2.1). This clade was highly supported by Särkinen *et al.* (2013), and comprises the subclades Morelloid, Dulcamaroid, Normania, African Non-Spiny and Archeasolanum. The two last subclades are exclusively distributed in Africa (including Madagascar) and Australia (including New Guinea and New Zealand), respectively. However their number of lineages contrasts with number of the Old World spiny lineages found in those same regions (e.g., 17 species of African Non-spiny vs 76 species of the Old World spiny clade distributed in Africa; and 8 species of Archaesolanum vs 119 species of the Old World clade distributed in Australia). These contrasting patterns in diversity between spiny and non-spiny species in Africa and Australia may represent differences in ecological opportunities between these two groups of *Solanum* or differences in their time of diversification.

Looking into this further, Vorontsova *et al.* (2013) notes that the spiny solanums occurring in continental Africa are predominantly concentrated in arid areas or vegetation types such as the “Somalia-Masai regional centre of endemism” sensu White (1983). In contrast, the continental species of the African non-spiny clade are confined in more mesic regions such as the wet tropical forests of central Africa, the Cape region and tropic-montane regions sensu Linder (2014). Likewise, the spiny solanums endemic to Madagascar occupy very different vegetation types from those occupied by the endemic African non-spiny solanums. So generally the African species of the spiny solanums are found in xeric conditions (Vorontsova *et al.*, 2013), in contrast with the African non-spiny, which are mostly found in forested and mesic regions (Knapp & Vorontsova, 2016). Similar to the distribution of spiny and non-spiny solanums in continental Africa and Madagascar, those in Australia are also associated with distinct environmental conditions. Non-spiny solanums in Australia are mainly represented by the Archaesolanum clade also known as the Kangaroo apples. With eight endemic species, Archaesolanums are mainly confined in temperate and mesic areas of the SW Pacific region (i.e., Australia, Tasmania, New zealand and Papua New Guinea) (Poczai *et al.*, 2011). In contrast, the 119 species of spiny solanums are predominantly distributed in warmer and arid regions of Australia (Whalen, 1984; Bean, 2004; Symon *et al.*, 1981). The lineages in this group have an average net diversification rate of 0.68 *lineages/Ma*, which is almost three times that of the Archaeosolanum, 0.25 *lineages/Ma*. This successful expansion and diversification of lineages of spiny solanums in arid zones, in contrast with the more mesic distribution and less explosive diversification of non-spiny species in Africa and Australia, could indicate that the factors that have promoted the radiation of the spiny solanums in the Old World are associated with the adaptation to warm, arid and open environments. The integration of these patterns of diversification with other data sources such as traits, climate data and historical biogeography at the species level will allow us to understand in more detail the specific factors that triggered this significant radiation of *Solanum*.

There are no obvious novel morphological or physiological traits associated with the diversification of Old World species of spiny solanums. However, the signal of change in diversification at the base of spiny solanums in Figure 2.3 and Figure B.9, could give us some indication that distinctive traits of spiny solanums, as a whole, such as stellate indumentum and prickles (Whalen, 1984) could have played an important role in the diversification of the Old World clade. In this context, the Old World clade could be an example of an exaptive radiation where previously acquired traits – originally shaped by different selective forces – are advantageous under a new selective regime (Simões *et al.*, 2016).Furthermore, a significant increase of diversification at the base of the Old World spiny clade (i.e., at the base of the group with most of the spiny solanums distributed out of the Americas) could indicate a correlation between biogeographic changes and shifts in diversification. This nested radiation of lineages of the Old World clade within spiny solanums might be a classic example of Key Confluence sensu Donoghue & Sanderson (2015), where shifts in diversification are associated with the interaction between key innovations and extrinsic factors such as biogeographic movements and/or environmental changes (Moore & Donoghue, 2007). Another typical example of this diversification scenario is described in Drummond *et al.* (2012), where the radiation of the genus *Lupinus* in the Neotropics was the result of the interaction between environmental change — the colonisation of montane regions – with morphological innovations — perennial habit. In the next chapter, I will explore the association between these patterns of diversification in *Solanum* and its historical biogeographic movements and environmental affinities.

2.5.4 Other potential shifts in diversification

With the MEDUSA analysis, I found evidence for potential diversification shifts rates in two groups of *Solanum* — a significant decrease in diversification in the Thelopodium and a significant increase in the Torva clade. The Thelopodium clade, represents a well-supported and isolated clade of *Solanum* (Bohs, 2005; Särkinen *et al.*, 2013). This group is the first branching group of *Solanum* and is the outgroup of the two main clades of the genus. With only three component species, in contrast with more than 1,200 species from its sister clade, it is not surprising that the Thelopodium clade was identified as a significantly depauperate lineage. Several scenarios could explain the significant decrease in diversification in this group. Assuming a diversity-dependent scenario, a significantly depauperate lineage could represent a group of surviving species from a once larger group that experience lower speciation rates and a constant extinction rate (Strathmann & Slatkin, 1983; Donoghue & Sanderson, 2015). Another process which could explain this shift in diversification is the "taxon pulse" scenario of Erwin (1985) where a clade starts diversifying in an ancestral environment but then moves and radiates in another habitat experiencing high extinction rates in the ancestral habitat. Since the signal of evolutionary history in depauperate lineages is eventually lost, it is very problematic to find a robust association between traits, biogeography or other factors with decreases in diversification. This is mainly due to sampling error and also to difficulties in inferring extinction rates using the current methods in macroevolution (Ricklefs, 2006; Rabosky, 2010; Purvis, 2008; Donoghue & Sanderson, 2015).

Unlike Thelopodium, the Torva clade is considered one of the most species-rich groups within the spiny solanums (Stern *et al.*, 2011). This group is distributed mainly in the New world but recently studies have shown recent introductions to the Old World (Aubriot *et al.*, 2016). This section has been of particular interest in phylogenetic and recent biogeographic studies (Stern *et al.*, 2011; Aubriot *et al.*, 2016) since some of its lineages were previously considered to belong to the Old World clade. The phylogenetic structure of this clade is still very unstable and few of its species have been included in molecular phylogenetic studies, therefore, it is important to increase the species sampling and use more molecular markers in its phylogeny, before making any statement about its evolutionary dynamics.

2.5.5 Reliability of methods for estimating heterogeneity in diversification rates

There is an ever-increasing number of available methods to study diversification dynamics which entail different fundamental questions in macroevolution (see Morlon 2014 for a review of these methods). The development of approaches that aim to identify significant diversification-rate shifts along branches is very recent and the consequences of the assumptions of these various approaches are still poorly known. However, recent publications such as Rabosky & Goldberg (2015), Moore *et al.* (2016) and May & Moore (2016) have triggered a heated debate about model adequacy and the reliability of inference of widely applied approaches in diversification studies including MEDUSA and BAMM; as well as approaches that detect diversification-rate correlates such as BiSSE (Binary State Speciation and Extinction, Maddison *et al.* 2007). The nature of these debates tends to be extremely technical focusing mainly in the software implementation, the algorithms and the mathematics behind each approach. In this analysis, I did not aim to compare the performance and reliability of the different approaches available in diversification studies, instead, I used different approaches to testing the robustness of the results drawn in this study.

The MEDUSA approach (Alfaro *et al.*, 2009) was one of the most popular ways to detect lineage-specific diversification rates before the advent of BAMM (Rabosky, 2014). Although several concerns about the reliability of model selection and accuracy of rate parameters in MEDUSA have recently raised (May & Moore, 2016), this approach provides a useful exploratory tool as it is computationally very efficient. Using this framework we can identify groups with unusual shifts in diversification and later use this information as a priori hypotheses to test whether these predefined diversification-rate shifts are significant using approaches such as BayesRates (Silvestro *et al.*, 2011). One of the shortcomings of MEDUSA is that it assumes constant diversification through time; it also only provides point estimates without including uncertainty in the parameter estimates. The BAMM approach, which is the main technique used in this study, was developed to resolve these issues. However, as with any approach that implements a Bayesian statistical framework, it is important to evaluate the consequences of the prior assumptions on the location and magnitude of estimates in diversification.

Moore *et al.* (2016) recently raised concerns about the reliability of the results obtained from BAMM analyses. One of the main issues claimed in this paper is the extreme sensitivity of the posterior distribution of the diversification-rate inferred to the assumed prior. Here I demonstrated that the estimated numbers of diversification shifts from the BAMM analysis in *Solanum* are robust to the choice of prior models (see Figure

S2-S6). These results are consistent with the findings presented in Mitchell & Rabosky (2016) and Rabosky *et al.* (2017), who claim that the unusual effect of priors on the estimated numbers of shifts found in Moore *et al.* (2016) can be explained by an error in a previous software implementation of BAMM (version < 2.3.1). Other BAMM issues such as errors in the likelihood function and the use of compound Poisson distribution, which describes the prior distribution of diversification rate shifts, are more difficult to evaluate here. However, to test the robustness of the results in this study I used other recently developed approaches, TESS and RevBayes, to detect diversification shifts.

The TESS method (Höhna *et al.*, 2016b) detects tree-wide shifts in diversification through time. In the TESS analysis, although I found significant diversification shifts through time using the original Särkinen *et al.* (2013) phylogeny (Figure B.7), there was no evidence of tree-wide diversification shifts through time using the pooled phylogenies produced by the polytomy resolver PASTIS (Figure B.8). As in other approaches such as TreePar (Stadler, 2011), the TESS analysis assumes that the diversification processes occur tree-wide and that all lineages at a given time slice have the same rates. These contrasting results in TESS when the complete species phylogenies were analysed could be explained by the significant proportion of species included by PASTIS in some of the *Solanum* clades (e.g., more than 90% within the Geminata clade) that infer the timing of the missing splits under a constant-rate birth-death model (Thomas *et al.*, 2013). This conservative way to include missing species into the Särkinen *et al.* (2013) phylogeny could have created biases towards detection of constant-rate models when whole-tree methods such as TESS was implemented.

Another recently developed approach to identify diversification rates along branches is the RevBayes analysis (Höhna *et al.*, 2016a). This approach has not been implemented in any empirical example so far. Like BAMM, this analysis follows a Bayesian statistical framework to identify and quantify branches with significant changes in diversification rates. Using a distribution of 20 trees produced by PASTIS, I found, surprisingly, very similar results in the location and the magnitude of diversification rates along branches to those that I found with the BAMM analysis (Figure B.9 and Figure B.10). As in the BAMM analysis, the Old World spiny clade was supported as the most rapid radiation within *Solanum* and the Petota subclade as the most rapid radiation within the non-spiny solanums. Interestingly, other clades such as the Torva clade, previously identified in MEDUSA but not in the BAMM analysis, showed differences in rates compared with the background. This higher heterogeneity in rates found in Revbayes could indicate that the dynamics of diversification in *Solanum* are more complex than inferred by BAMM. However, as in any recently developed approach, the consequences of the prior assumptions in the RevBayes analyses are poorly known. Therefore, it is important to consider this analysis of *Solanum* as exploratory. The study of diversification patterns is a dynamic and ongoing field with recently developed approaches whose reliability still needs to be assessed. Therefore, it is essential to be critical of any results until the consequences of the different assumptions and parameters have been explored. However, despite the implementation and conceptual challenges generated by the growing biological data, the careful application of these imperfect inference tools help us to reveal fundamental insights into the processes that generate current patterns of biodiversity.

2.5.6 Concluding remarks

The study of recent radiations can provide a more accurate window into the processes that generate and maintain biological diversity. Here I demonstrated that *Solanum* is a remarkable study system for understanding the diversification of plants. Not only because it contains a recent radiation into the Old World, but also because the diversity of this genus is supported by several clades with different and sometimes contrasting evolutionary dynamics. There are still many drivers that need to be explored before we have the whole picture of the diversification of *Solanum*, especially correlations with biogeographic events and arid-adapted traits. However, thanks to the collected effort and data from *Solanum* specialists worldwide, we are beginning to uncover many aspects of the macroevolution of this mega-diverse plant genus.

CHAPTER 3

Historical biogeography and climate affiliations of *Solanum*

3.1 Abstract

The imbalance of species richness across space reflects the differences in the dynamics of historical events and diversification among lineages. Here I synthesise phylogenetic and distributional data for an ongoing radiation — the megadiverse plant genus *Solanum* L. — to show how dispersal and climatic change have interacted to shape diversification. In *Solanum*, the Old World lineages have radiated more rapidly than those in the New World, with an explosive diversification coinciding with a long-distance dispersal events from the Neotropics ca. 6 Ma. Within Australia, the lineages adapted to arid environments have radiated whereas more mesic lineages show average net rates of diversification, consistent with the continent's long-term climatic trend. These findings provide a clear example of how successful colonisation of new areas and niches can — but do not always — drive explosive diversification.

3.2 Introduction

Changes in the earth's environmental and biogeographic history have left an imprint in the distribution, the size and the composition of regional species pools. The evolution of the enormous richness of plant species has been modelled by the interplay between intrinsic (morphological and/or physiological novelties) and extrinsic events (biogeographic and/or climatic events) through time (Moore & Donoghue, 2007; Donoghue & Sanderson, 2015). In order to disentangle the effect of any of these variables on the current patterns of diversity is essential to integrate historical and ecological information (Ricklefs *et al.*, 1999; Ricklefs, 2006; Wiens & Donoghue, 2004).

Differences in diversification rates (i.e., the net production of lineages per unit time) among taxa lead to both species-rich and depauperate groups. These differences in the per-lineage rates of speciation and extinction among groups are usually the result of differential responses to ecological opportunities through the extinction of antagonists, the acquisition of a key innovation or the colonisation of a new habitat (Simpson, 1955; Schluter, 2000; Moore & Donoghue, 2007; Yoder *et al.*, 2010; Purvis *et al.*, 2011). Integrating historical biogeography into diversification studies provides an opportunity to explore the effect of biogeographic movements on the diversification of taxa. This integration not only helps us to model the expansion, migration, diversification and extinction of certain taxa but also helps us to uncover the ecological processes that influenced their current patterns of diversity (Moore & Donoghue, 2007; Donoghue & Sanderson, 2015). By exploring the historical biogeography of taxa, we can potentially understand why certain lineages have been excluded from regional biota whereas others have successfully colonised, expanded and diversified (Moore & Donoghue, 2007; Donoghue & Sanderson, 2015; Wiens & Donoghue, 2004; Cavender-Bares *et al.*, 2016).

The growing availability of large phylogenies along with the development of macroevolutionary methods provides an opportunity to explore complex biogeographic events usually found in diverse and cosmopolitan taxa. Here, I investigate the spatial and temporal patterns of continental-scale diversity of a widely distributed and economically important plant genus *Solanum* L (Solanaceae). This genus is one of the ten most species-rich genera in Angiosperms (Frodin, 2004) and has more than 1200 species. *Solanum* occurs in all temperate and tropical regions in the world with a predominant distribution in the south hemisphere. With almost 70% of *Solanum* species, the Neotropics represents its main centre of diversity and endemism (Olmstead & Palmer, 1997), however, significant diversity also occurs in some places of the Old World such as Australia and Africa (Knapp & Vorontsova, 2016; Vorontsova *et al.*, 2013; Aubriot *et al.*, 2016; Poczai *et al.*, 2011). This diversity distribution varies between the two principal divisions of *Solanum* — spiny and non-spiny *Solanum*s. With a strong monophyletic support, the spiny solanums (known as the subgenus *Leptostemonum* Bitter) has been long recognised as a major division of *Solanum.* This group is distinctive by their stellate trichomes, long tapering anthers and usually prickled traits (Vorontsova *et al.*, 2013). With approximately 420 species, the centre of diversity of spiny solanums is divided mainly among the Neotropics (ca. 150), Australia (ca. 130) and Africa (ca. 79). The last two are mainly represented by the distribution of lineages of the most diverse monophyletic group within spiny solanums known as the Old World clade (Stern *et al.*, 2011; Särkinen *et al.*, 2013). This high diversity of spiny solanums out of the Americas contrasts with the distribution of the diversity of the non-monophyletic group the non-spiny solanums, which has more than 90% of its diversity distributed in the Neotropics.

Previous historical biogeographic studies have shown that the family Solanaceae and the major clade Solanae (*Solanum* + *Jaltomata genera*) have a neotropical origin with several dispersal events to the Old World (Dupin *et al.*, 2016; Olmstead, 2013). The large size of *Solanum* has limited the study of its biogeography to smaller groups and regions (Poczai *et al.*, 2011; Aubriot *et al.*, 2016; Weese & Bohs, 2010; Vorontsova *et al.*, 2013; Hijmans *et al.*, 2007; Olmstead *et al.*, 2008). However, due to the recent species-level phylogeny of Solanaceae (Särkinen *et al.*, 2013) and the collaborative effort to compile species-level distribution through the web resource Solanaceae Source (`http://www.solanaceaesource.org`), it is now possible to explore the historical biogeography of the genus *Solanum* as a whole.

Divergence time estimates for the crown clade of *Solanum* range from 13 to 18 Mya* and for the stem clade from 15 to 19 Mya (Särkinen *et al.*, 2013). Based on the current distribution of *Solanum* lineages, it is likely that its diversification occurred in the Neotropics with several successful colonisations to the Old World, especially to Africa and the South-West Pacific (Olmstead *et al.*, 2008; Bohs, 2005; Olmstead & Palmer, 1997). Moreover, a single colonisation event could explain the monophyletic relationships of the majority of species of spiny solanums distributed in the Old World (Levin *et al.*, 2006; Olmstead *et al.*, 2008). In addition, the wide range of habitats that *Solanum* lineages occupy suggests that the evolution into a new environment could have occurred along with geographic movements or in situ as the environment changed. Therefore, understanding the main historical geographic movements and the extent of niche conservatism within the genus can help us reveal the factors that affected the diversification and colonisation of *Solanum* across the globe.

In this study, I aim to reconstruct the spatiotemporal evolution of *Solanum* using the most up-to-date phylogenetic and distributional information. Specifically, I ask: (1) What are the main historical biogeographic events that have shaped the current distribution of lineages in *Solanum*? (2) Are any of these biogeographic events associated with shifts in diversification?

3.3 Methods

Geographical patterns of diversification in *Solanum*

To determine which regions have accumulated and are accumulating a higher or lower number of *Solanum* lineages across the globe, the average diversification rates per species

---

*million years ago

obtained in the BAMM analysis in Chapter 2 were displayed to a 1 x 1 degree grid scale map.

84,606 botanical records from 1005 *Solanum* species were extracted from the Solanaceae Source database `http://solanaceaesource.org/` on 13 March 2016; and 34,462 records from 215 species were compiled from the Australian National Herbarium, accessed through the Atlas of Living Australia (ALA) website `http://www.ala.org.au/`, on 16 April 2016. I applied a series of quality control filters to discard the following:

1. Records with latitude and longitude coordinates of 0° , 0° .
2. Records matching coordinates from major herbaria or political centroids at all administrative divisions extracted from Edwards *et al.* (2015).
3. Any records with the described country conflicting with the country extracted from their coordinates. This step was performed by overlaying the records with a global administrative polygons extracted from the R package "rworldmap" (South, 2011) using the function *"over"* from the R package "sp" v 1.1-0 (Pebesma & Bivand, 2005).
4. Records with identical coordinates for a given species
5. Records from non-native, cultivated or naturalized species

Taxonomic names were updated for all records to correct for synonymy using as a reference the accepted name list from `http://solanaceaesource.org/`. After cleaning, I was left with 64,826 unique records for 1,096 taxa.

Species occurrences were then converted into a presence-absence matrix of a 110 x 110 km equal area grid using the function *"lets.presab.points"* in the R package "letsR" v 2.1 (Vilela & Villalobos, 2015b). Using the mean species-specific rates, I estimated the mean assemblage diversification rates as the geometric mean of all species' rates present in a grid cell. I also computed a weighted version of this, dividing the mean species-specific diversification rates by the inverse of their range size — log of the area (sqm) occupied by each species, to correct for the overestimation of rates in an area as a result of the occurrence of widespread species.

Finally, I reconstructed lineages-through-time plots for the three principal diversity centres for *Solanum* (Neotropics, Australia and Africa) using the R package "paleotree" v 2.3 (Bapst, 2012) to visualise the geographical differences in lineage accumulation dynamics among continents.

3.3.1 Ancestral range reconstruction

I investigated the historical biogeography of *Solanum* using the R package "BioGeoBEARS" (Matzke, 2012, 2014) — BioGeography with Bayesian (and likelihood) Evolutionary Analysis in R Scripts. This approach provides a statistical framework to compare traditional models in biogeography such as DIVA (Dispersal-Vicariance Analysis; Ronquist, 1997), DEC (Dispersal-Extinction-Cladogenesis; Ree *et al.*, 2005), and BayArea (Landis *et al.*, 2013). Each of these models assumes different processes to reconstruct the ancestral range of lineages. In all of the models, species ranges are allowed to change along the branches by anagenetic evolution through two main events: dispersion (range expansion) and extinction (range contraction). The events allow in cladogenesis vary depending on the fitted model. For example, sympatric speciation subset, or peripatric speciation (i.e., the ancestral range, which is widespread, is completely inherited to one of the daughters whereas the other inherits the range of just a single area) is only allowed in the DEC model. Unlike the DIVA and the BayArea models, the DEC model only allows one daughter to inherent widespread distributions during a vicariant event, since the widespread distribution of both daughters would assume additional events such as post-speciation dispersal (Ree *et al.*, 2005). In all models, sympatric speciation (i.e., within-area speciation or duplication) is allowed. In addition to these classic events of biogeography, BioGeoBEARS includes a founder-event speciation event (+j) which considers the influence of speciation through long-distance dispersal that are common in Island-like models (see Figure 3.1).

Phylogenetic relationships among species of *Solanum* were taken from the time-calibrated clade credibility (MCC) tree from Särkinen *et al.* (2013). This phylogeny represents the current state-of-the-art of *Solanum* relationships published so far. Several species were pruned from the tree as follow: (1) species which are considered widely cultivated or with ambiguous native distribution, (2) species considered as synonyms, and therefore duplicated in the tree, and (3) species with a low support of a direct ancestor-descendant relationships leading to negative branches in the MCC. The final pruned phylogeny used for subsequent analyses contained 386 species of *Solanum* with the addition of *Jaltomata andersonii* as the outgroup.

Using the extracted distribution records, I created an occurrence matrix of *Solanum* species into six biogeographic areas — Africa, Australia, Indo-Pacific, Neotropics, Nearctic and Palearctic based on the floral kingdoms defined by Cox (2001). The distribution of widespread species was double checked and corrected for potentially recent cultivated and/or naturalised species using the native descriptions available from `http://solanaceaesource.org/` or specific taxonomic monographs (see Appendix C.1). The

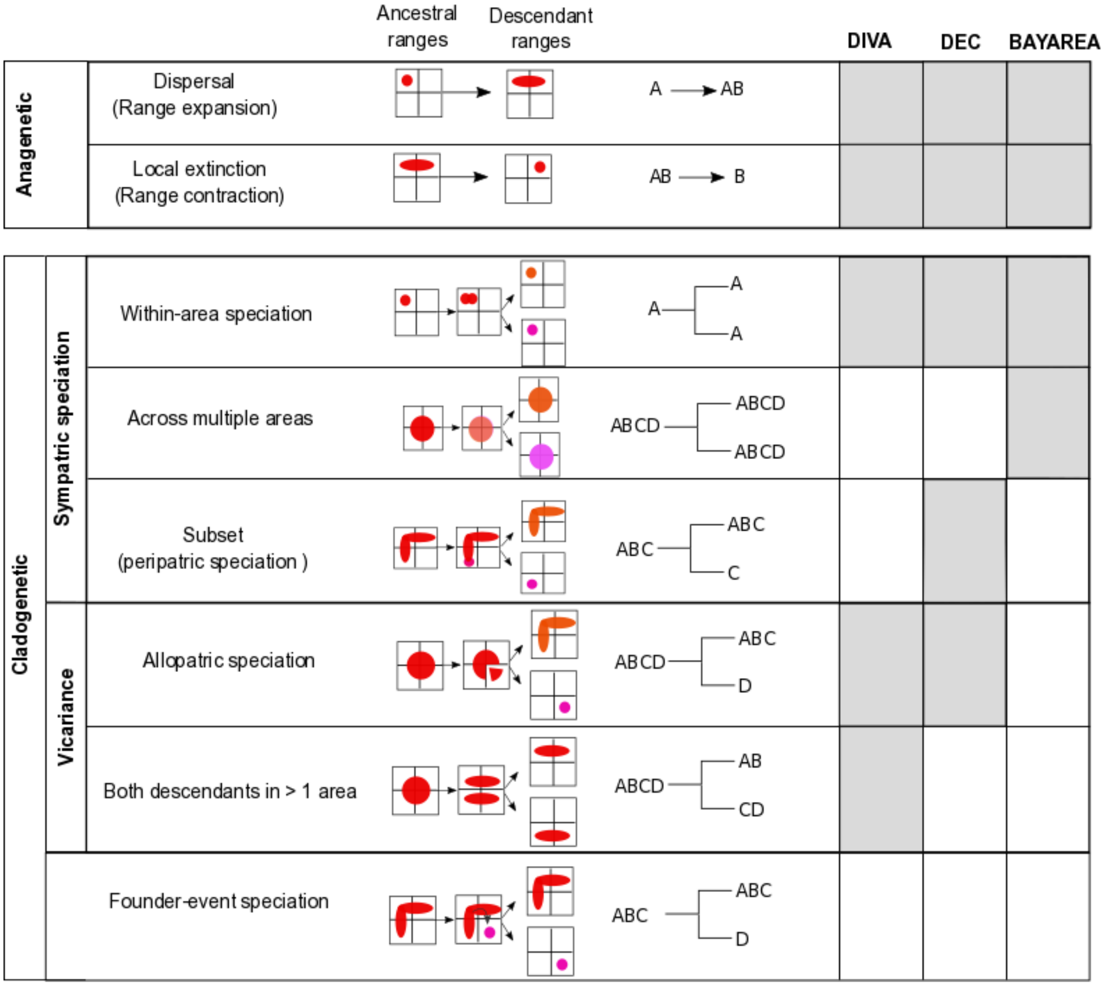


Figure 3.1: Historical biogeographic events model in BioGeoBEARS. Summary of anagenetic and cladogenetic events of geographic range evolution allowed in the biogeographic models DIVA (Ronquist & Sanmartín, 2011), DEC (Ree *et al.*, 2005) and BAYAREA (Landis *et al.*, 2013) implemented in BioGeoBEARS. Modified from Ronquist & Sanmartín (2011)

maximum number of areas was set to three to avoid the influence of obscured species distribution by extensive human transport and also to reduce the complexity and computational time in the analysis.

Three biogeographic models were fitted in the *Solanum* phylogeny and associated geographic distribution using BioGeoBEARS — DIVALIKE, DEC and BAYAREALIKE. The influence of founder-event speciation event (+J) was also included into each model, resulting in a total of six models. An additional set of six biogeographic models were fitted using a dispersal matrix multiplier to weight the dispersal probability of adjacent areas as 1, 0.5 and 0.001 for easy, medium, and hard dispersal (see Table C.2). The model that best describe the empirical data (i.e., optimal fixed model structure) was then chosen using a stepwise selection from the candidate models ranking under the Akaike Information Criterion (AIC; Burnham & Anderson, 2002).

Once the best-fitting model was identified, I estimated the overall probabilities of the anagenetic and cladogenetic events conditional on the model, the phylogeny, and the geographic distributions from 100 Biogeographic Stochastic Maps (BSM) to account for the uncertainty in the state transitions and ancestral range reconstructions (Matzke, 2016). These stochastic maps are similar to simulations of trait change along phylogenies (Huelsenbeck *et al.*, 2003) using transition rate models (Pagel, 1999). Given the observed range data, the phylogeny and the best fitting model of biogeographic events, the BSM simulates possible histories constraint to the observed ancestral ranges. An average of all the probabilities of the simulated histories will result in the same ancestral state probability obtained under the best fitting model (Matzke, 2016).

3.3.2 Climatic space

For each of the species records collated, I extracted the 19 BIOCLIM variables from Worldclim (Hijmans *et al.*, 2005) using the package "raster" (Hijmans, 2015) at a spatial resolution of 0.5 degrees (55.6 km at the equator). I also extracted the mean potential evapotranspiration (PET) and aridity index (AI) from the CGIAR-CSI Global-PET and Global-Aridity Database (Zomer *et al.*, 2007, 2008), respectively, at a spatial resolution of 30 arc-seconds (ca. 1km at the equator) to measure potential vegetative growth. Geographical manipulation and spatial analyses were performed in the R packages "maptools" v. 0.8-39 (Bivand & Lewin-Koh, 2016). Each of the variables was standardised (to zero mean and unit standard deviation) and transformed to follow a Gaussian distribution (using either log10 or sqrt). Species with only one record were excluded from the analysis.

Using the median species values, I reduce the dimensionality in the correlated climatic variables through Principal Component Analysis (PCA) using the R package "vegan" v. 2.4-1 (Oksanen *et al.*, 2016). I then used the first two principal components to visualise the climatic space among main sections of the *Solanum*.

To investigate the evolution of biome transitions among lineages of *Solanum*, I reconstruct the ancestral biome states along the tree of the genus using the R package "BioGeoBEARS" in the same way as in the historical biogeographic analysis (Matzke, 2012, 2014). Distributions of all species were map against the 14 World Wildlife Fund "Ecoregions", which are based on Olson *et al.* (2001), to create a presence-absence matrix of species in the biomes. Unlike the biogeographic areas used in the ancestral range reconstruction, the biomes used in this analysis do not represent shared biogeographic history but instead share climatic conditions among species. The biome ancestral reconstruction was then analysed setting the maximum number of regions per species to four to reduce computational complexity. Additionally, 50 Biogeographic Stochastic Maps (Matzke, 2016) were run to account for uncertainty in the frequency and the directionality of events across biomes. Transition frequencies and directionality were plotted using the R package "qgraphs" (Epskamp *et al.*, 2012).

3.4 Results

3.4.1 Geographic patterns of diversification

Although most of the diversity and endemism of *Solanum* is found in the Neotropics (ca. 70% of species, see Figure 3.2A), the highest diversification rates are seen in lineages mainly concentrated in Australia, Africa and the Indo-Pacific (Figure 3.2B and 3.2C). The heterogeneity of rates in Africa results from the distribution of species from groups with standard diversification rates such as Morelloids (0.27 *lineages/Ma*) and African non-spiny (0.25 *lineages/Ma*), and species from the group with high diversification rates — the Old World clade (0.68 *lineages/Ma*). Australia shows an interesting latitudinal heterogeneity of diversification rates (Figure 3.3) shaped by the distribution of lineages of different evolutionary origins — species from the Archaesolanum clade mainly distributed in the south with 0.20 *lineages/Ma* and the widespread distribution of the Old World spiny species with 0.68 *lineages/Ma*.

Lineages-through-time curves (Figure 3.4) show that most of the diversification in *Solanum* has occurred within the Neotropics. At the global level, there is a slight acceleration in the number of lineages of *Solanum* in the last 5 Ma, which is likely to be shaped by the considerable increase of lineages in Australia around the same time.

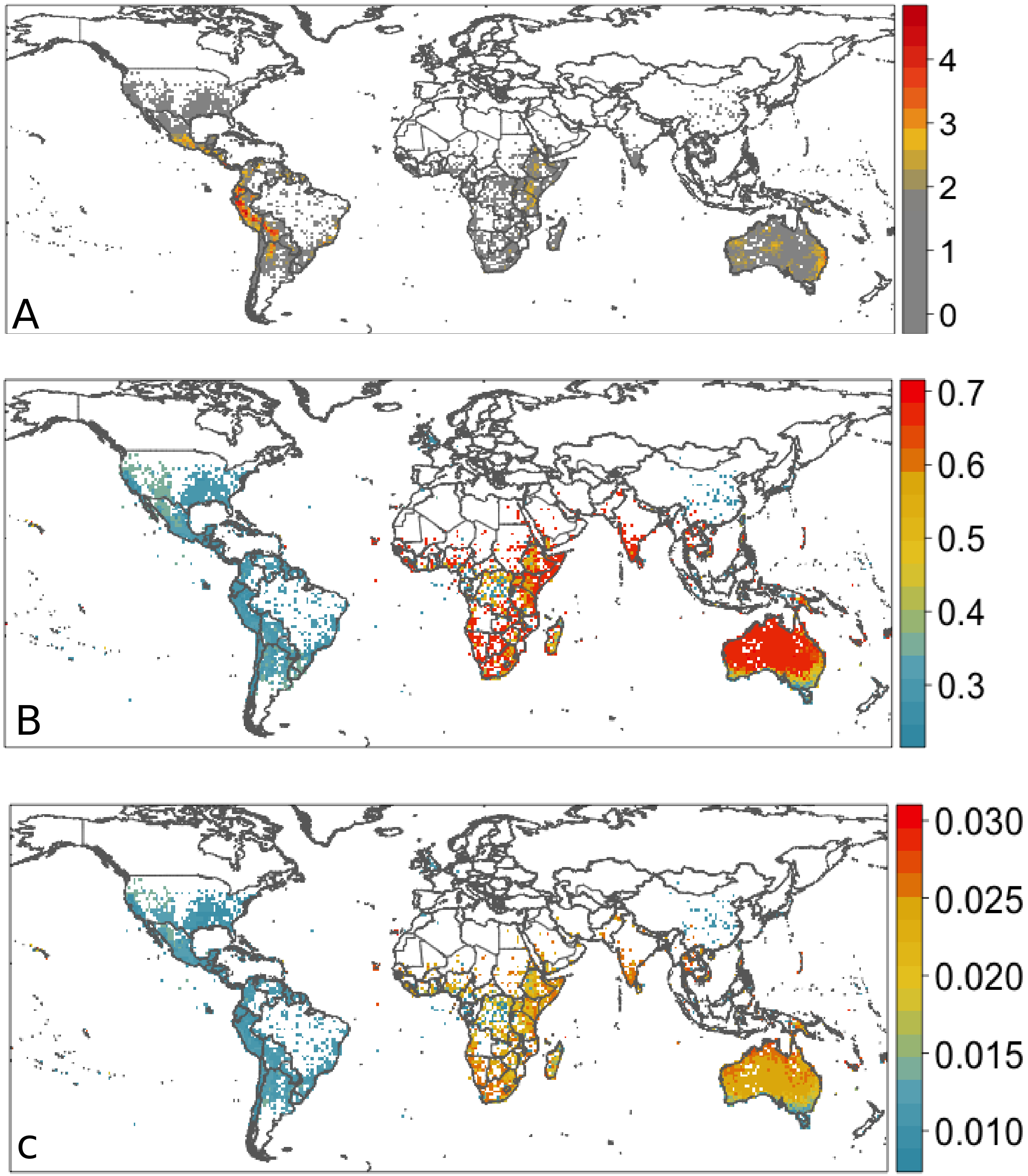


Figure 3.2: Global distribution of the net diversification rates of *Solanum*. A. Global distribution of species richness in *Solanum* in a log_{10} scale **B.** Mean lineage net diversification rates with a grid resolution of 1 x 1 degrees. **C.** Weighted mean lineage net diversification rates. This was calculated as the mean net diversification rate for all the species present in a grid cell assemblage, weighted by the inverse of their range size, $log_{10} m^2$.

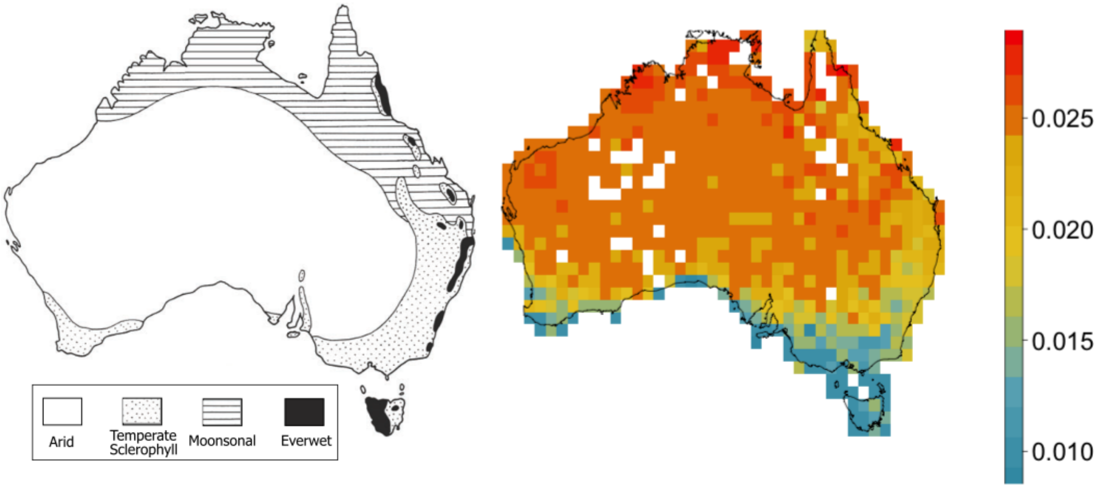


Figure 3.3: Distribution of the net diversification rates of *Solanum* in Australia. A. Main contemporary biomes of Australia based on Crisp *et al.* (2004) and Crisp & Cook (2013). **B.** Distribution of the weighted net diversification rates of *Solanum* in Australia (calculated as the mean net diversification rate for all the species present in a grid cell assemblage, weighted by the inverse of their range size, $log_{10}m^2$. Grid resolution is 1 x 1 degrees).

3.4.2 Historical biogeography

Ancestral range reconstructions across the *Solanum* phylogeny are best explained by the DEC M1 model (Dispersion, Extinction and cladogenesis), which allows for equal probabilities of dispersion from the Neotropics to both Africa and Australia (M1, Table C.2). Including the founder-event speciation event into the model did not improve the AIC values significantly (Δ AIC = 0.5, see Table 3.1), therefore, all the historical biogeographic results, including the Biogeographic Stochastic Mapping, in this study were based on the simpler model, DEC M1.

The genus *Solanum* appears to have risen in the Neotropics ca. 15 Mya (95% HPD[†] 13-18) as well as the majority of its subclades (see Figure 3.5). The distribution of *Solanum* appears to have been modelled mainly via within area sympatric speciation and dispersion, with vicariance only supporting the 3% of the total number of events (see Table 3.2). The distribution of *Solanum* in the Nearctic region is supported by at least seven dispersal events out of the Neotropics. Although several subclades are currently distributed in the Old World, only two subclades appear to have arisen via historical events (contrary to recent introductions). These dispersal events from the Neotropics to the Old World occurred at different times and in two different groups of *Solanum*. The first movement from the Neotropics to Africa + Australia is likely to have occurred

---

[†] highest posterior density 95% confidence intervals

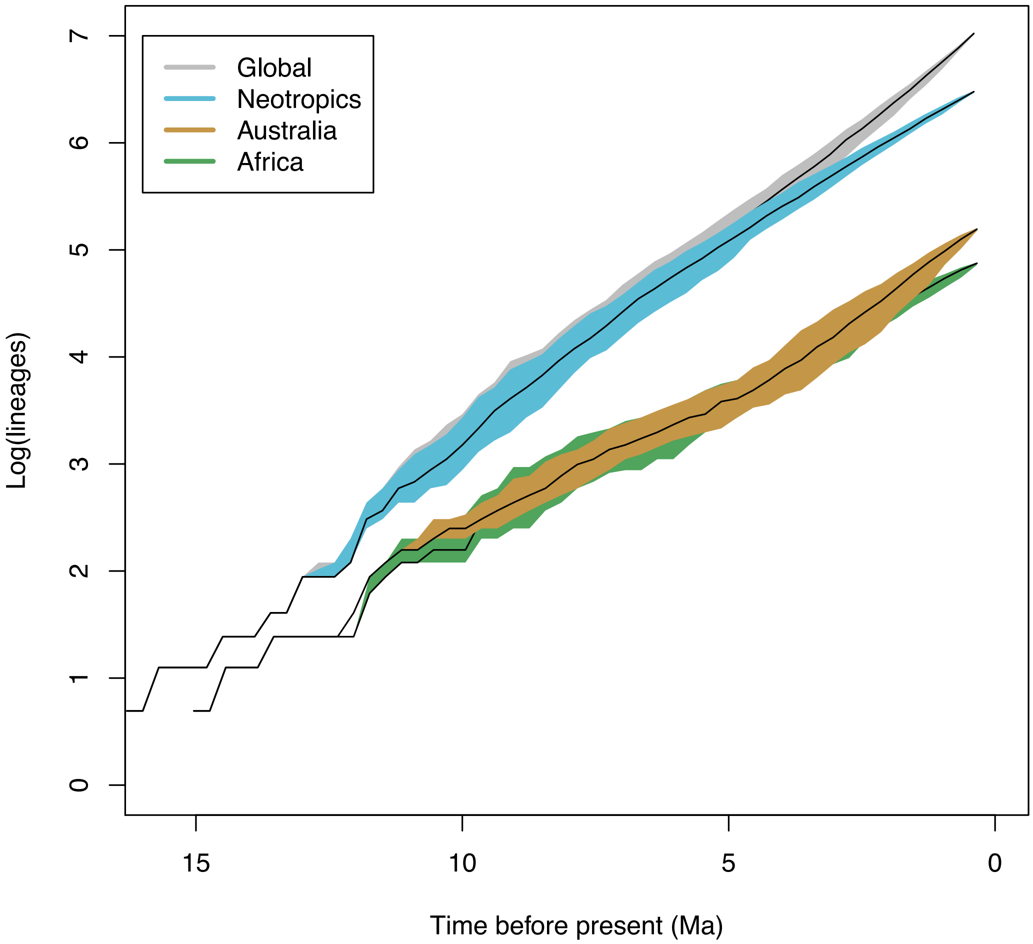


Figure 3.4: Lineages-through-time plot of *Solanum* species for the global and three regional diversities. Region around the median curve is the two-tailed 95 % lower and upper quantiles of the number of lineages across the 100 *Solanum* trees obtained in chapter 2

Models	lnL	*k*	*d*	*e*	*j*	AIC	Δ AIC
DEC+J M1	-302.53	3	0.017	0	0.003	611.06	0
DEC M1	-303.81	2	0.019	0	–	611.62	0.55
BAYAREALIKE+J M1	-305.92	3	0.013	0.002	0.009	617.84	6.77
DIVALIKE M1	-307.94	2	0.02	0	–	619.89	8.82
DIVALIKE+J M1	-306.96	3	0.019	0	0.003	619.92	8.85
BAYAREALIKE+J M0	-319.65	3	0.006	0.001	0.005	645.31	34.24
DEC+J M0	-320.19	3	0.008	0	0.002	646.39	35.33
DEC M0	-322.08	2	0.009	0	–	648.16	37.09
DIVALIKE+J M0	-327.68	3	0.009	0	0.002	661.36	50.30
DIVALIKE M0	-329.10	2	0.01	0	–	662.21	51.15
BAYAREALIKE M1	-364.11	2	0.017	0.019	–	732.23	121.17
BAYAREALIKE M0	-378.47	2	0.008	0.019	–	760.94	149.88

Table 3.1: Estimated parameters, log-likelihood and AIC values of the biogeographic models tested in BioGeoBEARS. Models were ranked based on their AIC values. +j models allowed founder-events. M1 models included a matrix that weight the dispersal probability of adjacent areas. lnL= log-likelihood, *k*= number of parameters, *d*=rate of range expansion, *e*=rate of range contraction, *j*= rate of jump dispersals.

ca. 10 Mya (95% HPD 7-12) within the non-spiny solanums, in the crown group of the subclades *Solanum valdiviense* + African-non-spiny + Normania + Archaesolanum (from now on called the M1 subclade, see Figure 3.5). The second dispersal event occurred ca. 6 Mya (95% HPD 5-7) within the spiny solanums in the subclade Elaeagnifolium + Old World spiny clade (see Figure 3.5). The stochastic mapping estimates suggest that the direction of these dispersal events from the Neotropics to either Africa and Australia is still unresolved in both the non-spiny solanums (with a probability of 30% to disperse from the Neotropics to Africa, and 24% to Australia) and the spiny solanums (with a probability of 45% to disperse from the Neotropics to Africa , and 29% to Australia). Biogeographic movements at several points within the Old World clade are still not well resolved.

Overall, the Neotropics was the main source of *Solanum* movements with more than the 60% of the estimated dispersal events (Figure 3.6). The highest number of dispersals involved movements from the Neotropics to the Nearctic with ca. 40% of the total estimated events. Movements from the Neotropics to Australia appear to be more frequent than those from the Neotropics to Africa (8.7 ± 1.1 vs 2.9 ± 0.9). Dispersals from Africa to Australia are more common than those in the opposite direction (5.5 ± 1.2 vs 1.64 ± 0.9).

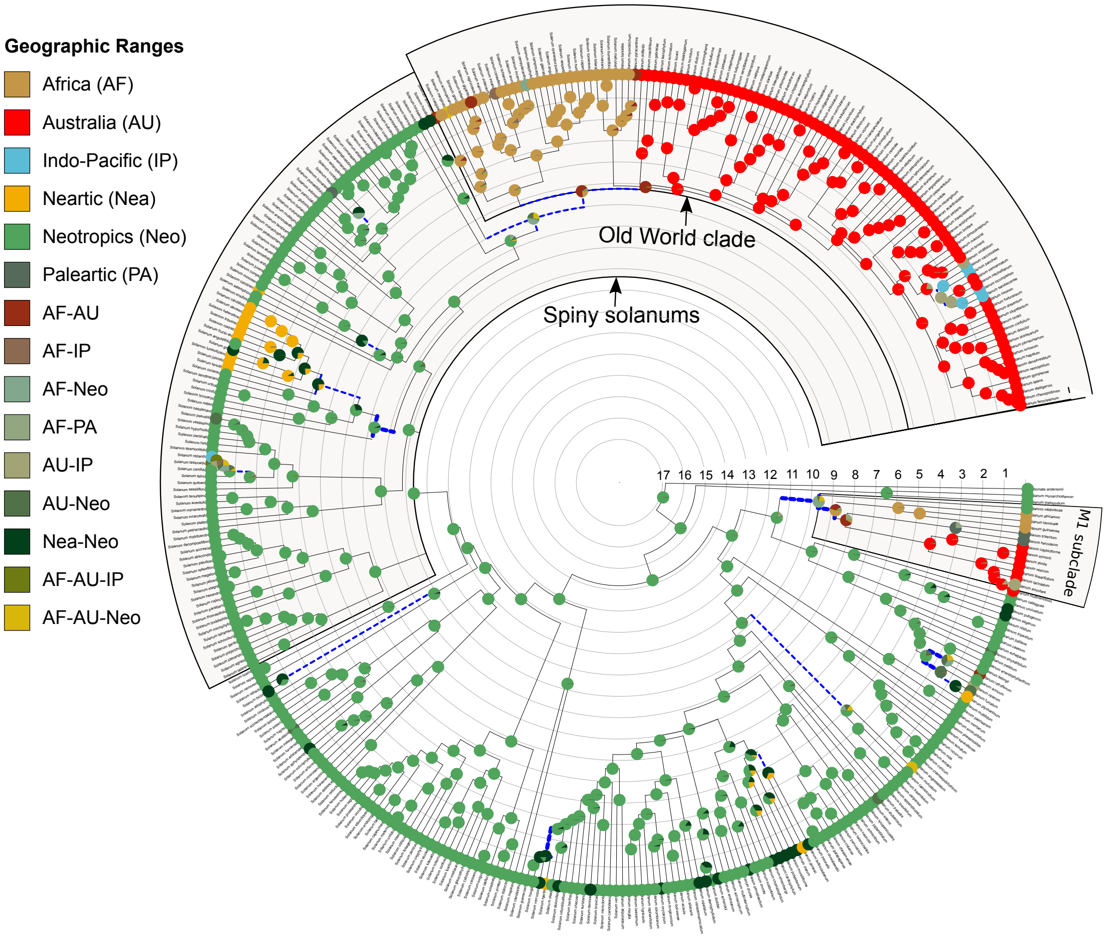


Figure 3.5: Reconstruction of the historical biogeography of *Solanum* under the DEC model implemented in BioGeoBEARS. Pies at each node represent the probability that each region (or the combination of regions) is, according to the model, the ancestral range distribution. The highlighted branches represent the dispersals events inferred in at least the 50% of BSM simulations. Thicker branches show dispersals inferred more than the 95% of the BSM simulations. The timescale is given in Ma.

Event	Type	Mean	SD	Percentage
Sympatric speciation	within-area speciation	354.6	3.65	80.0
	subset (peripatric speciation)	17.36	4.16	3.9
Dispersal	Range expansions	57.52	1.19	13.0
	Range contractions	0	0	
	Founder events	0	0	
Vicariance	Vicariance	13.99	1.17	3.2
Total events		443.47		100.0

Table 3.2: Mean number of biogeographic events estimated across the 100 biogeographical stochastic simulations using the DEC M1 model. No range contractions or founder events were estimated since the inclusion of these parameters did not improve the model significantly.

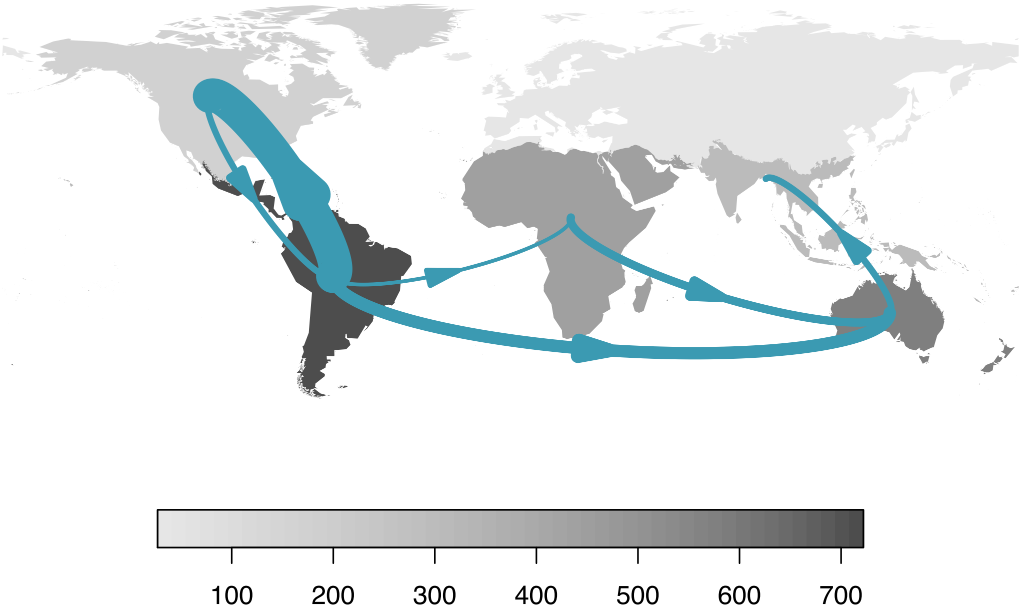


Figure 3.6: Summary of dispersal events within the main biogeographic regions of *Solanum*. The arrows between regions represent the frequency and direction of dispersal events. The bar represents the total number of species within each region. Only dispersal events with a mean of two or more counts are shown. The thick of the arrows describes the frequency of the events.

3.4.3 Climate space and biome transitions

The botanical records included in the principal component analysis included 55,926 georeferenced *Solanum* specimens, with an average of 59 specimens per species (ranging from 2 to 1,863). According to this analysis, the climatic variables that explained most of the variation in the climatic space of *Solanum* species are annual mean temperature, annual mean precipitation, max temperature of warmest month, min temperature of coldest month, temperature seasonality, and isothermality. Climate PC1 is most strongly related to precipitation, whereas Climate PC2 is more supported by extreme temperatures (as a limiting environmental factor) and seasonality (see Figure 3.7A). Most of the seasonal and extreme heat tolerant species in *Solanum* belong to the Old World clade, specially those species occurring in Africa and Australia. Several species of the Old World also occurred in humid habitats but these are mainly distributed in the Indo-Pacific region (see Figure 3.7A)

Mapping individual climate space on the phylogeny of *Solanum* reveals that most of the species of the spiny solanums occupy drier and aseasonal environments (Figure 3.7B). However, this tendency could be bias by the distribution of the potato clade species which occupy high mountainous regions in the Neotropics which tend to be cold and arid. In addition, there is a prevalence of species with higher heat tolerance and more seasonal species within the Old World clade than in any other group of *Solanum*.

The mapping of the frequency and direction of main movements of *Solanum* among biomes reveals that tropical forest is the main source of diversity of this genus with 43% of total transitions. This also shows that transitions among tropical biomes are more common than in temperate one (see Figure 3.8 and Table C.3). Arid biomes such as tropical savannas, dry forests and deserts represent the main sinks of *Solanum* diversity from tropical forests.

3.5 Discussion

3.5.1 Early diversification of *Solanum*

The current patterns of diversity and distribution of the megadiverse plant genus *Solanum* appear to have been shaped by a mixture of radiations, range expansions and biogeographic movements. These events have created a great heterogeneity of species richness and diversification rates of *Solanum* lineages across time and space (see Figure 3.2). There are two regions which show distinct patterns in their structure of diversity of *Solanum*; the Neotropics with the highest richness but average rates of diversification, and the Old

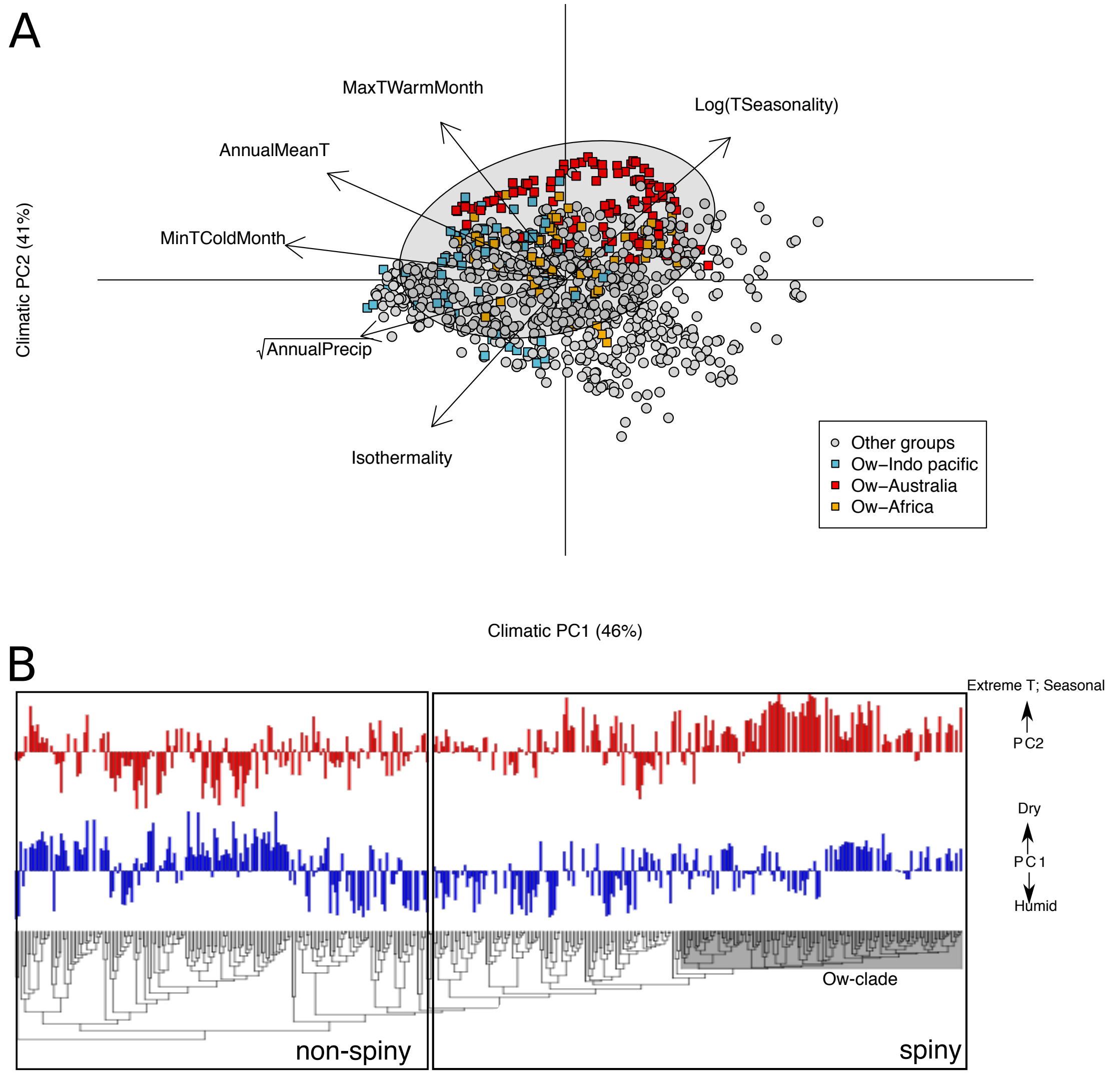


Figure 3.7: Principal component analysis (PCAs) for the climate preferences in *Solanum*. (A) PCA of climate variables of *Solanum* based on median climatic values of 907 species. PC1 explains 46% and PC2 41% of the total variation of the data. Ellipse represents the climate space of the Old World clade. AnnualMeanT, Annual Mean Temperature; AnnualPrecip, Annual Precipitation; MaxTWarmMonth, Max Temperature of Warmest Month; MinTColdMonth, Min Temperature of Coldest Month; TSeasonality, Temperature Seasonality. **(B)** *Solanum* phylogeny (Särkinen *et al.*, 2013) and corresponding species habitats as described by the two principal components eigenaxis. Highlighted clade represents the Old World clade.

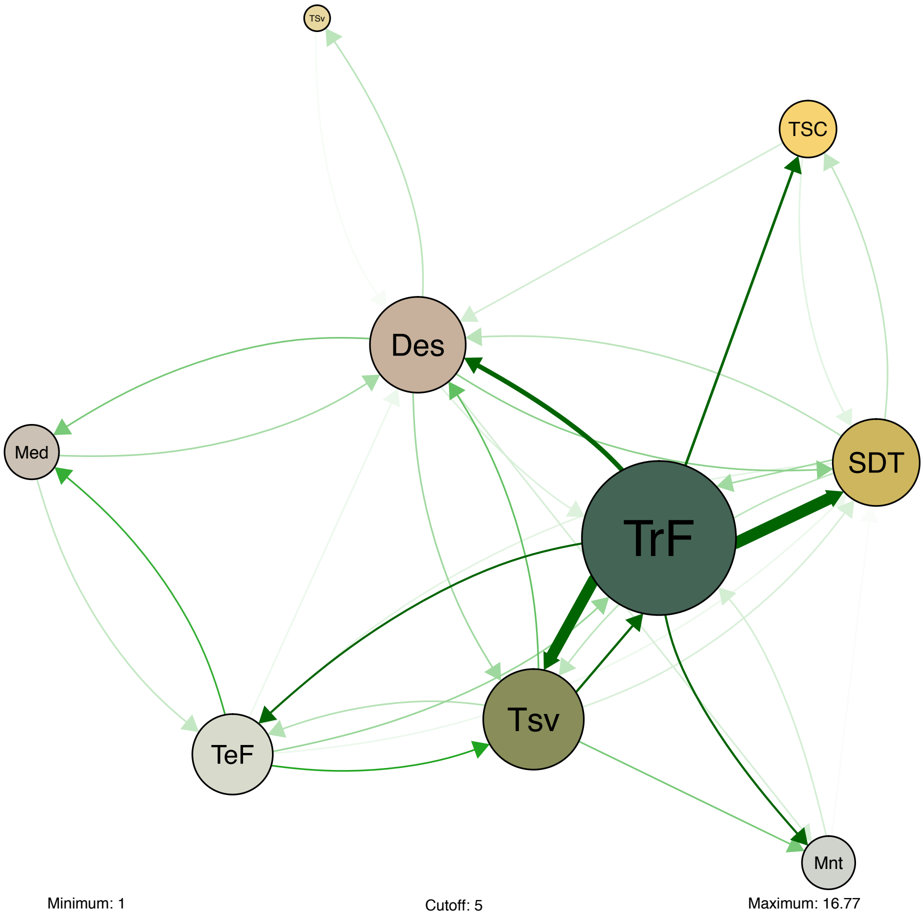


Figure 3.8: Summary of main biome transitions along the phylogeny of *Solanum*. The width of the arrows represent the mean frequency of the movements among biomes. The colour intensity becomes stronger as the frequency of the movement approach the cutoff of 5. The size of the circle represents the relative richness of *Solanum* species in each biome. TrF, Tropical and Subtropical Moist Broadleaf Forests; Tsv, Tropical and Subtropical Grasslands, Savannas and Shrublands; SDTF, Tropical and Subtropical Dry Broadleaf Forests; TSC, Tropical and Subtropical Coniferous Forests; Des, Deserts and Xeric Shrublands; Med, Mediterranean Forests, Woodlands and Scrub; Mnt, Montane Grasslands and Shrublands; TeF, Temperate Broadleaf and Mixed Forests.

World supported by species with the highest rates of diversification, especially in Africa and Australia (see Figure 3.2B). This distinction could be attributed to differences in their age of diversification in association with ecological and biogeographic opportunities.

More than one-third of the historical radiation of *Solanum* appears to have occurred within the Neotropics before the first lineages established in other continents (see Figure 3.5). With most of the richness and endemism of *Solanum* (Figure 3.2A), the Neotropics is the region where *Solanum* species experienced the highest expansion in their ranges, which has also been demonstrated as the highest in the family Solanaceae as a whole (Dupin *et al.*, 2016). The long history of diversification and expansion of *Solanum* within the Neotropics has driven a greater lineage accumulation through time compared with other regions as shown in Figure 3.4. This region not only supported the early evolution of *Solanum* but also represents the main source of diversity and the origin of biogeographic movements across the globe (see Figure 3.6 and 3.8). However, the vast diversity of *Solanum* in the Neotropics is not directly supported by higher diversification rates compared with other regions. In fact, the areas that contained the highest rates of diversification are regions such as Australia, Africa and the Indo-pacific (here refer as the Old World). This mismatch between the high diversity of *Solanum* within the Neotropics and its net diversification rates could be explained by the early and long history of diversification of *Solanum* in this region. Although *Solanum* lineages in the Old world appear to have evolved more rapidly, *Solanum* lineages within the Neotropics had more time and opportunity to accumulate before establishing in other regions, creating this great imbalance of diversity (McPeek & Brown, 2007). Therefore, this pattern of highest richness in the Neotropics could be explained mostly by the greater age of *Solanum* species in this region (Weir & Schluter, 2007).

Due to the long history of diversification of *Solanum* within the Neotropics, we could expect the diversification in this region to decrease with time as a result of an early filling of ecological space by older radiations (i.e., a negative effect of diversity on diversification known as diversity-dependent diversification; See Etienne *et al.* 2011; Morlon 2014; Phillimore & Price 2008; Rabosky & Glor 2010). However, as shown in Figure 3.4, the Neotropics is still experiencing a gradual accumulation of lineages suggesting that diversity-dependent diversification has not limited the accumulation of species in the genus. These results indicate that the Neotropics has provided and still provides important ecological opportunities for the diversification of *Solanum*.

3.5.2 Diversification of non-spiny solanums in the Old World

The non-spiny and the spiny solanums within Africa and Australia have contrasting patterns of present-day diversity; they have evolved as a result of several recent and two early-branching dispersal events from the Neotropics to the Old World (Africa, Australia, and the Indo-Pacific mainly). These differences are reflected in the heterogeneity of net rates of diversification in these regions seen in Figure 3.2 and 3.3. To identify the factors that model these distinct patterns of diversification of *Solanum*, it is essential to understand the historical context of the different dispersal events to the Old World.

The first long-dispersal event from the Neotropics to the Old World occurred ca. 10 Mya (95% HPD 7-12) within the non-spiny solanums, specifically, in the crown of the M1 clade which supports *Solanum valdiviense* + African-non-spiny + Normania + Archaesolanum. The lineages of this group are much less diverse than the spiny solanums in the Old World and are usually restricted to temperate and forested areas. In Africa, for example, the non-spiny solanums are represented by the African-non-spiny clade (ANS). This group is significantly older than the spiny solanums in Africa (derived from the Old World clade) and are mostly woody, although some species can be shrubby in some habitats (e.g., *S. terminale*) and some grow as epiphytes (Knapp & Vorontsova, 2016). These species are distributed in the forested and temperate areas of the Cape region (Austro-temperate region sensu Linder 2014), the central zone of the Tropic-montane flora and Madagascar. In Australia, the diversity of non-spiny solanums is represented by the Archaesolanum group known as the Kangaroo apples. With eight species in total (Poczai *et al.*, 2011), this group occurs in the temperate and forested areas of the South West Pacific (e.g., Australia, Tasmania, New Zealand and Papua New Guinea).

Linder (2014) suggest that the Austro-temperate flora is related to other flora from the southern hemisphere, which may explain not only the relationships between the ANS lineages and the Archaesolanum group but also their similarity in their occupied niches and their distribution (mostly restricted to southern Africa and Australia, respectively). Therefore, a possible explanation for the distribution and the type of habitats that the non-spiny solanums occupy in the Old World might be that the first colonisation to the Old World occurred via the southern hemisphere, at a time when forested regions were more widespread than today, perhaps using Antarctica as stepping stone. One issue with this hypothesis, however, is that the age of the dispersion of non-spiny solanums in the phylogeny is dated to ca. 10 Mya, which occurred long after the initiation of the primary Antarctic glaciation and therefore after the contraction of forested regions in the south (Lawver & Gahagan, 2003). However, recent evidence such as the discovery of a fossil of Solanaceae in the Patagonia (Wilf *et al.*, 2017) and the incompatibility of ages of the current Solanaceae phylogeny with the ages of obligate herbivores (Garzón-Orduña *et al.*, 2015), could suggest that the nodes in the Solanaceae family are significantly older. This new evidence may considerably push back the timing of the evolution of the Solanaceae family and its subclades. For example, using as a reference the change in the age of divergence of Physalinae, which Wilf *et al.* (2017) suggested should move from ca. 11 Mya to 52.2 Mya, the first dispersion of *Solanum* to the Old World could shift from 10 Mya to approximately 33 Mya. This shift in the age of the events could imply that *Solanum* lineages dispersed to the Old World just before the contraction of rainforests and the expansion of sclerophyllous vegetation in the southern hemisphere. This dispersal event would then allow not only a successful colonisation of non-spiny solanums to the Old World but also a subsequent contraction of their ranges due to the climatic changes that occurred soon after their colonisation (Byrne *et al.*, 2011). However, further sampling of lineages of the ANS group and its sister clades is necessary to consider this hypothesis further and to understand the enigmatic distribution of other non-spiny solanums in the Old World such as Normania (which encompasses two endemic species to Macaronesia).

3.5.3 Diversification of spiny solanums in the Old World

The second long-dispersal event from the Neotropics to the Old World occurred ca. 6 Mya (95% HPD 5-7). This event took place at the crown of the Old World spiny clade which includes the majority of the spiny solanums in Africa and Australia. This group contains the species with the highest net diversification rates (see Chapter 2), whose distribution provides the spatial heterogeneity of rates seen in Figure 3.2B and 3.2C. This association between a dispersal event and a shift in diversification could represent a "dispersification" scenario sensu Moore & Donoghue (2007), where the dispersal to a new area promotes a change in diversification either by exploiting newly available resources or through ecological release due to the lack of competitors. However, not all the *Solanum* groups distributed in the Old World and that were associated with dispersal events had shifts in their diversification rates (e.g., the non-spiny solanums in the Old World). This finding and the differences in the timing of the dispersal events between non-spiny and spiny solanums suggests that the diversity of the latter was shaped by different ecological opportunities and evolutionary regimes.

One finding that could lead to understanding these distinct evolutionary regimes of the spiny solanums in the Old World is the comparison of temporal patterns of diversification among major regions in Figure 3.4. This figure shows a distinctive signature of a rapid diversification within Australia ca. 5 Mya, which has been shown in a wide range of other groups that form the Australian biota (Crisp *et al.*, 2004; Crisp & Cook, 2013, 2007; Harrington *et al.*, 2012; Crayn *et al.*, 2006; Blom *et al.*, 2016; Rabosky *et al.*, 2007; Harmon *et al.*, 2003; Rabosky, 2006; Byrne *et al.*, 2008, 2011). The dramatic climate change that Australia has experienced since the breakup of Gondwana not only triggered the radiation of several taxa within Australia (Byrne *et al.*, 2008) but also the extinction of others (Byrne *et al.*, 2011). The initial stage of aridification in the mid-Miocene (ca. 15 Ma), caused the assembly of new biomes, such as the arid, monsoonal and alpine regions, which may have encouraged the establishment of arid-adapted lineages and/or the shift of mesic-adapted lineages to schlerophyll environments (e.g., Tremandraceae, see Crayn *et al.* 2006). The late stage of aridification of Australia in the last 1-4 Ma, which formed the central Australian desert, appears to have triggered the expansion and radiation observed in many Australian groups (Byrne *et al.*, 2008; Crisp *et al.*, 2004; Crisp & Cook, 2013; Rabosky *et al.*, 2007). The majority of the species of the Old World spiny solanums, which have the highest net diversification rates in *Solanum*, occur in the arid and semi-arid regions of Australia, suggesting a potential link between the aridification and radiation of this group. This pattern, and the congruence of the timing of diversification with other arid-adapted Australian groups, indicates that after the Old World spiny solanums arrived by a transoceanic dispersal, they may have established and diversified as the arid environments in Australia expanded.

A general question we can ask about the Old World spiny solanums is "what were the evolutionary processes leading to this diversification?" The widespread distribution of this group in Australia could suggest a positive effect of area on diversification rates. The expansion of a habitat over a large area, such as occurred with the arid zone of Australia, could create the opportunity for geographic isolation and therefore diversification by allopatric speciation (Losos & Schluter, 2000; Davies *et al.*, 2005; Rabosky *et al.*, 2007). However, this may be only part of the explanation for the increasing diversification of the Old World spiny solanums in Australia. Although trait data is needed to explicitly test the hypothesis, the phylogenetic clustering of Old World spiny solanums with critical thermal maximum temperatures (Figure 3.7B) suggests that traits related to thermal physiology might underlie this extraordinary radiation in the arid zones of Australia.

The direction of the dispersal of spiny solanums to the Old World is still unclear. The relationships between Hawaiian endemic species (*S. sandwicense*) with New Caledonia endemic species (*S. dunalianum*), as well as the relationships of Asian spiny with African and Australian solanums (see Aubriot *et al.* 2016) could indicate a potential South Pacific dispersal (Levin *et al.*, 2006; Olmstead *et al.*, 2008). However, the lack of sampling from species distributed in the South Pacific, as well as in some regions of Asia and Africa, limits our ability to consider this hypothesis further.

3.5.4 Biome transitions and climate preferences of *Solanum*

The assembly of the arid biota in Australia involved the establishment of both pre-adaptive lineages and the in situ evolution and habitat shift of rainforest-adaptive groups (Crisp *et al.*, 2004; Crisp & Cook, 2013; Byrne *et al.*, 2008; Crayn *et al.*, 2006). Crisp *et al.* (2009) show that biome shifts in the southern hemisphere are rare, supporting a niche conservatism hypothesis in the assembly of the southern biota. The frequency of the successful establishment of colonisers after transoceanic dispersals is likely to be small, although it should be easier in a biome to which the species are pre-adapted (Donoghue & Edwards, 2014; Crisp *et al.*, 2009). In general, *Solanum* has experienced several biome shifts with frequent transitions from tropical forest to arid environments (e.g., savannas, dry forests, mountains and deserts, see Figure 3.8). Congruent with Crisp *et al.* (2009), older biomes such as the tropical forests serve as the source of diversity for *Solanum*, whereas relatively young biomes such as the savannas, dry forest, and deserts serve as the sinks or recipients. The dispersal of the Old World spiny into Africa and Australia appears to be associated with a biome shift from tropical forests into arid regions. However, it is likely that this transition involved the conservatism of microclimatic preferences since according to field descriptions, spiny solanums in tropical forests are usually found in forest gaps, forest margins and disturbed areas (Bean, 2004; Vorontsova *et al.*, 2013; Knapp & Vorontsova, 2016), which experience higher temperature and aridity than the closed-canopy locations (Senior *et al.*, 2017; Frishkoff *et al.*, 2015). This pre-adaptive condition of the Old World solanums is also supported by the climatic preferences of its sister clades. For example, the clade Elaeagnifolium has seven species native to temperate and highly seasonal regions of North and South America. This clade contains important invasive species such as *Solanum elaeagnifolium*, which tolerates high levels of salinity and drought. The close historical relationship between Elaeagnifolium and the Old World spiny may indicate a pre-adaptation of these groups to the modern arid-zone environments that developed during the Pliocene.

3.5.5 Limitations and future studies

The number of biogeographical events in this analysis should be considered as minimum estimates since the phylogeny used in this study sampled only the 34% of the expected numbers of species of *Solanum* and did not contain any extinct species. Since all the models used in BioGeoBEARS assume cladogenesis, the range-change processes of missing cladogenetic events could underestimate the numbers of events modelled in this biogeographic analysis. The use of polytomy resolvers such as PASTIS (Thomas *et al.*, 2013) to account for the biogeographic information provided by missing taxa it is not appropriate in this context, since this approach can break down the natural patterns of geographic distribution in the phylogeny (Rabosky, 2015b). It is therefore important to increase the sampling of species in the phylogeny in undersampled groups and regions, especially in Africa, Asia and the Indo-Pacific to test the support of these hypotheses.

The difficulty of including the distribution of extinct species into biogeographic analysis may be somewhat limited the results of this study. Not only in this study but also other historical biogeography studies, the ancestral geographic ranges and past biogeographic events are usually inferred using molecular phylogenies of extant species and current distributional data (Lomolino *et al.*, 2010). Extinction events can erase evidence of past speciation events, and its high incidence can affect our inferences significantly (Lieberman, 2005). For example, models such as Dispersal-Extinction-Cladogenesis (DEC), which is the model that best explained the data in this study, assume that accurate ancestral range estimates are obtained only if the rates of extinction and dispersion are low in relation with cladogenesis (Ree & Smith, 2008). Failing to include extinct species, when extinction is particularly high within a clade or area, could lead to spurious reconstructions and artificial distributional patterns (Lieberman, 2005). One possible solution is to include fossil records to account for the distribution of extinct biodiversity and ideally consider the climatic preferences of ancestral lineages such as in Meseguer *et al.* (2014). Incorporating the geographic distribution of extinct species has shown to change dramatically biogeographical inferences (Meseguer *et al.*, 2014; Wood *et al.*, 2012) and ideally different sources such as paleoclimatic data should be considered to infer the biogeographic history of taxa. Unfortunately, *Solanum* does not count with good fossil record to account for the distribution of extinct species in this analysis. In future studies, it might be possible to combine differences sources of evidence such as paleoclimatic data and fossil records to provide more realistic inferences of the past distribution of *Solanum*. However, it is necessary first to encourage the exploration of the fossil record of *Solanum*, or at least in the family Solanaceae, in areas such as the Neotropics, Africa and Australia. Nonetheless, the relatively recent diversification history of this genus might help to minimise the biases associated with potential extinction events.

3.5.6 Concluding remarks

This study shows the great dynamism behind the diversification of genera with high diversity such as *Solanum*, whose current richness appears to be the result of a spatiotemporal variation in diversification and dispersal events. Differences in the diversity between non-spiny and spiny solanums reflect the capacity of the latter to colonise and diversify to arid and high seasonal regions, implying that this group not only colonise the right place at the right moment but perhaps with the right traits. Future work should aim to increase the sampling of species in the phylogeny and quantify the influence of traits in the radiation of this megadiverse plant genus.

Although the expansion of *Solanum* to dry, warm open habitats was natural, human actions are now causing them to expand with unprecedented rapidity: 25% of the world's forest has been permanently cleared in the last 300 years (Hurtt *et al.*, 2011), with often intense effects on the microclimate experienced by species (e.g., Senior et al. in press). The successful expansion of *Solanum* in the Old World could reveal the possible evolutionary capacity of this group to expand and colonise disturbed and open areas. This evolutionary legacy could describe *Solanum* as a winner under current environmental changes, an epithet which is supported by the importance of some *Solanum* species being considered as weeds such as *S. elaeagnifolium*, *S. torvum*, *S. nigrum.* However, the species of *Solanum* that evolved to greater tolerance to high temperatures may be close to their critical thermal limits, which could make them more vulnerable to smaller increases in the temperature of their microhabitat (Stillman, 2003; Frishkoff *et al.*, 2015). Several studies have shown that species' responses to land-use change vary widely, depending on their ecological traits (Newbold *et al.*, 2014; De Palma *et al.*, 2015). Intriguingly, Frishkoff *et al.* (2016) showed that, in a set of over 300 Neotropical forest birds, the species adapted to drier climates are better able to colonise agricultural land. A broader test of the link between climate adaptation and tolerance of human land-use would help to clarify whether this pattern applies more generally. In the next chapter, I, therefore, undertake a global analysis to assess whether the climatic preferences that have shaped macroevolution of *Solanum* are now also shaping plant conservation macroecology worldwide.

CHAPTER 4

Association between plant species' responses to land use change and their climatic adaptations.

4.1 Abstract

Current human activities are rearranging biodiversity patterns which have been modelled over millions of years of evolutionary history. The future of biodiversity will depend on how species are able to respond to these changes. Recent studies have shown that land-use change alters the microclimate of the species mainly by increasing of their local temperature. This study performs the first global-scale assessment of how plant species respond to land-use change and how these responses are associated with their climatic adaptations. Here, I found that plant species' responses to land-use change vary widely depending on their climatic adaptations; species with large ranges which are adapted to arid and seasonal regions appeared to benefit from land-use change, whilst species with narrow ranges which are distributed in mesic and aseasonal environments fared much worse. These responses appear to be associated with the species' ability to adjust to the microclimatic changes produced by land-use change. Thus, species distributed in a stable climate seem to have greater climatic specialisation and, therefore, greater vulnerability to changes in their local climate. The results of this study show that land-use and climate change both appear to favour species adapted to similar microclimatic conditions, indicating not only a homogenisation but also a "thermophilization" of the global biodiversity.

4.2 Introduction

Although the dynamic nature of biodiversity has produced fluctuations of the global distribution of diversity over time, the unprecedented contemporary environmental changes have triggered a global redistribution of life on Earth (Pecl *et al.*, 2017). While climate change is producing rapid and major shifts in species ranges (Chen *et al.*, 2011), land-use change is altering dramatically ecological communities, producing significant changes in their richness and composition (Pereira *et al.*, 2010; Sala *et al.*, 2000; Alkemade *et al.*, 2009; Newbold *et al.*, 2015, 2016a). The fate of future biodiversity will depend on how species respond to these current environmental changes (Jetz *et al.*, 2007), and while many species are negatively impacted by these changes, others appear to thrive in the Anthropocene (Swihart *et al.*, 2003; Newbold *et al.*, 2013; De Palma *et al.*, 2015).

Understanding the combined effects of climate and land-use change is essential to predict future biodiversity impacts and to identify and prioritise taxa and regions at risk (Brook *et al.*, 2008). However, these efforts are complicated by the fact that multiple threats can interact between each other. For example, habitat fragmentation can limit the dispersal ability of species that are trying to track their climatic niches (Honnay *et al.*, 2002; Travis, 2003). Traditionally, studies that have assessed the combined effects of land-use and climate change have focused mainly on how land-use change exacerbates the impacts of climate change (e.g., expansion or contractions in species ranges). However, few studies have evaluated the species responses to changes in microclimatic conditions directly produced under land-use change (but see Frishkoff *et al.* 2016, 2015; Köster *et al.* 2013).

Conversion of natural and semi-natural forests into human-dominated land uses (e.g., pastures, plantations, croplands, urban areas, etc) affects drastically the vegetation structure and plant cover, causing an increase in the local temperature and evapotranspiration rates, which decreases the moisture available for plants (i.e., increasing local aridification). A recent meta-analysis by Senior *et al.* (2017) has shown that increases in local temperature by land-use change can even surpass the temperatures projected by the maximum warming scenario (IPCC, 2013). These drastic changes in microclimatic conditions are likely to favour species which are adapted to landscapes with similar environments to the ones produced by habitat conversion, and negatively impact those which are closely associated with forested areas. For instance, some studies have shown that human-dominated habitats are usually exploited by warm and dry-affiliated species of birds, amphibians and plants (Frishkoff *et al.*, 2015, 2016; Köster *et al.*, 2013; Stevens *et al.*, 2015). Future projections of increasing global temperature, droughts, and rates of land-use change (IPCC, 2013), especially in the tropics, are thus likely to favour these warm-affiliated species producing a long-term homogenisation and “thermophilization” of the biodiversity (Stevens *et al.*, 2015; Frishkoff *et al.*, 2016). However, there are few studies which have evaluated the plant species responses to changes in microclimatic conditions under land-use change.

These are usually restricted to specific organisms (Köster *et al.*, 2013) or particular regions (Stevens *et al.*, 2015; García-Valdés *et al.*, 2015).

Evaluating the correlation between thermal and drought tolerances with responses to land-use change, I aim to identify which plant species are more likely to thrive ("winners") and fail ("losers") under anthropogenic changes. To do so, I use one of the most spatially extensive datasets of the abundances of plant species in different land uses from different surveys across the globe. By characterising the general association between species' responses to land use change and their climatic adaptations, I hope to provide the basis to understand the mechanisms that drive some species to be more or less tolerant to land use change, and in this way, predict and protect the most vulnerable species under current threats.

4.3 Methods

To evaluate the effect of species climatic tolerances on their responses to land use change, I followed three main steps. First, as the response variable, I modelled the species-specific responses to land-use change (ca. 4000 plant species globally distributed), using the species differences in abundance among sites with different treatments of land use. Second, I extract several geographic and climatic variables as a proxy for the species climatic tolerances. Third, I assessed whether any of these proxies best explained the variation in the species responses to land-use change.

4.3.1 Modelling species' responses to land-use change.

In this analysis, I used the site-level biodiversity data compiled in the PREDICTS database from field surveys across the globe (Hudson *et al.*, 2017) that compare diversity sites facing different land-use pressures; many such studies contrast human-dominated sites and nearby primary or secondary vegetation. This database is publicly available from the Natural History Museum's Data Portal (`doi:http://dx.doi.org/10.5519/0073893`) and details about data collation can be found in Hudson *et al.* (2014) and Hudson *et al.* (2017). The site-level biodiversity data used in this analysis included only studies that sampled at least one plant species (see Figure 4.1). Species were included only if their abundance (i.e., number of individuals) was measured at multiple sites and if they were sampled in both land-use classes included in the analysis.

Plant species' responses to land-use change were modelled as the log-response ratio of abundances of each plant species in natural habitat sites (i.e., primary and secondary forests) versus human-dominated sites (i.e., croplands, pastures, plantations and urban areas), hereafter called Hum/Nat, equation 4.1. Within this broad land-use categorisation, differences in vegetation complexity in different habitats can influence species responses according to their climatic tolerances in different ways. The land-use classes, therefore, were reclassified, allowing the comparison of species abundances between sites with complex vegetation cover (i.e, primary forests, mature and intermediate secondary forests, and plantations) and simpler vegetation cover (young secondary vegetation, croplands, pastures, and urban areas), hereafter called Complex/Simple, equation 4.2.

$$ln(RR)_{Hum/Nat} = ln(\frac{\text{mean abundance of } species_i \text{ in human-dominated habitat}}{\text{mean abundance of } species_i \text{ in natural habitat}}) \quad (4.1)$$

$$ln(RR)_{Simp/Comp} = ln(\frac{\text{mean abundance of } species_i \text{ in simpler vegetation cover}}{\text{mean abundance of } species_i \text{ in complex vegetation cover}}) \quad (4.2)$$

For each plant species, abundance was modelled as a function of the land-use classification (natural vs human-dominated or complex vegetation vs simpler vegetation) using Bayesian generalized linear models (bayesglm function, arm package, Gelman *et al.* 2009), with a Poisson error structure or quasi-Poisson when overdispersion was detected (Crawley, 2012), and weakly informative priors. This modelling approach was implemented since the coarse land-use classification in this analysis sometimes produce extreme estimates and standard errors which are unreliable using Generalized linear models (e.g., when a species is always absent from human-dominated sites and still present in Natural sites). Study identity was used as a covariate to account for differences in sampling methodologies and biogeography. The land use coefficient estimates and standard errors for each species were then extracted from the model and considered as the response variable in subsequent analyses.

De Palma (in prep) mainly developed the code for this log-response ratio as a species-specific measure to land-use change. Positive values of the log-response ratio indicate an increase in the abundance of the species (i.e., “winners”) when forested sites are converted to human-dominated ones or when Complex vegetation cover are converted to Simpler vegetation cover, depending on the treatment of land-use change used. The first treatment of land-use change, Hum/Nat, included a dataset of 4,441 plant species from 2,252 sites. Simp/Comp included a dataset of 3,976 plant species from 2,296 sites (differences in the number of species and sites are due to differences in the number of land use comparisons allowed in the data).

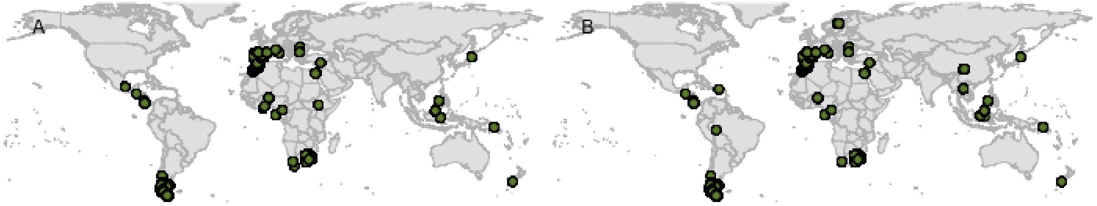


Figure 4.1: Map of sites use in this study where the log-response ratio of plant species abundance was calculated for (A) Hum/Nat and (B) Simp/Comp land-use change treatments.

4.3.2 Climatic data

To assess each species' climatic adaptations, I downloaded the georeferenced records for all the plant species included in the analysis from the Global Biodiversity Information Facility (`www.gbif.org`). I then followed a number of quality control steps, namely: (1) eliminating records with latitude and longitude coordinates of 0 and 0; (2) eliminating records which were found to be outside the land; and (3) eliminating records which contained countries that do not match the coordinates provided after correcting any sign mistakes using the R package "rangeBuilder" (Rabosky *et al.*, 2016). From 58,275,398 records downloaded, a total of 55,091,596 were used for the analysis. Using the R package raster (Hijmans, 2015), I extracted the values of four climate variables from the WorldClim version 2.0 dataset (annual mean temperature, annual mean precipitation, temperature seasonality and max temperature of warmest month Fick & Hijmans, 2016). I chose to work at a medium-scale spatial resolution of 2.5 min, as the optimal trade-off between local accuracy and the tractability of the data due to the worldwide scale of the analysis and the difficulty of verifying the accuracy of the occurrence data on a record-by-record basis. Additionally, I extracted the CGIAR-CSI Global Aridity Index (Trabucco & Zomer, 2009) provided online by the `http://www.cgiar-csi.org/` and the International Center for Tropical Agriculture (CIAT). This index quantifies the proportion of precipitation available after counting for atmospheric water demand (i.e., mean annual precipitation over mean annual potential evapotranspiration). The values of this index increase for more humid conditions (>0.65) and decrease with more arid ones. For each species, I calculated the median values of each of the climate variables as well as their standard deviation for further analysis. Across all the species, the distribution of the aridity index and annual precipitation was slightly skewed to the left. To reduce this skew, the values of these variables were square root transformed. Temperature Seasonality was slightly more skewed than aridity and annual Precipitation, therefore, this variable was transformed using cube root. In addition to the climatic variables, range size was included in the analysis. This variable was calculated for each species as the summed area of cells in which each species occur, using the function "lets.rangesize" from the R package "letsR" (Vilela & Villalobos, 2015a), which accounts for the Earth's curvature. Since the scale of the variables used here can vary greatly, all variables in this analysis were standardised (scale and centre) by subtracting each observation by the mean of the variable among all species and dividing by the standard deviation.

4.3.3 Statistical Analyses

Across all the species, the log-response ratios range from -8 to 7.4, though these extreme values were usually associated with large standard errors (see Figure D.1). Under the Hum/Nat land-use treatment, 1,033 species have log-response ratios significantly different from zero (58% negative and 42% positive), whereas 811 species under the Simp/Comp treatment have significant log-response ratios (49% negative and 51% positive). To assessed whether the variation of the species responses to land-use change is best explained by the species' climatic adaptations, I fit linear mixed effects models (lmer4 R package Bates *et al.*, 2015), weighted by the inverse square of the standard errors of the species log-response ratio. Genus nested within Families was considered as random effects (both random intercepts and slopes), to account for differences in the responses among taxonomic groups. In this analysis, I first determined the best random-effects structures using likelihood ratio tests (Zuur *et al.*, 2009), comparing all formulations. Then, I fitted an initial maximal fixed-effects model structure with all two-way interactions among climatic variables and range size, and used stepwise model simplification on likelihood ratio tests to reduce model complexity. Full models were assessed for multicollinearity using generalized variance inflation factors (GVIFs Zuur *et al.*, 2009). To estimate the relative importance of the variables in the minimum adequate models, I calculate the decrease in explanatory power when each term was excluded from the model, estimating the marginal R^2GLMM, conditional R^2GLMM, AIC weight values and the variance within the random effects. The marginal and conditional R^2GLMM values (i.e., the variance explained by the fixed effects alone and by the fixed and random effects combined, respectively) were calculated using the "MuMIn" R package (Bartoń, 2013).

To avoid multicollinearity among the variables, I fitted different models using only one precipitation variable (aridity index or mean annual precipitation), one temperature variables (mean annual temperature, temperature seasonality or max temperature of the warmest month) and range size. This gives a total of six different sets of models for each treatment of land-use change (see Table 4.1). I followed this approach since correlated variables can still represent different ecological responses to land-use change. For example, although annual mean temperature and max temperature of the warmest month are highly correlated, the latter can be more informative in the context of land-use change as it represents a limiting environmental factor. The models with the highest explanatory power for each land-use change treatment were then selected by their marginal and conditional R^2. The importance of each of the variables, for the models with the highest explanatory power, was assessed using the sum of Akaike weights for all the combination of models where the term appears.

4.4 Results

The variation of the log-response ratio of plant species abundances in natural habitats and nearby human dominated areas (i.e., ln(RR) Hum/Nat) was better explained by the fixed terms than the variation log-response ratio of plant species abundances in complex vegetation cover habitats and nearby simpler vegetation cover ones (i.e., ln(RR) Simp/Comp) expressed by their marginal R^2 (see Table 4.1).

In both land-use change treatments (Hum/Nat and Simp/Comp), models using aridity index returned equivalent qualitative results to those using mean annual precipitation. I therefore focus on the models which use mean annual precipitation in the main text (see Table 4.1).

LU Treatment	**Precipitation**	**Temperature**	**Minimum adequate model**	**R^2 marg.**	**R^2 cond.**
Hum/Nat	Aridity index	MAT	ln(RR) ∼ Aridity + Range size + Temp + Range size:Temp	0.039	0.386
		Maximum	ln(RR) ∼ Aridity + Tmax + Range size + Range size:Tmax + Aridity:Tmax	0.032	0.377
		Seasonality	lln(RR) ∼ Seasonality + Range size + Range size: Seasonality	0.042	0.369
	MAP	MAT	ln(RR) ∼ Precip + Temp + Range size + Range size:Temp	0.037	0.373
		Maximum	ln(RR) ∼ Precip + Tmax + Range size + Precip:Tmax	0.030	0.366
		Seasonality	ln(RR) ∼ Precip + Seasonality + Range size + Range size:precipitation	0.046	0.367
Simp/Comp	Aridity index	MAT	ln(RR) ∼ Temp + Range size + Range size:Temp	0.018	0.322
		Maximum	ln(RR) ∼ Aridity + Tmax + Range size + Range size:Tmax	0.022	0.328
		Seasonality	ln(RR) ∼ Seasonality + Range size + Range size:Seasonality	0.018	0.323
	MAP	MAT	ln(RR) ∼ Precip + Temp + Range size + Precip:Temp	0.022	0.326
		Maximum	ln(RR) ∼ Precip + Tmax + Range size + Precip:Tmax	0.023	0.321
		Seasonality	ln(RR) ∼ Seasonality + Range size + Range size:Seasonality	0.017	0.319

Table 4.1: Minimum adequate models from each of the set models used in this study. MAT, Mean Annual Temperature; MAP, Mean Annual Precipitation; Tmax, maximum temperature of the warmest month.

4.4.1 Natural vs Human-dominated habitats

Among the different temperature variables, temperature seasonality explained more the variation of the log-response ratio of the species abundance in this land use treatment, with a reduction of 29% of the variation compared with a reduction of 25% and 6% for mean annual temperature and maximum temperature of the warmest month, respectively.

Response variable	**Explanatory variables**	**Estimate**	**SE**	χ^2	**d.f.**	**Sig.**
$ln(RR)_{Hum/Nat}$ (N=4441)	(Intercept)	-0.313	0.046	46.952	1	***
	Precipitation	0.020	0.043	0.220	1	
	Range size	0.218	0.035	39.219	1	***
	Seasonality	0.224	0.057	15.368	1	***
	Range size:Precipitation	0.126	0.030	17.489	1	***
$ln(RR)_{Simp/Comp}$ (N=3974)	(Intercept)	-0.314	0.051	38.190	1	***
	Precipitation	-0.156	0.038	17.225	1	***
	Range size	0.204	0.030	47.300	1	***
	Tmax	0.168	0.040	18.101	1	***
	Precipitation:Tmax	0.213	0.042	25.672	1	***

Table 4.2: Anova table for the minimum adequate models for both the Hum/Nat and Simp/Comp land-use treatments. Stars indicate the level of significance (Sig.): <0.05*, <0.01** and <0.001***. Random effects: NatHum, (1+Range size+Seasonality+Precipitation|Family)+(Range size|Genus); ComSimp, (1+Range size+Tmax+Precipitation+Precipitation:Tmax|Family)+(Range size|Genus).

The best supported model in the Hum/Nat treatment included temperature seasonality and the interaction between range size and mean annual precipitation (full model vs null model, df=4, χ^2=82.16, P<0.001, Table 4.2). The best random effects in this model included random intercepts and slopes for range size among genera and random intercepts and slopes for range size, seasonality and precipitation among families. Under this model, seasonal species had higher log-response ratios than aseasonal species (estimate=0.22, bCI=0.11-0.34, P<0.001, see Figure 4.2). How mesic plant species responded to habitat conversion depended strongly on their range size: small-range species fared much worse (i.e., had much lower Hum/Nat log-response ratios) than large-range species (see Figure 4.2 and Figure D.3).

4.4.2 Simple vs Complex vegetation cover habitats

The temperature variable that explained most of the variation of the log-response ratio of species abundance under drastic changes in vegetation cover was maximum temperature of the warmest month (Tmax). This variable explained 31% of the variation compared with 20% and 10% for mean annual temperature and temperature seasonality, respectively.

The best supported model under this land-use change treatment included the fixed terms interactions between Tmax and mean annual precipitation with the additive effect of range size (full model vs null model, df=4, χ^2=70.53, P<0.001, Table 4.2). The best random effects in this model also included random intercepts and slopes for range size among Genera and random intercepts and slopes for the same terms of the fixed effects

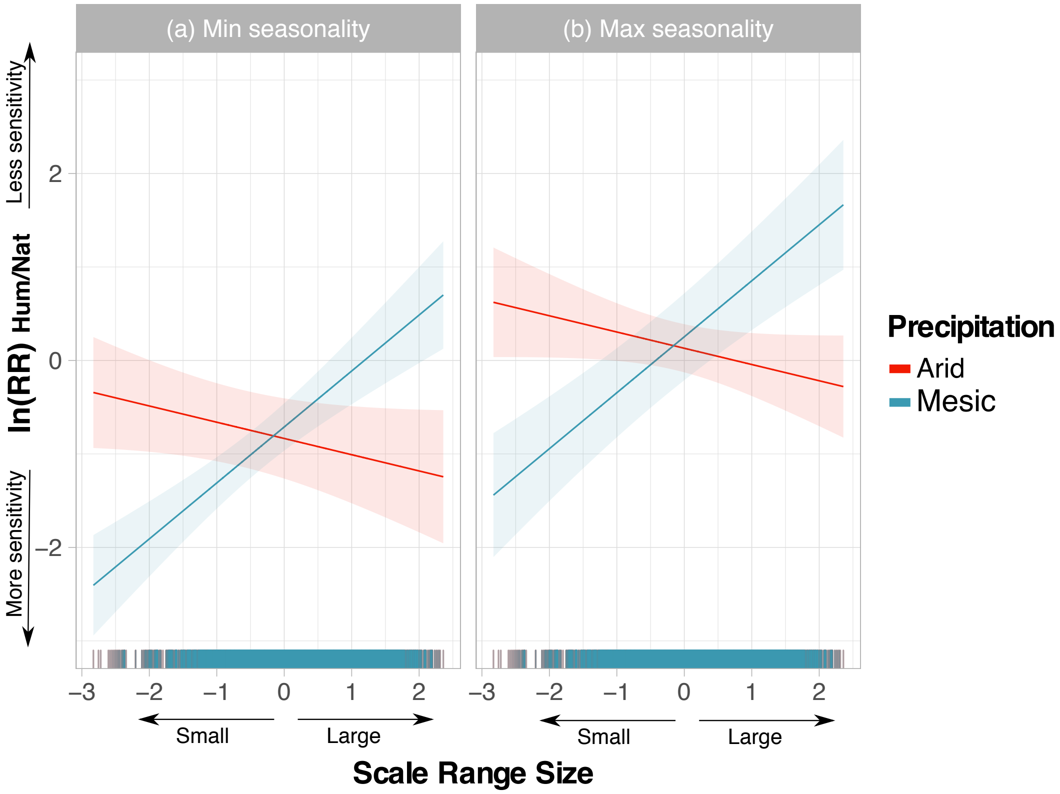


Figure 4.2: Log-response ratios of species abundances in a natural habitat and a nearby human dominated areas, ln(RR) Hum/Nat, along the interaction between range size and mean annual precipitation for different values of seasonality. Error bars represent 95 % confidence intervals. In precipitation, Arid $\leq$ 1358 mm, Mesic $>$ 1358 mm. 1358 mm correspond to the mean value of annual precipitation of all species in the Hum/Nat treatment.

among Families. Under this model, large-range species had higher log-response ratios than small-range species (estimate=0.20, bCI=0.15 - 0.26, P<0.001, see Figure 4.3). How mesic or arid species responded to habitat conversion depended strongly on their heat tolerance (i.e., Tmax values): mesic species with lower heat tolerance fared much worse (i.e., had much lower Simp/Comp log-response ratios) than species with higher heat tolerance; whereas, dry species with lower heat tolerance had much higher log-response ratios than species with higher heat tolerance (see Figure 4.3 and Figure D.4).

In the minimum adequate models from both land-use change treatments, range size was the most important predictor, and temperature was the second most important predictor of the variation of log-response ratio of species abundance. Precipitation and interaction terms were better predictors in the Simp/Comp treatment than in the Hum/Nat one (see Figure 4.4).

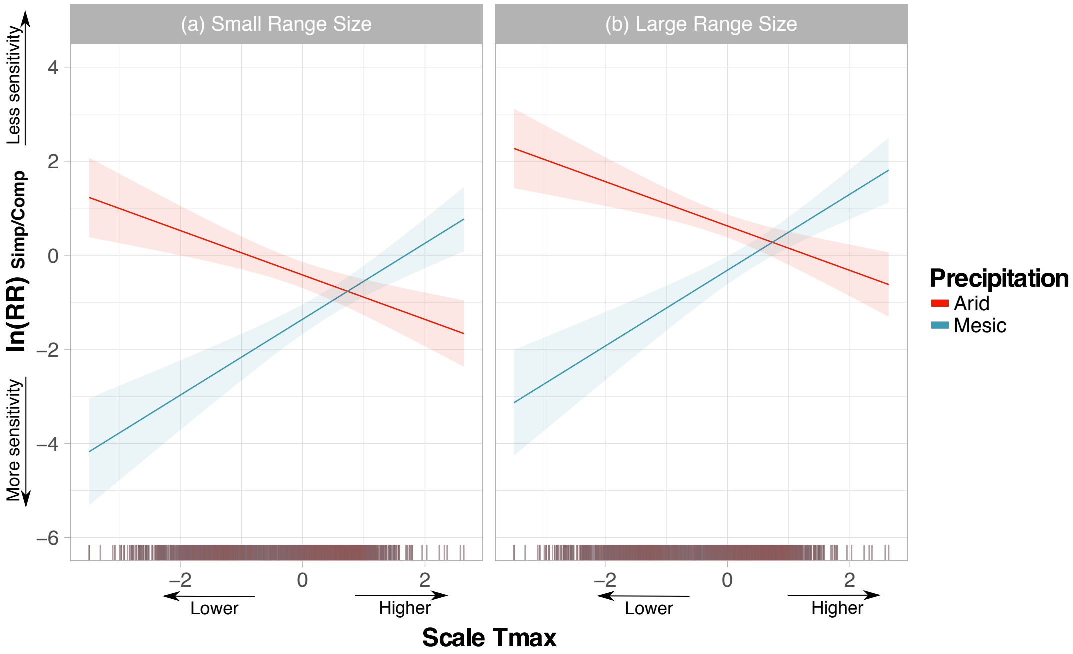


Figure 4.3: Log-response ratio of species abundances in sites with complex vegetation cover and nearby sites with simpler vegetation cover, ln(RR) Simp/Comp, along mean annual precipitation and maximum temperature for different values of range size. Error bars represent 95 % confidence intervals. In precipitation, Arid $\leq$ 1333 mm, Mesic $>$ 1333 mm. 1333 mm correspond to the mean value of annual precipitation of all species in the Simp/Comp treatment.

In both land-use change treatments, the random effects structure explained a significant portion of the variation (Hum/Nat, marginal R^2=0.05 and conditional R^2=0.37; Simp/Comp, marginal R^2= 0.02 and conditional R^2=0.33). The importance of the random effects on both models and the fact that model residuals did not vary significantly among families in both treatments, Nat/Hum ($F_{284,4156}$=0.966, p=0.65), and Simp/Comp ($F_{276,3697}$=0.0.858, p=0.95), indicate that the random structure did a reasonable job of removing phylogenetic non-independence.

4.5 Discussion

4.5.1 Range size and land-use change responses

The spatial distribution or extent of species (i.e., range size) has been widely recognised as a key feature to measure species vulnerability to changing environments, as those with smaller ranges tend to respond more negatively to human impacts than those with larger ranges (Purvis *et al.*, 2000; Whitton *et al.*, 2012; Newbold *et al.*, in Review; Gaston, 1996). In this study, range size was the most important predictor of variation in the log-response

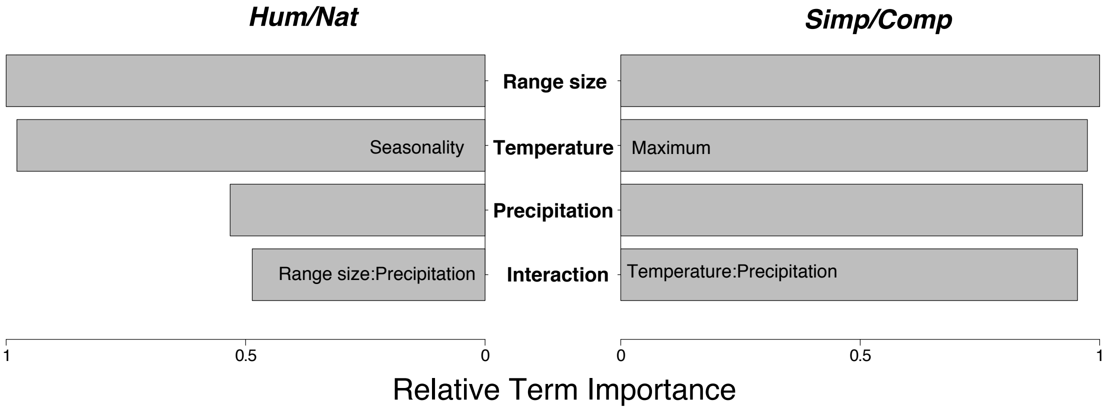


Figure 4.4: Relative importance of variables in explaining the variation of log-response ratios in the minimum adequate models for both treatments of land-use change. The relative importance was calculated as the sum of Akaike weights for models where the term appears.

ratios of species abundances for both land-use change treatments (see Figure 4.4). In both cases, the log-response ratios of small-range species were significantly lower than for widespread species, implying small range-size species generally respond more negatively to land-use change. These results are consistent with other studies such as Lozada *et al.* (2008), Köster *et al.* (2013) and Newbold *et al.* (in Review), which also found a strong correlation between range size and vulnerability to human-induced habitat change.

Climate tolerance and dispersal limits are known to be the most common driving factors of range size (Brown, 1984; Morin & Lechowicz, 2013; Nogués-Bravo *et al.*, 2014). Range-restricted species are often particularly vulnerable to extinction because of their narrow niche breadth or their ecological specialisation (Slatyer *et al.*, 2013). These species usually have lower plasticity in responding to changes in microclimatic habitat conditions (Köster *et al.*, 2013), which makes them very vulnerable to the changes in vegetation cover associated with land-use change, as is shown in this study. In contrast, widespread species are often considered ecological generalists because they usually have wider climatic tolerances (Slatyer *et al.*, 2013); additionally, they are associated with higher dispersal abilities owing to their intrinsic evolutionary legacies or as a result of human activity (Arribas *et al.*, 2012; Van Kleunen *et al.*, 2015). Therefore widespread species may be even more successful under land-use change, not only due to their tolerance to changes in microclimatic conditions but also to their tolerance to other consequences of land-use change such as habitat fragmentation.

Due to the great importance of range size on classifying species under the IUCN Red List of Threatened Species, it is not surprising that most of the species in this study that are listed as threatened (i.e., Critically Endangered, Endangered or Vulnerable) on the IUCN

Red List are associated with significantly negative log-response ratios (see Figure D.1 and Figure D.2). This presents the possibility that the log-response ratios methodology in this study could be used to complement the information of data-deficient species on the IUCN Red List.

4.5.2 Plant species' climatic adaptation and sensitivity to land-use change (Natural vs human dominated habitats)

In this study, nearly all aseasonal species experienced significant reductions in abundances when compared to seasonal ones. One possible explanation for this could be that species living in seasonal environments already experience a wider range of environmental conditions which allow them to fare better when land-use change causes large microclimatic changes (Williams *et al.*, 2007). In addition, many seasonal species in this analysis appear to be associated with arid environments as shown by the large confidence intervals of log-response ratios in seasonal, mesic species in Figure 4.2 (although the correlation was not so strong as to influence results, as assessed using Generalized Variance Inflation Factors, see Table D.3). Arid, seasonal species are often affiliated with simpler vegetation habitats such as savannas and shrublands, therefore the significant increases in abundance of these species in human-dominated habitats are likely to be explained by similarities in microclimatic conditions between their native distribution and the conditions that land-use change creates. These results are consistent with those from Frishkoff *et al.* (2016) who found that bird species affiliated to drier climatic conditions were more able to colonise and persist in agricultural habitats.

The lower log-response ratios of abundance in aseasonal, mesic affiliated species could be attributed to changes in vegetation cover structure. Precipitation is considered the best predictor of plant height, with a reduction of precipitation usually translating to losses of biomass (Moles *et al.*, 2009; Liu *et al.*, 2013). In addition, changes in precipitation have been shown to be a limiting factor on the distribution of tropical species (Esquivel-Muelbert *et al.*, 2017; Baltzer *et al.*, 2008; Brenes-Arguedas *et al.*, 2009), which are very likely to be represented by the aseasonal, mesic affiliated species in this analysis.

Most of the species listed as threatened on the IUCN Red List that experience significant changes in their abundances in the Hum/Nat land-use treatment are associated with mesic, aseasonal environments (see Table D.1). In contrast, widely cultivated species, effective colonisers and important weed species such as *Tribulus terrestris* (Zygophyllaceae) and *Aster squamatus* (Asteraceae) are among the species found to have significant increases in abundances and are associated with arid, seasonal adaptations (see Table D.2).

4.5.3 Plant species' climatic adaptation and sensitivity to changes in vegetation complexity

Contrary to expectations, effect sizes in species' sensitivity to land-use change were no higher when only vegetation cover complexity was taken into account (i.e., the Simp/-Comp treatment). The reduction in the explanatory power in the Simp/Comp treatment and the more balanced proportion of "losers" and "winners" in this analysis may suggest that the additional complexity given by land-uses such as plantations is not sufficient to buffer the microclimatic changes created by land-use change. However, these results should be interpreted with caution as the sampling methods commonly used in the studies underlying this analysis (and most ecological studies) are usually designed to sample habitats with simpler vegetation cover, which could lead to an underestimation of the real differences in vegetation complexity among land-use habitats.

The importance of the climatic variables in the land-use treatment when complexity of vegetation cover was taken into account (i.e., Simp/Comp) was higher than in the analysis of responses to changes from natural to human dominated habitats (i.e., Hum/Nat), suggesting that the sensitivity of species to drastic changes in vegetation cover may depend strongly on how those species respond to changes in their microclimatic adaptations. When only changes in vegetation cover are taken into account, the responses of mesic and arid species to habitat conversion depended strongly on their heat tolerance (i.e., Tmax). Mesic, small-range species with lower heat tolerance were found to have lower log-response ratios (i.e., they were the most sensitive to changes in vegetation cover). Among these very sensitive species are found shade species such as *Homalomena pendula* (Araceae) and *Begonia convallariodora* (Begoniaceae), or epiphytes species such as *Philodendron verrucosum* (Araceae).

In general, arid affiliated species show significantly higher log-response ratios than mesic species, implying arid species tend to be less sensitive to land-use change. However, arid species with higher heat tolerances (i.e., higher values of Tmax) were found to fare worse than mesic species with higher heat tolerance. This could be because arid affiliated species are usually associated with habitats with excessive light and strong climatic fluctuations. These strong climatic variations may produce the filtering of species based on their thermal limits, meaning that species in these environments may be close to their critical thermal limits, making them very vulnerable to small changes in temperature. These results are consistent with other studies which have found that species that have evolved greater tolerance to high temperatures have done it at the expense of acclimation, and therefore they are the most vulnerable to smaller changes in microhabitat temperatures (Stillman, 2003; Frishkoff *et al.*, 2015).

4.5.4 Limitations

Based on the marginal R^2, little of the variation of the log-response ratios in either land-use treatment was explained by the fixed effects of the models suggesting that additional factors need to be taken into account such as ecological traits (e.g., growth rate, phenotypic plasticity, leaf area, self-fertilization, among others) or dispersal abilities (Nogués-Bravo *et al.*, 2014; Gallagher *et al.*, 2015; Daehler, 2003). In addition, the importance of random slopes in both models of land-use change (rather than random intercept) suggests that standard phylogenetic comparative analyses – which assume a constant relationship between y and x – might not be fully appropriate, but that the relationship itself varies across the phylogeny. This variation might be due to differences in species traits. With the increasing availability of data and methods that allow filling gaps on trait data such as Schrodt *et al.* (2015), there is a great scope for future studies to assess the influence of species traits in the responses to land-use change.

The results of this study rely on correlative models to assess links between species responses to land use change and their climatic adaptations using space-for-time substitution. While these represent powerful and data-rich approaches to elucidate trends in response, to truly understand the direct or indirect mechanisms, it is necessary to perform long-term studies of plant responses at different spatial and temporal scales. Furthermore, log-response ratio of abundance might not be a perfect measure for plants because, particularly for large plants, more young plants than mature plants can fit into a plot. As a result, disturbance can increase abundance despite reducing biomass. Because the data sources rarely if ever report traits measured on the plants in the plots (e.g., height), I was unable to convert abundances into meaningful estimates of biomass.

Climatic adaptations based on regional climate average across species' current geographic range does not account directly for microclimatic conditions created under forest cover or topography (Harwood *et al.*, 2014). Biogeographic affinities can be then used as an alternative to describe microclimatic tolerances based on the climatic conditions of the regions and time periods under which species have evolved and diversified (Wiens & Donoghue, 2004; Cavender-Bares *et al.*, 2016). Due to the global scale of this analysis and computational limitations, it was not possible to evaluate the biogeographic affinities of each of the species used in this study. However, with the increasing availability of plant phylogenies and computational efficiency biogeographic affinities would be able to be implemented in the near future. Moreover, although there are large amounts data in this analysis with a global scope, there are still geographical biases, especially in the Americas. A further collection of studies could greatly enhance the predictive ability of the models of this analysis.

These analyses have focused on land-use — and associated microclimatic changes — effects on individual species, but understanding effects on diversity within ecological assemblages is also important, given the role of local diversity in determining ecosystem function (Isbell *et al.*, 2011). The low R^2 values I have found here caution against estimating assemblage-level consequences from these species-based models. In the next chapter, I therefore take an assemblage-level approach to estimate how biodiversity has been affected by land-use change.

4.5.5 Concluding remarks

Human activities are currently driving significant rearrangements in the composition of almost every ecosystems on the globe. This study shows that land use change is selecting plant species with broad ranges and usually those adapted to arid and seasonal environments while adversely affecting plant species with narrow ranges and greater climatic specialisation. Since land use and climate change appear to favour similar species, these activities are likely to homogenise biodiversity faster than expected. However, more studies are necessary to understand not only the effects of climate and land-use change on the ecosystem compositions but also their synergistic effects.

CHAPTER 5

Modelling and projecting the response of local assemblage composition to land-use change across Colombia

5.1 Abstract

Understanding the impact of land use change within assemblages is fundamental to mitigation policies at local and regional scale. Here, we aim to quantify how site-level terrestrial assemblages are responding to land use change in Colombia a mega-diverse country and to project future biodiversity under different scenarios of land use change associated with climate change policies. We collated original biodiversity data from 17 publications (285 sites) that examined how human impact affects terrestrial biodiversity in Colombia. From each site we estimated compositional intactness (i.e. compositional similarity to undisturbed sites). We fitted generalized linear mixed-effects models to estimate how these measures of local biodiversity vary across land use habitats. Using space-for-time substitution, we applied our estimates to hindcast biodiversity changes since 1500 and project future changes under climate change policies of the four representative concentration pathways (RCPs). Assemblages in urban, cropland and pasture sites were compositionally very different from those in primary vegetation. We infer that average compositional intactness has been reduced by 18% across Colombia to date, with strong regional variation. The best RCP scenario for future biodiversity is GCAM-RCP4.5, a path that favours the expansion of secondary forests under a strong carbon market; while the worst is MESSAGE-RCP8.5, "the business-as-usual" scenario. Land use change has driven an increasing change in the composition of ecological assemblages in Colombia. By 2095, the implementation of carbon markets policy of climate change from GCAM-RCP4.5 could mitigate these changes in community composition. In contrast,

the business-as-usual scenario MESSAGE-RCP8.5 predicts a steep community change placing the quality of ecosystems at risk.

5.2 Introduction

The socio-economic benefits of exploiting natural resources have resulted in pressures that pose a serious threat to ecosystems (Chapin III *et al.*, 2000; Foley *et al.*, 2005; Green *et al.*, 2005; Millennium Ecosystem Assessment, 2005; Sala *et al.*, 2000). Landscape transformation, disruption of biogeochemical cycles, invasive species and climate change are among the main consequences of a human-modified world (Chapin III *et al.*, 2000; Foley *et al.*, 2005). Land-use activities are currently the most important global driver of change in terrestrial biodiversity (Alkemade *et al.*, 2009; Jetz *et al.*, 2007; Pereira *et al.*, 2010; Sala *et al.*, 2000; van Vuuren *et al.*, 2006). Pressures that alter the composition and diversity of species assemblages are likely to also alter many ecological processes and services (Cardinale *et al.*, 2012; Díaz *et al.*, 2013; Hooper *et al.*, 2012; Srivastava & Vellend, 2005). Given the ongoing growth of human populations and the rising demand for food and fresh water, it is important to understand how assemblage-level (i.e., local) community composition and richness are affected by human activities and how they would be affected under different future socio-economic assumptions. Understanding biotic effects of human impacts is a particularly urgent priority in tropical forests, because of their combination of high biodiversity and rapid land-use change (Dirzo & Raven, 2003; Hansen *et al.*, 2013; Sala *et al.*, 2000).

Colombia, in the northwestern part of South America, is among the most biodiverse and vulnerable countries on Earth (Brooks *et al.*, 2002; Mittermeier *et al.*, 2003; Myers *et al.*, 2000; Olson & Dinerstein, 2002). The variation in altitude, geological substrates and rainfall among its regions make Colombia a land of climatic and ecological contrasts: such distinct ecosystems as deserts, rainforest and paramos can all be found within a small area. As a result of the biophysical, socioeconomic and political variety across its main five regions Andean, Caribbean, Pacific, Orinoco and Amazonian Colombia shows regional patterns of landscape transformation and human impact (Etter *et al.*, 2011; Etter & van Wyngaarden, 2000). Colombia's environmental and socioeconomic gradients make it a valuable case study on how biodiversity responds to human impact in the tropics.

The exploitation of natural resources is an important asset in Colombia's economy more than 6% of GDP in 2015 came from agriculture: forestry, fishing, cultivation of crops and livestock production (Dinero, 2015; The World Bank, 2015) so unplanned land-use change may be a major cause of biodiversity decline in the country. Previous studies have identified the expansion of cattle industry, agriculture and deforestation (mainly due to illicit crop production) as main reasons for land-use change in Colombia, with additional influences of soil fertility, proximity to roads, rural population density and forced migration (Armenteras *et al.*, 2011; Etter *et al.*, 2006a,b). Quantifying the average impacts of land-use change on local ecological assemblages is important because conservation and land-management decisions are typically implemented at national or lower levels; such understanding can also be combined with approaches such as the planetary boundaries framework (Steffen *et al.*, 2015) to mitigate societal impacts of land-use change.

Several studies have quantified the impact of land-use change on diversity, but they have usually (though not always: Gilroy *et al.* 2015, 2014) been restricted geographically and taxonomically, which is problematic given that responses can vary among regions and taxa (Gibson *et al.*, 2011; Newbold *et al.*, 2014). Additionally, studies so far have focused mostly on species-richness which can mask large changes in community composition if losses of species are balanced by colonisation of new species or spread of disturbed-tolerant native species (McCune & Vellend, 2013; Thomas, 2013). We have therefore collated data from many Colombian field studies that have assessed land-use effects on biodiversity, and which between them cover a wide range of taxonomic groups and represent four of Colombia's five main regions. Assuming spatial comparisons can be used in lieu of time-series data, we model responses of community composition to land-use change and two related pressures (human population density and proximity to roads).

Because Colombia's land-use history is relatively well known (Etter *et al.*, 2008; Etter & van Wyngaarden, 2000), we are able to couple our models with historical land-use maps to explore regional variation in the biotic impact of human activities to date. Likewise, we use scenario-based estimates of future land use in Colombia (Hurtt *et al.*, 2011) to project changes in biodiversity through this century under a range of possible climate policies (Jantz *et al.*, 2015; Newbold *et al.*, 2015).

5.3 Methods

5.3.1 Diversity data

Search criteria

We collated field data from reports and scientific literature in English and Spanish of Colombian studies that aimed to quantify how human pressures affect site-level biodiversity. Nearly all such studies have used space-for-time substitution, comparing biodiversity at otherwise-matched sites facing different land use or use intensity under the assumption that differences in community composition are caused by difference in pressure. Although this approach has limitations (Luck, 2007; Wearn *et al.*, 2012), it is the only approach currently feasible to obtain a large evidence base to support dose-response models of land-use change (Alkemade *et al.*, 2009; Newbold *et al.*, 2015, 2014). An initial pool of 120 promising papers was identified in January 2013 from a Web of Knowledge search that used the following terms:

((habitat OR forest* OR land*) AND (impact OR effect OR influence OR role) AND (species abundance OR species diversity OR biodiversity) AND (natural* OR semi-natural* OR primary OR manag* OR unmanag* OR virgin OR old-growth OR remnant* OR ancient* OR silviculture OR cut* OR clear-cut* OR felling OR clear-fell* OR clearfell* OR select* cut* OR thinning* OR coppic* OR logging OR unlogging OR logged OR unlogged OR regeneration OR plantation* OR planting OR drainage OR ditching OR intensification OR old OR abandonment OR drug OR unrest) AND (Colombia* OR equator* OR neotrop*))*

As well as examining these 120 papers, we also searched Colombian journals and university library catalogues for relevant papers and dissertations (see Appendix E in Supporting Information). We retained all studies that:

1. Sampled, in a comparable way, multiple terrestrial sites or landscapes having different land-use or intensity impact; and
2. Considered the impact of a human pressure on a set of taxa; and
3. Reported any metric of species' occurrence or abundance, or of site-level diversity (e.g., species richness or diversity indices).

Subsequently, S. E.-L. contacted the corresponding authors (73 in total) from 61 retained sources (papers and dissertations) to request the following information, if it was not presented:

1. Locations of the sampling points, as precisely as possible;
2. An indication of the predominant land use at each site (e.g., primary forest, secondary forest, intensely-farmed crop, hedgerow between two fields);
3. Data on occurrence or ideally the abundance of each species at each site (many papers included only summaries).

We obtained the raw biodiversity data and matching land-use information relating to 17 sources, 26 studies and 285 within-study sites (Figure 5.1 and Table E.1). Some sources provided multiple sets of data, collected using different sampling methods. We treated each such data set as a different study, because diversity cannot be compared directly between samples collected in different ways, while recognising this risks a degree of pseudoreplication. The available data are strongly biased geographically (Figure 5.1): 73% of studies are from the Andean region, while one of the five regions (the Pacific region) is not represented. Taxonomic bias is also present but less pronounced: arthropods are the main taxon sampled (71% of studies), followed by vertebrates (18%), plants (7%) and fungi (4%). A total of 2582 species were sampled.

Figure 5.1: Map of Colombia displaying the distribution of sources collected in this study. The main polygons correspond to the main regions in Colombia and blue dots represent the 285 within-study sites.

Site-level variables

Each site was allocated to one of eight land-use classes (primary vegetation, mature secondary vegetation, intermediate secondary vegetation, young secondary vegetation, plantation forest, croplands, pastures, and urban) by applying the criteria in Table E.1 in Hudson *et al.* (2014) to descriptions in the source papers or directly from the authors. These decisions were all made by one person (S. E.-L., a native Spanish speaker) after training and with discussion where necessary, to maximise consistency in the use of categories (intensity classes were assigned to all sites, but the design was too unbalanced and sample sizes too small to permit use in modelling, see Table E.2). Only primary forests were reported in our studies as primary vegetation type (i.e., no other primary vegetation types such as natural grasslands were used in our studies). Each site was also characterised by its human population density and distance from the nearest road, two pressures shown to predict site-level diversity in a global analysis (Newbold *et al.*, 2015); these variables were logarithmically transformed $\log(x + 1)$ prior to analysis to reduce skew.

We use compositional intactness as a proxy of site-level biodiversity, which has shown to respond more sensitively to disturbance than other species diversity metrics (Dornelas *et al.*, 2014; Magurran & Henderson, 2010). Compositional intactness is a more sensitive indicator of biodiversity change than species richness, because it captures gains and losses of species, rather than only the difference between them (Dornelas *et al.*, 2014; Thomas, 2013). We estimated compositional intactness as the mean Sørensen similarity index (Magurran, 2004) between primary vegetation sites and sites within each other land use within each study (or within each block, where present), scaled by the mean Sørensen index between primary vegetation sites within the same study (Newbold *et al.*, 2015). This scaling avoids conflating natural spatial turnover with human impact, but means values can exceed 1. Compositional intactness was estimated for studies that include primary vegetation sites as reference (15 out of 26). The independence of explanatory variables was assessed using Generalized Variance Inflation factors (GVIF, Zuur *et al.*, 2009), which never breached the threshold of 10 (see Table E.3).

5.3.2 Statistical analysis

Estimating biodiversity response to human pressures

The studies use in this analysis sampled biodiversity in different ways: sampling effort, sampling method, area sampled, temporal duration of sampling and taxonomic focus all differed among studies. Because such differences will affect site-level diversity, we used generalized linear mixed-effects models (Bolker *et al.*, 2009), to control for among-study (and, for studies with blocked designs, among-block) differences. All the models were fitted using the lmer function from the LME4 package version 1.1-6 (Douglas Bates & Bolker, 2013) in the software environment R 3.0.0 (R Core Team, 2013). Given the small sample size and unbalanced distribution of stages in secondary vegetation, mature and intermediate secondary vegetation were merged as "I-MSV" (see Figure 5.2). Because our response variable can exceed 1, we did not apply transformations prior to analysis; model diagnostics suggested our treatment was reasonable. The optimal structure of random effects was first found using stepwise selection among (i) random slopes and intercepts, (ii) random intercepts only, or (iii) no random effects (Bolker *et al.*, 2009; Zuur *et al.*, 2009). The minimal adequate model (optimal fixed structure) was then chosen using stepwise selection from the candidate models ranking under the second order Akaike Information Criterion (AICc), which increases the penalty for model complexity when sample size is small in comparison to the number of estimated parameters. We used a threshold of 4 AICc units to adopt a simpler model or to drop a term (Burnham & Anderson, 2002).

Hindcasting and projecting

Spatiotemporal estimates of land use are needed in order to estimate how average compositional similarity to primary vegetation have changed through history, and how they may change throughout this century. We used gridded (0.5 x 0.5 degree resolution) historical estimates of how the area under each land-use class changed from 1500-2005, from the History Database of the Global Environment HYDE 3.1 Klein Goldewijk *et al.* (2011). HYDE's historical land-use maps of Colombia are based on Etter & van Wyngaarden (2000) and Etter, McAlpine & Possingham (2008) statistics for historical landscape change in each region, which were derived from contemporary and historical maps, statistical data (socio-economical, demographic and geographic variables) and historical writings (books and chronicles). Projecting compositional intactness from 2005-2100 used the gridded land-use change data associated with the four Representative Concentration pathways (RCPs): IMAGE-RCP2.6, GCAM-RCP4.5, AIM-RCP6.0 and MESSAGE-RCP8.5 (Hurtt *et al.*, 2011). The Intergovernmental Panel on Climate Change (IPCC) developed the RCPs as a reference group of climate-change projections having different rates and magnitudes of land-use changes and greenhouse gas emissions (Moss *et al.*, 2010; Van Vuuren *et al.*, 2011). Each RCP aims to meet its specified level of radiative forcing (denoted by the number in its name, in W/m2) using a different combination of socio-economic, demographic, technological and policy assumptions (Harfoot *et al.*, 2014; Moss *et al.*, 2010).

We applied the coefficients from the minimal adequate model for compositional intactness with the historical and future land-use estimates to produce maps of compositional intactness, i.e., similarity to primary vegetation every 15 years from 2005-2095 for all the RCPs except MESSAGE-RCP8.5 (2005 and then every 20 years 2010-2090). We then

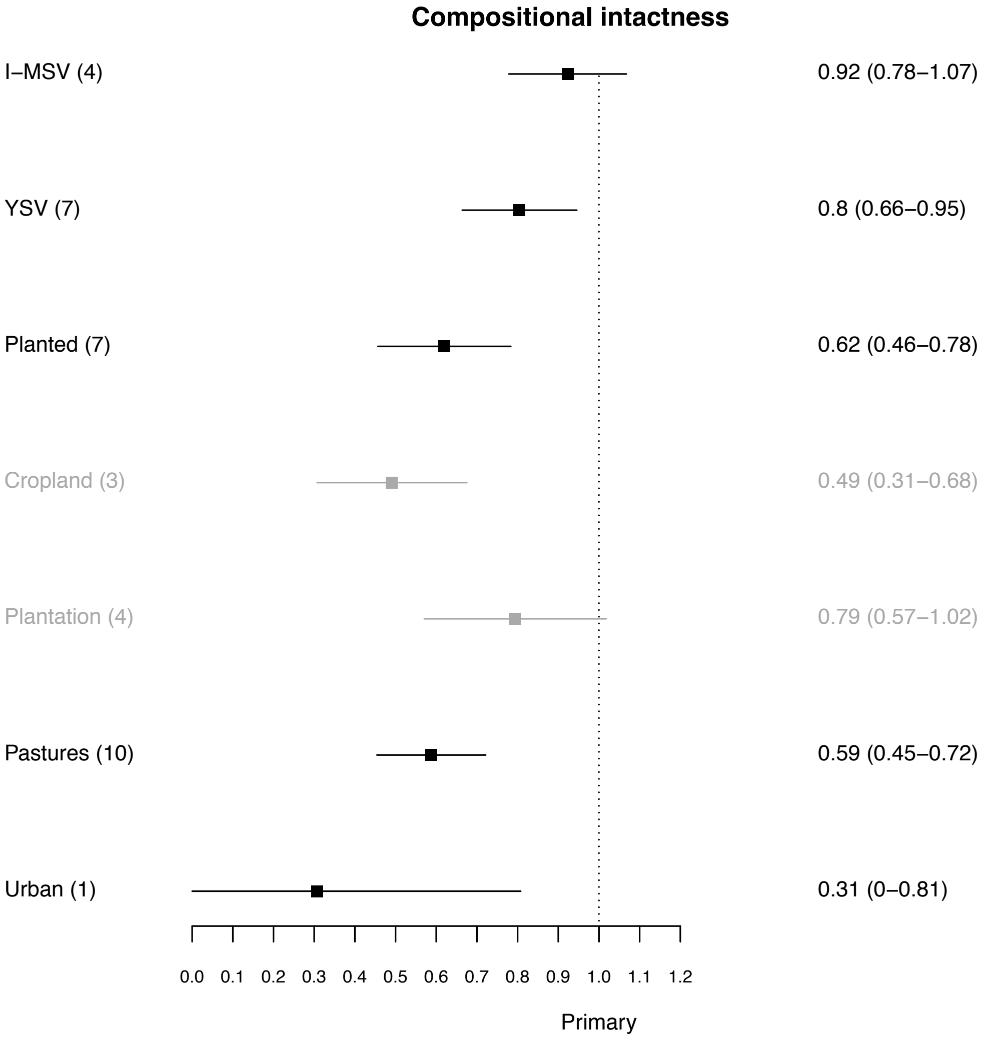


Figure 5.2: Community similarity for each land-use classes compared to primary vegetation as measured by the Sørensen similarity index. Error bars show coefficients estimates $\pm 95\%$ CI. Black points correspond to the estimates from the minimal adequate model (optimal fixed structure). The studies in cropland and plantations (grey points) were merged as "planted" in this model to keep a balanced distribution of studies among land use classes. Numbers within the left parenthesis represent the number of studies used per land-use level. I-MSV = Intermediate and mature secondary vegetation combined, YSV = Young secondary vegetation, Planted = Cropland and plantation.

aggregated the map for each date to give time series of the estimated average change within each region and across Colombia as a whole. When aggregating in this way, we weighted grid cells by their land area and current vertebrate species richness (Newbold *et al.*, 2015), to reflect that changes in more diverse regions might be of more concern (though vertebrate richness is at best an imperfect surrogate for invertebrate richness). We used the uncertainty in the model coefficients and model structure to place 95% confidence intervals around the biodiversity estimates for past and future dates, rescaling uncertainty to be zero in 2005 to separately show uncertainty in hindcasts and projections. However, we were unable to integrate uncertainty in the pressure data, because no uncertainty estimates are available for them (Newbold *et al.*, 2015).

5.4 Results

5.4.1 Biodiversity response to land use change

Compositional intactness varied significantly among land uses ($\chi^2 = 51.98$; d.f = 5; $P < 0.001$). While species composition in secondary vegetation, especially older secondary vegetation was similar to that in primary vegetation, the composition in all human-dominated land uses was markedly different (Figure 5.2). Human population density and distance to roads did not explain significant additional deviance (Table 5.1).

Model	d.f.	Δ AIC	Rank
LU	9	0.00	1
LU + logdistRd	10	6.32	2
LU + logHPD	10	8.01	3
Null	4	17.64	4
logdistRd	5	21.18	5
logHPD	5	24.44	6
LU + logHPD × logdistRd	12	25.58	7
LU × logdistRd + logHPD	15	38.55	8
LU × logHPD + logdistR	15	42.13	9
LU × logHPD × logdistRd	24	100.27	10

Table 5.1: Change in Akaike's information criterion (Δ AIC) and model rank for all models fitted in the community similarity analysis. LU, Land use; logHPD, log(human population density+1); logdistRd, log(distance to roads+1). See Figure 5.2 as a reference for land use classes used in the analysis. $N = 115$ sites in 15 studies.

Maps for the year 2005 of compositional intactness (Figure 5.3) show markedly lower values outside the Amazonian region, especially in the Andean region. Across Colombia as a whole, we estimate that average compositional intactness was 82% in 2005 (see Figure E.1).

5.4.2 Historical and future projections of diversity

Our hindcasts suggest that the extent and timing of declines in compositional intactness have varied markedly among regions (Figure 5.4). Although all regions show a tendency for accelerating decline through the 20th century, this acceleration is most pronounced in the Pacific and especially in the Caribbean region. While the Amazonian region was still

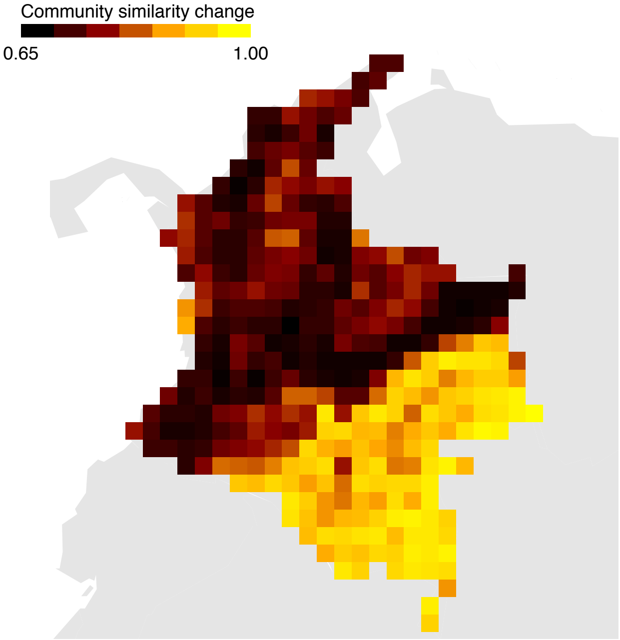


Figure 5.3: Net change in community composition in Colombia caused by land use by 2005 using 0.5 x 0.5 degrees resolution map.

92% compositionally intact in 2005, the estimates for the other regions are between 75% and 79%. Projections of compositional intactness differ strongly among RCP scenarios (Figure 5.4). It is projected to increase under GCAM-RCP4.5, because of a projected expansion of secondary forest outside Amazonia. In all regions compositional intactness is projected to decrease under MESSAGE-RCP8.5 scenario.

5.5 Discussion

5.5.1 Time course of compositional change in Colombia

During the history of Colombia, the complexity of primary vegetation has increasingly been replaced with relatively simple croplands, plantations, urban environments and especially pastures (Etter *et al.*, 2008; Etter & van Wyngaarden, 2000). We have shown (Figure 5.2), using Colombian data from a wide range of taxonomic groups, that these land uses support assemblages whose composition differs markedly from those in primary vegetation. The consequence of ongoing habitat conversion has been to drive down the average compositional intactness of assemblages across Colombia to 82% (Figure E.1), with

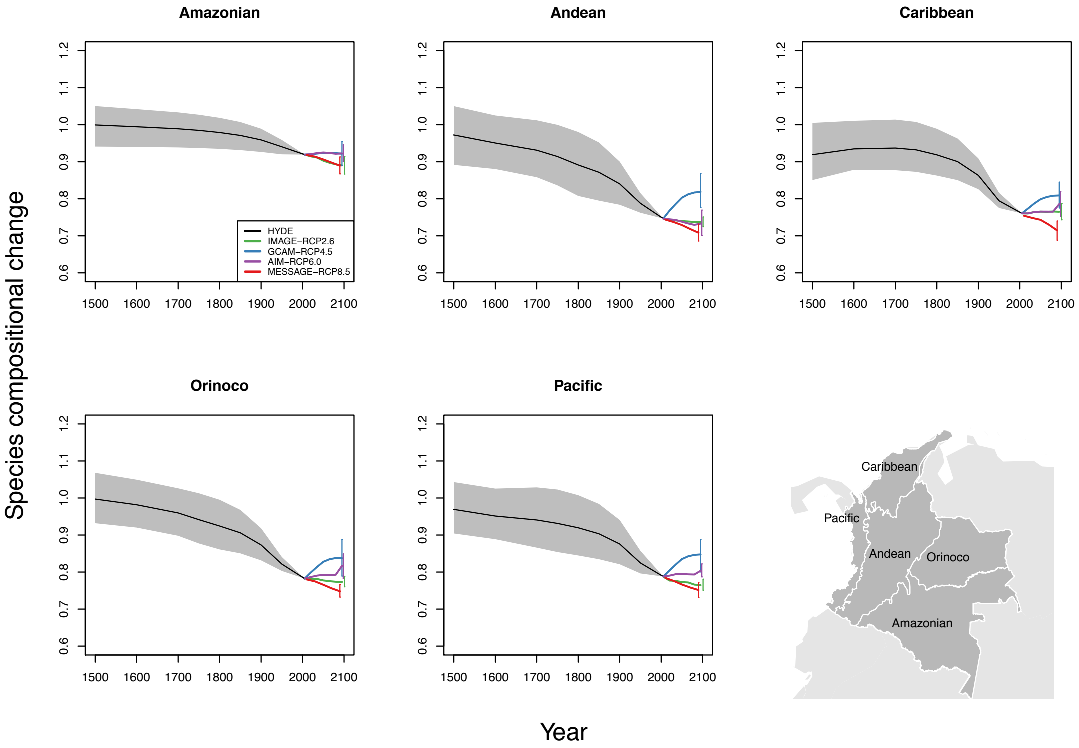


Figure 5.4: Projected net change in compositional intactness from 1500 to 2095 by region in Colombia. Grey shading (historical) and error bars (future) show ± 95% confidence intervals. Future projections are based on the four RCP scenarios (Table 5.2). We used the uncertainty in the model coefficients and model structure to place 95% confidence intervals around the biodiversity estimates for past and future dates, rescaling uncertainty to be zero in 2005 to separately show uncertainty in hindcasts and projections. We were unable to integrate uncertainty in the pressure data, because no uncertainty estimates are available for them (Newbold *et al.*, 2015).

the greatest reductions being in regions where conversion has been widespread (Figure 5.3). Only the Amazonian region is inferred to have largely intact assemblages ($>$ 90% intact: Figure 5.3 and 5.4).

In this analysis, three land uses in particular are associated with low compositional intactness: pasture, cropland and urban. Of these, pasture is most important in reducing overall compositional intactness: historically cattle grazing have expanded at expense of forests and croplands dominating over the 90% of cleared areas Colombia in 2000 (Etter *et al.*, 2008; Etter & van Wyngaarden, 2000).

Anthropogenic effects on assemblage composition are likely to involve biotic homogenisation (McKinney & Lockwood, 1999), as the simplification of habitats favours ecological generalists over specialists (Gámez-Virués *et al.*, 2015), reducing assemblage-level functional diversity (Olden *et al.*, 2004). The average decline in compositional intactness shown in Figure 5.4 is therefore potentially of concern from the perspective of ecosystem function as well as conservation.

Our projections of how compositional intactness may change in the future within each region (Figure 5.4) should not be viewed as exact predictions; rather they provide a practical way to explore the possible effects of different climate-change policies on ecological assemblages. Each representative concentration pathway (RCP) makes different socio-economic assumptions (summarised in Table 5.2) to achieve its target level of radiative forcing. Among the scenarios, MESSAGE-RCP8.5 (business-as-usual) projects the fastest decline, with compositional intactness falling to an average of 79% (77%-81%) by 2090. Under this RCP, despite yield improvements, forested areas will be replaced by croplands and pastures to meet rising demands for food, energy and fibre from a rapidly-growing human population (Brooks *et al.*, 2002; Harfoot *et al.*, 2014; Hurtt *et al.*, 2011). In contrast, GCAM-RCP4.5 is projected to offer the best outcome for biodiversity in Colombia — intactness rising to an average of 86% (82%-90%) by 2095 — agreeing with conclusions of a global study (Newbold *et al.*, 2015). This RCP is based on a fairly low stabilization scenario of greenhouse-gas (GHG) emissions through an expansion of forested areas by means of effective carbon pricing. Under this scenario, croplands and pastures will be replaced by forested areas, with yield improvement, dietary shifts and international trade able to satisfy food demand from a human population that is smaller than under MESSAGE-RCP8.5 (Harfoot *et al.*, 2014; Hurtt *et al.*, 2011). The expansion of secondary forest, which is permitted to mature in GCAM-RCP4.5, is projected to improve average compositional intactness in Colombia (Figure 5.4). However, these results focus only on land use change, ignoring the effects of climate change on biodiversity. The combination of land-use change and climate could increase the impact of the GCAM-RPC4.5 scenario on biodiversity compared with the most ambitious climate-change mitigation target IMAGE-RCP2.6. Therefore, future analyses should study the effects of climate change, land-use change and their interaction to evaluate in more detail the potential impact of future climate policies on biodiversity.

Scenario	**IMAGE-RCP2.6**	**GCAM-RCP4.6**	**AIM-RCP6.0**	**MESSAGE-RCP8.5**
Climate change policy	Very-low greenhouse concentration pathway. Mitigation of air pollutants through energy efficient policies based on renewable energy and bio-fuels.	Medium-low greenhouse concentration pathway. Mitigation based on carbon storage pricing, diet shift, decrease of energy consumption, crop yield improvement.	Medium baseline greenhouse concentration pathway. Mitigation based on technology development. Growing economy and population density.	High baseline greenhouse concentration pathway. No mitigation policies. High population growth and lower rate of technology development
Primary	Decrease	Decrease	Decrease	Decrease
Secondary	Medium increase	Significant increase	Significant increase	Medium increase
Cropland	Significant increase	Decrease	Medium increase	Medium increase
Pasture	Constant	Medium decrease	Decrease	Medium increase
Urban	Constant	Constant	Increase	Increase

Table 5.2: Main features of the RCP scenarios. Information based on Van Vuuren *et al.* (2011); Harfoot *et al.* (2014).

5.5.2 Secondary vegetation

Assemblages in maturing (but not young) secondary vegetation had similar composition to those in primary vegetation (Figure 5.2). Such assemblages are also often comparable to those in primary vegetation in terms of species richness (Barlow *et al.*, 2007; Gilroy *et al.*, 2014; Martin *et al.*, 2013; Newbold *et al.*, 2015; Wright & Muller-Landau, 2006) and plant biomass (Gilroy *et al.*, 2014; Martin *et al.*, 2013; Poorter *et al.*, 2016), suggesting that natural forest regeneration may provide important conservation and ecosystem services to buffer human disturbances. However the conservation value of secondary vegetation is likely to depend on the configuration of the landscape (Gardner *et al.*, 2009), which we were unable to incorporate in our analyses. Patches of recovering forest that are near to mature forest are likely to recover more quickly and more completely than are isolated patches, given many forest specialists are dispersal-limited (Hermy & Verheyen, 2007). Edge effects can mask the real contribution of regenerated habitats since generalist species can move through secondary vegetation from nearby pristine habitats (Brook *et al.*, 2006; Chazdon *et al.*, 2009; Norden *et al.*, 2009; Pulliam & Danielson, 1991). Because the surveys in our analyses mostly considered sets of sites that were close together (the median separation between the most distant pair of sites within each study was 1.6km), the inferred high compositional intactness of maturing secondary vegetation should be viewed as being contingent on there being nearby primary forest.

Assemblages in plantations and young secondary vegetation were less compositionally intact than those in primary or older secondary vegetation, but more so than those in cropland or pasture. This finding is consistent with Gilroy *et al.* (2015) multi-taxon assessment of community composition in natural forest, oil palm plantations and improved pasture in the Colombian Llanos in the Orinoco region. They found the amount of forest cover was an important determinant of diversity in birds, but not of dung beetles, ants or herpetofauna. They also reported depauperate diversity and reduction of compositional intactness in pastures compared with plantations. Once again the proximity of remnant forest could play an important role in maintaining compositional intactness in plantations and secondary vegetation. Other variables such as forest cover could also have a positive effect on the probability of occurrence of certain species (Newbold *et al.*, 2014). Further research regarding the role of indirect factors such as edge effects, habitat fragmentation and forest cover would help to improve estimates of the impacts of these land uses on diversity.

5.5.3 Biases and limitations

Although our models and projections infer a marked decline in compositional intactness in four of Colombia's five regions (Figure 5.4), there are several reasons why the decline may have been more severe than we estimate. First, published biodiversity surveys that set out to compare different habitat types may tend to underestimate the true diversity difference between habitats. Natural habitats tend to be more structurally complex than converted ones, often greatly so. Surveys comparing habitats of differing structural complexity are likely to use only those sampling methods that can easily be used in the simpler habitat. The sampling methods most commonly used in our data sets are baited/pitfall traps (six studies); fixed plots/quadrats (three studies) and visual encounter survey (three studies), all of which can be used in agricultural fields as well as in primary forest. A suite of methods widely used in mature forest (e.g., canopy fogging), and which sample different species pools (Longino *et al.*, 2002), are unlikely to be applied to structurally simpler sites. Therefore, even though our models indicate assemblages in pasture to be very distinct from those in primary vegetation, we may be underestimating how distinct they are. A related possible bias is that researchers may tend to make comparisons in which the result is not obvious. For example, intensively managed monoculture plantation forests obviously have fewer tree species than natural forest, but few surveys comparing the two habitats consider tree diversity (one exception being Phalan *et al.*, 2011). These biases apply not only to our study, but also to any synthetic comparisons that attempt to estimate overall effects of land-use change on species richness (e.g., Alkemade *et al.* 2009; Gibson *et al.* 2011; Newbold *et al.* 2015) or composition (e.g., Martin *et al.* 2013; Newbold *et al.* 2016b).

There are further grounds for caution in interpreting our statistical models and projections from them. First, different taxonomic groups may respond differently to human impacts (Gilroy *et al.*, 2014; Lawton *et al.*, 1998; Newbold *et al.*, 2014). Even the similarity between assemblages in primary and secondary vegetation can vary among taxa and functional groups: Barlow *et al.* (2007) multi-taxon comparison in Brazil showed that the proportion of primary-forest species also found in secondary forest ranged from below 40% (for trees, lianas, birds and grasshoppers) to over 80% (scavenger flies, large mammals and orchid bees). Martin, Newton & Bullock (2013) meta-analysis of mostly Neotropical (but not Colombian) studies reported that local species richness in secondary forest recovered much more quickly for trees than for epiphytes. Furthermore, in amphibians and reptiles assemblages in secondary forests in southeastern Mexico, Hernández-Ordóñez *et al.* (2015) showed that the recovery of local species richness is faster than abundance or composition. Our data set was not large enough or balanced enough to allow models to be fitted that explicitly test for and accommodate among-taxon variation. Whereas a multi-continent analysis of responses of tropical forest vertebrates showed marked differences among classes (Newbold *et al.*, 2014), a still-broader global analysis found responses of vertebrates, invertebrates and plants did not differ significantly (Newbold *et al.*, 2015). Differences in site size – especially within studies – could affect our model estimates; however, we did not find a significant relationship between maximum linear extent (MLE, our site size measure) and compositional intactness ($\chi^2_{1,84} = 0.016$, $p = 0.90$) (see Figure E.2). Only three studies reported differences in maximum linear extent among sites; removing these studies did not alter significantly our model estimates (see Figure E.3).

A second reason for caution is the uneven geographic coverage of our data. Within Colombia, biodiversity sampling appears to be biased towards the Andean region, from which around three quarters of our studies come (Figure 5.1). This bias is not surprising since this area harbours the main cities and therefore the main academic institutions in Colombia (Arbeláez-Cortés, 2013); proximity to research institutions has been identified as a strong biasing factor in the state of global biodiversity knowledge (Meyer *et al.*, 2015). Lacking enough data to test for regional differences, we have assumed that the compositional similarity between sites in different land uses is consistent among regions; the differences in the trajectories shown in Figure 5.4 arise because of different land-use histories rather than because the regional biotas differ in their intrinsic sensitivity to land-use change. Such differences have been suggested previously (e.g., Gibson *et al.* 2011, argued that Asia's biota is more sensitive than that of other regions). A priority is therefore to increase sampling effort in areas having limited information, perhaps focusing on identified hotspots of land-use change in the country (Armenteras *et al.*, 2013; Etter *et al.*, 2006a). Moreover, additional studies would improve the confidence in all parameter estimates and would allow more complex model structures (e.g., random slopes and intercepts models), as well as allowing investigation of variation among taxa and regions, and robust modeling of a wider range of biodiversity measures (e.g., species richness). Unfortunately, few data sets are available from regions with recent and ongoing transformation, places where the knowledge of species composition is scarce, and where there is a strong demand for natural resources such as the Orinoco and the Pacific regions. However, since the Andean region has a long history of transformation since the pre-Spanish colonization, further sampling is also needed in view of possible extinction debts (Tilman *et al.*, 1994), high levels of endemism (Orme *et al.*, 2005), unique evolutionary history (Madriñán *et al.*, 2013) and high risk of extinction (Brooks *et al.*, 2002).

Publication bias can cause negative (i.e., non-significant) results to be under-represented in the literature, especially in high-impact journals (Murtaugh, 2002); this potential bias is mitigated somewhat here as our literature search includes unpublished dissertations and articles published in Spanish in local journals as well as English-language papers in international journals. Eleven of the 17 sources had been published in local journals, and seven were written in Spanish, showing that there is still a language barrier to accessing biodiversity studies in highly diverse and vulnerable countries like Colombia. Since most of the biodiversity data are still hidden in unpublished information (Corlett, 2011), more effort should be invested in translation and migration of local field-data studies especially in the tropics where most of the diversity and recent land-use change is concentrated.

A final caveat is that the uncertainty in hindcasts and projections of compositional uncertainty through time (Figure 5.4) incorporates only the uncertainty of parameter estimates, meaning that the true uncertainty is greater than we have been able to accommodate. This reflects the lack of any uncertainty estimates associated with the historical and future land-use data we used Hurtt *et al.* (2011).

5.5.4 Concluding remarks

This study represents the first attempt to hindcast and project land-use impacts on local biodiversity in Colombia, and highlights the irreplaceability of natural forest. We estimate that land-use change has already reduced average compositional intactness by 18% across this megadiverse country, though we caution that the decline may have been even larger. Our projections suggest that future trends depend on the socioeconomic path that is chosen: expansion of secondary forests under a strong carbon market (GCAM-RCP4.5) can reverse the decline, whereas “business as usual” (MESSAGE-RCP8.5) will exacerbate it. Given the rate of land-use change in the country, particularly in those areas considered vulnerable, additional field-based studies are urgently needed to improve the robustness of models such as ours and to account for differences among regions, taxa and studies making them better able to inform mitigation policies of land-use change.

CHAPTER 6

General discussion and conclusions

In this thesis I have used a range of case studies, from a variety of taxonomic groups and geographical scales, to show that — even at a clade-specific level — biodiversity has been shaped by a variety of mutually reinforcing drivers whose influence can change over time (Chapter 2 and 3); and that these legacies from the past can affect the performance of species under environmental change (Chapter 4). I have also shown that human impacts are changing the composition of assemblages in one of the most vulnerable and biodiverse places on the globe and that these changes are likely to continue unless sociopolitical decisions are taken to reverse them (Chapter 5). The results of this thesis make a significant contribution to the fields of plant macroevolution and conservation macroecology, by helping to (i) understand the dynamics of diversification in one of the most diverse and economically important plant groups; (ii) characterise which plant species are more sensitive or more resilient to land-use change; and (iii) show how neotropical assemblages are currently responding — and are likely to respond in the future — to land-use change.

Dynamics of diversification and single-cause hypotheses

One of the most important findings to emerge from the results of the study of diversification of *Solanum* in this thesis is that the dynamics shaping the diversity of large genera may be more complex than we thought. A variety of mutually reinforcing drivers appear to have formed the current patterns of *Solanum* diversity, which cautions against the single-cause hypotheses, such as the influence of “key innovations” often invoked in macroevolutionary studies (Donoghue & Sanderson, 2015). These findings also raised important questions about the reliability of global studies that assume a single model for the evolution of character change and diversification across large phylogenies (e.g., Zanne *et al.* 2014), or studies that use average values of diversity, trait distribution or geographical distribution in genus-level studies (e.g., Cornwell *et al.* 2014). For instance, Beaulieu *et al.* 2013 and Chira & Thomas 2016 demonstrate that heterogeneity of diversification rates can have significant consequences for the model selection of trait change that best fits comparative data. In general, therefore, the understanding of the magnitude and location of shifts in diversification rates should be included as prior information in macroevolutionary analyses, especially in analyses of trait evolution (Chira & Thomas, 2016; Morlon, 2014); or at least models of trait evolution must not implicitly assume that such shifts do not occur.

Chapters 2 and 3 highlight the importance of context dependency in the study of diversification dynamics (De Queiroz, 2002; Donoghue, 2005; Donoghue & Sanderson, 2015). These chapters show that, for *Solanum* species to radiate into the Old World, it was not only necessary for there to be a geographical movement, but the confluence between environmental changes and character innovation was also important to spark this diversification. Therefore, focusing on the effects of the *combination* of events and attributes on changes in diversification could represent a more productive framework to understand diversity processes (Donoghue & Sanderson, 2015). This argument could also explain that, although many correlates of diversification rates have been reported in the literature, few have either much generality or high explanatory power. These results suggest the need to consider more complex analyses that involve the integration of phylogenies with other biological and Earth system data sources including geography, climate, historical biogeography and physiology among others.

Global analyses vs multiple clade-specific analyses

Both global multi-clade analyses and smaller-scale clade-specific analyses make important contributions to our understanding of the processes that have shaped biodiversity. Global comparative analyses can elucidate general patterns of diversity, highlighting phenomena that are sufficiently general and widespread that their broad explanations must also be general. However, such analyses do not provide many insights into the underlying mechanisms (Schluter, 2016) and a cost of generality is often the omission of clade- and context-specific factors that would provide much greater statistical explanatory power. In contrast, multiple case studies of diversification, biogeography and trait change within clades offer a more reliable approach to understand these mechanisms (Donoghue & Edwards, 2014; Schluter, 2016). Studies of diversification focusing on specific clades can make use of higher-quality data (e.g., well-sampled and more secure phylogenies) than global multi-clade studies, and additionally allow more careful evaluation of taxonomic and geographic sampling biases, which are regularly ignored in global studies. Such clade-specific analyses can also benefit from expert evaluation of the potential biases and the merits of the multiple sources of evidence (as in Chapter 2 and 3), providing a solid foundation from which to draw robust generalities of drivers of diversification across multiple clades. Given the vast diversity of *Solanum*, however, it is still necessary to improve the sampling of DNA, traits and geographic information of undersampled subclades to obtain a complete and accurate picture of the diversification of this megadiverse genus.

Taxonomic stability effects on macroevolution

Taxonomic instability is an important and ongoing source of bias in macroevolutionary studies which has not received the same attention as other sources of error such as phylogenetic uncertainty (Purvis *et al.*, 1994; Duchêne & Lanfear, 2015; Rabosky, 2015b), fossil calibrations (Dornburg *et al.*, 2011; Toussaint & Condamine, 2015; Yang & Donoghue, 2016) or a misspecified evolutionary model (Diaz-Uriarte & Garland, 1996, 1998). Many factors can bias the taxonomic effort towards some groups rather than others. For example, taxa with large body size and geographic range size are usually described earlier and tend to have greater taxonomic maturity than those that are smaller or narrow-range endemics (Patterson, 1994; Gaston, 1991; Gaston *et al.*, 1995). Another important source of heterogeneity might be produced by differences in taxonomic opinions between "splitters" and "lumpers", which can create a false diversity contrast among groups. Using a series of simulations, Faurby *et al.* (2015) found that inferences in changes of diversification rates through time and the diagnosis of unique evolutionary histories (i.e., clades with different diversification dynamics compared with the background) were highly sensitive to variation in taxonomic opinion.

Due to the wide variety of species in *Solanum*, the taxonomy of this genus has been mostly limited to studies at an infra-generic level, with monographs usually focused on species-level treatments of smaller groups or specific regions. Crucially, because species description accumulates non-randomly among regions and clades, the current understanding of the systematics of any clade cannot be guaranteed to reflect what a complete understanding would show. As a result, the conclusions of Chapters 2 and 3 might be sensitive to the different states of taxonomic maturity among groups of *Solanum*, even after correcting for non-random incomplete taxon sampling in the phylogeny. For example, although there have been historical studies of solanums in the Old World, there has been a recent increase in the number of descriptions of new *Solanum* species in these regions in the last decade, especially in Australia, which could inflate the expected number of species in this region. In contrast, the number of accepted species of *Solanum* in the Potato group has decreased in the past decade due to a set of comprehensive taxonomic revisions such as those by Spooner *et al.* (2004, 2014). Although a preliminary analysis suggests that the taxonomic maturity of *Solanum* species from the Old World spiny clade is similar to the taxonomic maturity of the genus as whole (see Figures F.1 and F.2 in Supplementary Information F), a more detailed analysis is necessary to understand whether the variation of taxonomic maturity among groups affects the conclusions drawn from Chapters 2 and 3.

Limitations of current approaches for estimating deep time patterns of diversification

The use of phylogenetic comparative methods (PCMs) to investigate drivers and models of diversification has grown exponentially over recent years (see Figure 1 in Cooper *et al.* 2016). Unfortunately, the development of more user-friendly and extensively documented software implementing PCM approaches has focussed the attention away from their limitations and assumptions (Cooper *et al.*, 2016). For instance, methods for inferring speciation-extinction dynamics and diversification drivers from molecular phylogenies are unreliable under many realistic scenarios since these phylogenies do not provide direct information about extinction events (Rabosky, 2010; Moore *et al.*, 2016; Rabosky & Goldberg, 2015). Even assuming that phylogenies are accurate, they are usually undersampled or have sampling biases particularly near the "tips" (e.g., Jetz *et al.* 2012 and Zanne *et al.* 2014). One of the major focuses in diversification studies is the inference of extinction dynamics from molecular phylogenies, which leads to the question of the reliability of our interpretation of patterns in deep time (Ricklefs, 2007; Purvis, 2008; Rabosky, 2010; Schluter & Pennell, 2017). It is, therefore, necessary to integrate paleontological data such as the fossil record and paleoclimatic data to infer meaningful dynamics of speciation, extinction and past distributions in deep time (Fritz *et al.*, 2013); and focus on organisms with good fossil records such as foraminifera, bivalves or molluscs when understanding longstanding gradients of biodiversity (Aze *et al.*, 2011; Ezard *et al.*, 2011; Fenton *et al.*, 2016; Jablonski *et al.*, 2017).

Despite the problems of using molecular phylogenies to infer ancient patterns of diversity, the study of relatively recent and ongoing radiations such as *Solanum* might be less affected by extinction since lineages that emerged more recently have had less opportunity to have gone extinct (Nee *et al.*, 1992, 1994a; Jetz *et al.*, 2012; Schluter & Pennell, 2017). Chapters 2 and 3 provide relevant information about the maintenance of and recent changes in *Solanum* diversity gradients, which could be valuable to predict future dynamics. For example, these chapters show that *Solanum* has experienced a recent explosive radiation out of the Neotropics in regions associated with arid and hot conditions. These results along with field habitat descriptions could indicate that the diversity of *Solanum* is formed by opportunistic species that have the eco-evolutionary capacity to colonise and expand in open and disturbed areas. According to future land-use projections (IPCC, 2007), human-dominated habitats will continue to extend at the expense of forested regions and are likely to transform the local climatic conditions towards warmer temperatures, and land-use change can cause microclimates to warm much more rapidly than projected by global climate change models (Senior *et al.*, 2017). In this context, *Solanum* species are not only likely to persist under these conditions, but may even benefit and expand their ranges based on their evolutionary legacy. However, additional experimental and field studies, and a deeper understanding of the intraspecific differences in their responses to environmental change, are necessary to support this hypothesis.

Biotic responses to land-use change

The future of biodiversity will depend on how species respond to the current changes in the atmosphere and landscape. The results of Chapters 4 and 5 show that these responses vary widely among plant species and among regions of the Neotropics, but a high proportion of the variation in the models that best fitted the biotic responses to land-use change in these chapters was unexplained. This means that other sources of variation — geographic, taxonomic or trait-based — need to be included in future studies to give a more mature understanding.

Besides the idiosyncrasies in species' sensitivities, the importance of the random slopes in the models fitted in Chapter 4 suggests that the relationship between climatic affiliations and species' sensitivity varies significantly among taxonomic groups. The assessment of the phylogenetic signal of species' sensitivity could help us to provide a broader picture of the effect of the land-use change. For example, if species' sensitivities show phylogenetic signal, this could mean whole clades may be affected producing a significant risk to ecosystem services (Purvis *et al.*, 2000; Díaz *et al.*, 2013). However, these approaches do not give any insight into the mechanisms that drive the vulnerability of the species which ultimately inform conservation efforts (but see Fritz & Purvis 2010). For that, the collection of trait data would be more appropriate, allowing analysis of ecological differences that drive these species' responses (De Palma *et al.*, 2015). Although there is an increased availability of trait databases such as TRY (Kattge *et al.*, 2011) and BIEN (Enquist *et al.*, 2016), these are still sparse and limited. However, the development of new approaches for gap-filling and trait prediction (e.g., Schrodt *et al.* 2015) provide a promising future for studies looking at the influence of traits on species' responses to environmental change.

The PREDICTS dataset, which is the principal source of information for Chapter 4 and 5, is currently the most representative collation of data on how local land-use impacts terrestrial biodiversity (Hudson *et al.*, 2017). However, some regions need additional studies to have a more accurate picture of the responses of both plant species and neotropical biodiversity. Additional information is also required to examine other land-use effects such as legacies of past land uses (e.g., Perring *et al.* 2016), habitat fragmentation (e.g., Haddad *et al.* 2015, Phillips *et al.* 2017) and land-use intensity (e.g., Newbold *et al.* 2015), all of which are likely to explain some of the as-yet-unexplained variation in these studies. As in the macroevolutionary analyses of Chapters 4 and 5, context dependency is likely to be widespread in analyses of biotic responses to human pressures.

As discussed in Chapters 4 and 5, a major limitation of my analyses and those of others is the space-for-time approach (also discussed by Purvis *et al.* in press). This approach involves the comparison of diversity across sites that — although otherwise matched — differ in their current land use as a proxy to estimate how land-use change affected local diversity. Although the geographic and taxonomic span of the available information is much greater in these spatial comparisons than temporal data would permit, there are notable limitations that we need to consider. One significant constraint is the biotic lag in species responses (Essl *et al.*, 2015). The full impact of land-use transitions on assemblages and species might only become apparent after many years — centuries, even. Therefore, looking at temporal changes of diversity could provide a better understanding of the effect of the land-use change, for example, using time-series data or before-after-control-impact (BACI) studies. However, the availability of these time-series studies that include different land-use transitions remains very limited, and such studies are also sensitive to the sampling of temporal trends.

Concluding remarks

The biodiversity of the globe is currently facing significant changes in its composition as species' responses to changing environments are extremely heterogeneous. In this thesis, I have used multiple approaches and data sources to show how evolutionary processes have shaped current patterns of biodiversity (focusing on the genus *Solanum*), and how this diversity is responding to recent environmental changes. My results demonstrate the great dynamism involved in shaping global biodiversity, and the vast heterogeneity of the biotic responses to land-use change. Although there are still questions to be answered, the rapidly increasing availability and aggregation of biodiversity data will enable a more robust understanding of how life's variety arose, and the mechanisms that are driving the rearrangement of the assemblages under current global changes.

REFERENCES

Agapow, P.M. & Isaac, N.J. (2002) Macrocaic: revealing correlates of species richness by comparative analysis. *Diversity and Distributions*, **8**, 1, 41–43.

Alfaro, M.E., Santini, F., Brock, C., Alamillo, H., Dornburg, A. *et al.* (2009) Nine exceptional radiations plus high turnover explain species diversity in jawed vertebrates. *Proceedings of the National Academy of Sciences*, **106**, 32, 13410–13414.

Alkemade, R., van Oorschot, M., Miles, L., Nellemann, C., Bakkenes, M. *et al.* (2009) GLOBIO3: a framework to investigate options for reducing global terrestrial biodiversity loss. *Ecosystems*, **12**, 3, 374–390.

Arbeláez-Cortés, E. (2013) Knowledge of colombian biodiversity: published and indexed. *Biodiversity and conservation*, **22**, 12, 2875–2906.

Arbeláez-Cortés, E., Rodríguez, H.A. & Restrepo, M. (2011) Mixed bird flocks: patterns of activity and species composition in a region of the central Andes of Colombia. *Revista Mexicana de Biodiversidad*, **82**, 2, 639–651.

Armbrecht, I., Perfecto, I. & Silverman, E. (2006) Limitation of nesting resources for ants in colombian forests and coffee plantations. *Ecological Entomology*, **31**, 5, 403–410.

Armenteras, D., Cabrera, E., Rodríguez, N. & Retana, J. (2013) National and regional determinants of tropical deforestation in colombia. *Regional Environmental Change*, **13**, 6, 1181–1193.

Armenteras, D., Rodríguez, N., Retana, J. & Morales, M. (2011) Understanding deforestation in montane and lowland forests of the colombian andes. *Regional Environmental Change*, **11**, 3, 693–705.

Arribas, P., Velasco, J., Abellán, P., Sánchez-Fernández, D., Andujar, C. *et al.* (2012) Dispersal ability rather than ecological tolerance drives differences in range size between lentic and lotic water beetles (coleoptera: Hydrophilidae). *Journal of Biogeography*, **39**, 5, 984–994.

Aubriot, X., Singh, P. & Knapp, S. (2016) Tropical asian species show that the old world clade of 'spiny solanums' (solanum subgenus leptostemonum pro parte: Solanaceae) is not monophyletic. *Botanical Journal of the Linnean Society*, **181**, 2, 199–223.

Aze, T., Ezard, T.H., Purvis, A., Coxall, H.K., Stewart, D.R. *et al.* (2011) A phylogeny of cenozoic macroperforate planktonic foraminifera from fossil data. *Biological Reviews*, **86**, 4, 900–927.

Badgley, C. & Finarelli, J.A. (2013) Diversity dynamics of mammals in relation to tectonic and climatic history: comparison of three neogene records from north america. *Paleobiology*, **39**, 3, 373–399.

Baldwin, B.G., *Adaptive radiation of the Hawaiian silversword alliance: congruence and conflict of phylogenetic evidence from molecular and non-molecular investigations*, vol. 103. Cambridge University Press Cambridge, 1997, 1840–1843.

Baldwin, B.G. (2007) Adaptive radiation of shrubby tarweeds (deinandra) in the california islands parallels diversification of the hawaiian silversword alliance (compositae–madiinae). *American Journal of Botany*, **94**, 2, 237–248.

Baltzer, J., Davies, S.J., Bunyavejchewin, S. & Noor, N. (2008) The role of desiccation tolerance in determining tree species distributions along the malay–thai peninsula. *Functional Ecology*, **22**, 2, 221–231.

Bapst, D.W. (2012) paleotree: an r package for paleontological and phylogenetic analyses of evolution. *Methods in Ecology and Evolution*, **3**, 5, 803–807.

Barlow, J., Gardner, T.A., Araujo, I.S., Ávila-Pires, T.C., Bonaldo, A.B. *et al.* (2007) Quantifying the biodiversity value of tropical primary, secondary, and plantation forests. *Proceedings of the National Academy of Sciences*, **104**, 47, 18555–18560.

Barraclough, T.G., Harvey, P.H. & Nee, S. (1995) Sexual selection and taxonomic diversity in passerine birds. *Proceedings of the Royal Society of London B: Biological Sciences*, **259**, 1355, 211–215.

Barraclough, T.G. & Nee, S. (2001) Phylogenetics and speciation. *Trends in Ecology & Evolution*, **16**, 7, 391–399.

Bartoń, K. (2013) Mumin: Multi-model inference. R package version 1.9.13. `http://cran.r-project.org/package=MuMIn`.

Bates, D., Mächler, M., Bolker, B. & Walker, S. (2015) Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, **67**, 1, 1–48. doi:10.18637/jss.v067.i01.

Bean, A. (2004) The taxonomy and ecology of solanum subg. leptostemonum (dunal) bitter (solanaceae) in queensland and far north-eastern new south wales, australia. *Austrobaileya*, 639–816.

Beaulieu, J.M., O'Meara, B., Crane, P. & Donoghue, M.J. (2015) Heterogeneous rates of molecular evolution and diversification could explain the triassic age estimate for angiosperms. *Systematic biology*, syv027.

Beaulieu, J.M., O'Meara, B.C. & Donoghue, M.J. (2013) Identifying hidden rate changes in the evolution of a binary morphological character: the evolution of plant habit in campanulid angiosperms. *Systematic Biology*, **62**, 5, 725–737.

Bell, C.D. & Donoghue, M.J. (2005) Phylogeny and biogeography of valerianaceae (dipsacales) with special reference to the south american valerians. *Organisms Diversity & Evolution*, **5**, 2, 147–159.

Benton, M.J. (2009) The red queen and the court jester: species diversity and the role of biotic and abiotic factors through time. *science*, **323**, 5915, 728–732.

Bininda-Emonds, O.R., Cardillo, M., Jones, K.E., MacPhee, R.D., Beck, R.M. *et al.* (2007) The delayed rise of present-day mammals. *Nature*, **446**, 7135, 507–512.

Bivand, R. & Lewin-Koh, N. (2016) *maptools: Tools for Reading and Handling Spatial Objects*. R package version 0.8-39.

Blom, M.P., Horner, P. & Moritz, C., Convergence across a continent: adaptive diversification in a recent radiation of australian lizards. In *Proc. R. Soc. B*, vol. 283. The Royal Society, 2016, 20160181.

Bohs, L. (2004) A chloroplast dna phylogeny of solanum section lasiocarpa. *Systematic Botany*, **29**, 1, 177–187.

Bohs, L. (2005) Major clades in solanum based on ndhf sequence data. *Monographs in Systematic Botany*, **104**, 27.

Bohs, L. (2007) Phylogeny of the cyphomandra clade of the genus solanum (solanaceae) based on its sequence data. *Taxon*, 1012–1026.

Bohs, L. & Olmstead, R. (2001) A reassessment of normania and triguera (solanaceae). *Plant Systematics and Evolution*, **228**, 1-2, 33–48.

Bohs, L. & Olmstead, R.G. (1997) Phylogenetic relationships in solanum (solanaceae) based on ndhf sequences. *Systematic Botany*, 5–17.

Bolker, B.M., Brooks, M.E., Clark, C.J., Geange, S.W., Poulsen, J.R. *et al.* (2009) Generalized linear mixed models: a practical guide for ecology and evolution. *Trends in Ecology & Evolution*, **24**, 3, 127–135.

Brenes-Arguedas, T., Coley, P.D. & Kursar, T.A. (2009) Pests vs. drought as determinants of plant distribution along a tropical rainfall gradient. *Ecology*, **90**, 7, 1751–1761.

Brook, B.W., Bradshaw, C.J., Koh, L.P. & Sodhi, N.S. (2006) Momentum drives the crash: Mass extinction in the tropics. *Biotropica*, **38**, 3, 302–305.

Brook, B.W., Sodhi, N.S. & Bradshaw, C.J. (2008) Synergies among extinction drivers under global change. *Trends in ecology & evolution*, **23**, 8, 453–460.

Brooks, T.M., Mittermeier, R.A., Mittermeier, C.G., Da Fonseca, G.A., Rylands, A.B. *et al.* (2002) Habitat loss and extinction in the hotspots of biodiversity. *Conservation Biology*, **16**, 4, 909–923.

Brown, J.H. (1984) On the relationship between abundance and distribution of species. *The american naturalist*, **124**, 2, 255–279.

Bruneau, A., Dickson, E.E. & Knapp, S. (1995) Congruence of chloroplast dna restriction site characters with morphological and isozyme data in solanum sect. lasiocarpa. *Canadian Journal of Botany*, **73**, 8, 1151–1167.

Buckley, L.B. (2007) Linking traits to energetics and population dynamics to predict lizard ranges in changing environments. *The American Naturalist*, **171**, 1, E1–E19.

Burnham, K.P. & Anderson, D.R., *Model selection and multi-model inference: a practical information-theoretic approach.* Springer, 2002, 2nd ed.

Byrne, M., Steane, D.A., Joseph, L., Yeates, D.K., Jordan, G.J. *et al.* (2011) Decline of a biome: evolution, contraction, fragmentation, extinction and invasion of the australian mesic zone biota. *Journal of Biogeography*, **38**, 9, 1635–1656.

Byrne, M., Yeates, D., Joseph, L., Kearney, M., Bowler, J. *et al.* (2008) Birth of a biome: insights into the assembly and maintenance of the australian arid zone biota. *Molecular Ecology*, **17**, 20, 4398–4417.

Cabra-García, J., Bermúdez-Rivas, C., Osorio, A.M. & Chacón, P. (2012) Cross-taxon congruence of α and β diversity among five leaf litter arthropod groups in Colombia. *Biodiversity and Conservation*, **21**, 6, 1493–1508.

Calviño-Cancela, M., Rubido-Bará, M. & van Etten, E.J. (2012) Do eucalypt plantations provide habitat for native forest biodiversity? *Forest Ecology and Management*, **270**, 153–162.

Cardinale, B.J., Duffy, J.E., Gonzalez, A., Hooper, D.U., Perrings, C. *et al.* (2012) Biodiversity loss and its impact on humanity. *Nature*, **486**, 7401, 59–67.

Castillo, R.O. & Spooner, D.M. (1997) Phylogenetic relationships of wild potatoes, solanum series conicibaccata (sect. petota). *Systematic Botany*, 45–83.

Cavender-Bares, J., Ackerly, D.D., Hobbie, S.E. & Townsend, P.A. (2016) Evolutionary legacy effects on ecosystems: Biogeographic origins, plant traits, and implications for management in the era of global change. *Annual Review of Ecology, Evolution, and Systematics*, **47**, 433–462.

Chapin III, F.S., Zavaleta, E.S., Eviner, V.T., Naylor, R.L., Vitousek, P.M. *et al.* (2000) Consequences of changing biodiversity. *Nature*, **405**, 6783, 234–242.

Chazdon, R.L., Peres, C.A., Dent, D., Sheil, D., Lugo, A.E. *et al.* (2009) The potential for species conservation in tropical secondary forests. *Conservation Biology*, **23**, 6, 1406–1417.

Chen, I.C., Hill, J.K., Ohlemüller, R., Roy, D.B. & Thomas, C.D. (2011) Rapid range shifts of species associated with high levels of climate warming. *Science*, **333**, 6045, 1024–1026.

Chira, A.M. & Thomas, G.H. (2016) The impact of rate heterogeneity on inference of phylogenetic models of trait evolution. *Journal of evolutionary biology*, **29**, 12, 2502–2518.

Clark, J.L., Nee, M., Bohs, L. & Knapp, S. (2015) A revision of solanum section aculeigerum (the solanum wendlandii group, solanaceae). *Systematic Botany*, **40**, 4.

Collins, S.L., Suding, K.N., Cleland, E.E., Batty, M., Pennings, S.C. *et al.* (2008) Rank clocks and plant community dynamics. *Ecology*, **89**, 12, 3534–3541.

Condamine, F.L., Rolland, J. & Morlon, H. (2013) Macroevolutionary perspectives to environmental change. *Ecology letters*, **16**, s1, 72–85.

Cooper, N., Thomas, G.H. & FitzJohn, R.G. (2016) Shedding light on the âĂŸdark sideâĂŹof phylogenetic comparative methods. *Methods in ecology and evolution*, **7**, 6, 693–699.

Corlett, R.T. (2011) Trouble with the gray literature. *Biotropica*, **43**, 1, 3–5.

Cornwell, W.K., Westoby, M., Falster, D.S., FitzJohn, R.G., O'Meara, B.C. *et al.* (2014) Functional distinctiveness of major plant lineages. *Journal of Ecology*, **102**, 2, 345–356.

Cox, C.B. (2001) The biogeographic regions reconsidered. *Journal of Biogeography*, **28**, 4, 511–523.

Crawley, M.J., *The R book.* John Wiley & Sons, 2012.

Crayn, D.M., Rossetto, M. & Maynard, D.J. (2006) Molecular phylogeny and dating reveals an oligo-miocene radiation of dry-adapted shrubs (former tremandraceae) from rainforest tree progenitors (elaeocarpaceae) in australia. *American Journal of Botany*, **93**, 9, 1328–1342.

Crepet, W.L. & Niklas, K.J. (2009) Darwin's second "abominable mystery": Why are there so many angiosperm species? *American Journal of Botany*, **96**, 1, 366–381.

Crisp, M., Cook, L. & Steane, D. (2004) Radiation of the australian flora: what can comparisons of molecular phylogenies across multiple taxa tell us about the evolution of diversity in present–day communities? *Philosophical Transactions of the Royal Society B: Biological Sciences*, **359**, 1450, 1551–1571.

Crisp, M.D., Arroyo, M.T., Cook, L.G., Gandolfo, M.A., Jordan, G.J. *et al.* (2009) Phylogenetic biome conservatism on a global scale. *Nature*, **458**, 7239, 754.

Crisp, M.D. & Cook, L.G. (2007) A congruent molecular signature of vicariance across multiple plant lineages. *Molecular phylogenetics and evolution*, **43**, 3, 1106–1117.

Crisp, M.D. & Cook, L.G. (2013) How was the australian flora assembled over the last 65 million years? a molecular phylogenetic perspective. *Annual Review of Ecology, Evolution, and Systematics*, **44**, 303–324.

Daehler, C.C. (2003) Performance comparisons of co-occurring native and alien invasive plants: implications for conservation and restoration. *Annual Review of Ecology, Evolution, and Systematics*, **34**, 1, 183–211.

Davies, T.J., Savolainen, V., Chase, M.W., Goldblatt, P. & Barraclough, T.G. (2005) Environment, area, and diversification in the species-rich flowering plant family iridaceae. *The American Naturalist*, **166**, 3, 418–425.

De Palma, A., Kuhlmann, M., Roberts, S.P., Potts, S.G., Börger, L. *et al.* (2015) Ecological traits affect the sensitivity of bees to land-use pressures in european agricultural landscapes. *Journal of Applied Ecology*, **52**, 6, 1567–1577.

De Queiroz, A. (2002) Contingent predictability in evolution: key traits and diversification. *Systematic Biology*, **51**, 6, 917–929.

Dennis, E.B., Morgan, B.J., Roy, D.B. & Brereton, T.M. (2017) Urban indicators for uk butterflies. *Ecological Indicators*, **76**, 184–193.

Díaz, S., Purvis, A., Cornelissen, J.H., Mace, G.M., Donoghue, M.J. *et al.* (2013) Functional traits, the phylogeny of function, and ecosystem service vulnerability. *Ecology and Evolution*, **3**, 9, 2958–2975.

Diaz-Uriarte, R. & Garland, T. (1996) Testing hypotheses of correlated evolution using phylogenetically independent contrasts: sensitivity to deviations from brownian motion. *Systematic Biology*, **45**, 1, 27–47.

Diaz-Uriarte, R. & Garland, T. (1998) Effects of branch length errors on the performance of phylogenetically independent contrasts. *Systematic Biology*, **47**, 4, 654–672.

Dinero (2015) Composición de la economía colombiana 2015. Available at `http://www.dinero.com/economia/articulo/composicion-economia-colombiana-2015/214054` (accessed: 2016.09.29).

Dirzo, R. & Raven, P.H. (2003) Global state of biodiversity and loss. *Annual Review of Environment and Resources*, **28**, 1, 137–167.

Dominguez, Y. & Armbrecht, I. (2011) Response of ants and their seed removal in rehabilitation areas and forests at el cerrejón coal mine in Colombia. *Restoration Ecology*, **19**, 201, 178–184.

Donoghue, M.J. (2005) Key innovations, convergence, and success: macroevolutionary lessons from plant phylogeny. *Paleobiology*, **31**, S2, 77–93.

Donoghue, M.J. & Edwards, E.J. (2014) Biome shifts and niche evolution in plants. *Annual Review of Ecology, Evolution, and Systematics*, **45**, 547–572.

Donoghue, M.J. & Sanderson, M.J. (2015) Confluence, synnovation, and depauperons in plant diversification. *New Phytologist*.

Dornburg, A., Beaulieu, J.M., Oliver, J.C. & Near, T.J. (2011) Integrating fossil preservation biases in the selection of calibrations for molecular divergence time estimation. *Systematic biology*, **60**, 4, 519–527.

Dornelas, M., Gotelli, N.J., McGill, B., Shimadzu, H., Moyes, F. *et al.* (2014) Assemblage time series reveal biodiversity change but not systematic loss. *Science*, **344**, 6181, 296–299.

Douglas Bates, M.M. & Bolker, B. (2013) lme4: Linear mixed-effects models using s4 classes.

Drummond, C.S., Eastwood, R.J., Miotto, S.T. & Hughes, C.E. (2012) Multiple continental radiations and correlates of diversification in lupinus (leguminosae): testing for key innovation with incomplete taxon sampling. *Systematic Biology*, **61**, 3, 443–460.

Duchêne, S. & Lanfear, R. (2015) Phylogenetic uncertainty can bias the number of evolutionary transitions estimated from ancestral state reconstruction methods. *Journal of Experimental Zoology Part B: Molecular and Developmental Evolution*, **324**, 6, 517–524.

Dunal, M.F., *Histoire naturelle, medicale et economique des Solanum et des genres qui ont ete confondus avec eux.-Paris, Amand König 1813*. Amand König, 1813.

Dunal, M.F., *Solanorum Generumque Affinium Synopsis*. Montpellier, France, 1816.

Dunn, R.R. (2004) Recovery of faunal communities during tropical forest regeneration. *Conservation Biology*, **18**, 2, 302–309.

Dupin, J., Matzke, N.J., Särkinen, T., Knapp, S., Olmstead, R.G. *et al.* (2016) Bayesian estimation of the global biogeographical history of the solanaceae. *Journal of Biogeography*, **44**, 887âĂŞ899.

Edmonds, J.M. & Chweya, J.A., *Black nightshades, Solanum nigrum L. and related species*, vol. 15. Bioversity International, 1997.

Edwards, E.J., de Vos, J.M. & Donoghue, M.J. (2015) Doubtful pathways to cold tolerance in plants. *Nature*, **521**, 7552, E5–E6.

Ellis, E.C., Klein Goldewijk, K., Siebert, S., Lightman, D. & Ramankutty, N. (2010) Anthropogenic transformation of the biomes, 1700 to 2000. *Global ecology and biogeography*, **19**, 5, 589–606.

Enquist, B.J., Condit, R., Peet, R.K., Schildhauer, M. & Thiers, B.M. (2016) Cyberinfrastructure for an integrated botanical information network to investigate the ecological impacts of global climate change on plant biodiversity. Tech. rep., PeerJ Preprints.

Epskamp, S., Cramer, A.O.J., Waldorp, L.J., Schmittmann, V.D. & Borsboom, D. (2012) qgraph: Network visualizations of relationships in psychometric data. *Journal of Statistical Software*, **48**, 4, 1–18.

Erwin, T., *Taxonomy, phylogeny, and zoogeography of beetles and ants: a volume dedicated to the memory of Philip Jackson Darlington, JR.*, chap. The taxon pulse: a general pattern of lineage radiation and extinction among carabid beetles. Dordrecht, the Netherlands: Dr W. Junk Publishers, 1985, 437âĂŞ472.

Esquivel-Muelbert, A., Baker, T.R., Dexter, K.G., Lewis, S.L., Steege, H. *et al.* (2017) Seasonal drought limits tree species across the neotropics. *Ecography*, **40**, 5, 618–629.

Essens, T., van Langevelde, F., Vos, R.A., Van Swaay, C.A. & WallisDeVries, M.F. (2017) Ecological determinants of butterfly vulnerability across the european continent. *Journal of Insect Conservation*, 1–12.

Essl, F., Dullinger, S., Rabitsch, W., Hulme, P.E., Pyšek, P. *et al.* (2015) Delayed biodiversity change: no time to waste. *Trends in ecology & evolution*, **30**, 7, 375–378.

Etienne, R.S. & Haegeman, B. (2012) A conceptual and statistical framework for adaptive radiations with a key role for diversity dependence. *The American Naturalist*, **180**, 4, E75–E89.

Etienne, R.S., Haegeman, B., Stadler, T., Aze, T., Pearson, P.N. *et al.* (2011) Diversity-dependence brings molecular phylogenies closer to agreement with the fossil record. *Proceedings of the Royal Society of London B: Biological Sciences*, rspb20111439.

Etter, A., McAlpine, C., Phinn, S., Pullar, D. & Possingham, H. (2006a) Unplanned land clearing of Colombian rainforests: Spreading like disease? *Landscape and Urban Planning*, **77**, 3, 240–254.

Etter, A., McAlpine, C. & Possingham, H. (2008) Historical patterns and drivers of landscape change in Colombia since 1500: a regionalized spatial approach. *Annals of the Association of American Geographers*, **98**, 1, 2–23.

Etter, A., McAlpine, C., Seabrook, L. & Wilson, K. (2011) Incorporating temporality and biophysical vulnerability to quantify the human spatial footprint on ecosystems. *Biological Conservation*, **144**, 5, 1585–1594.

Etter, A., McAlpine, C., Wilson, K., Phinn, S. & Possingham, H. (2006b) Regional patterns of agricultural land use and deforestation in Colombia. *Agriculture, Ecosystems & Environment*, **114**, 2, 369–386.

Etter, A. & van Wyngaarden, W. (2000) Patterns of landscape transformation in Colombia, with emphasis in the andean region. *AMBIO: A Journal of the Human Environment*, **29**, 7, 432–439.

Ezard, T.H., Aze, T., Pearson, P.N. & Purvis, A. (2011) Interplay between changing climate and species' ecology drives macroevolutionary dynamics. *Science*, **332**, 6027, 349–351.

Faurby, S., Eiserhardt, W.L. & Svenning, J.C. (2015) Strong effects of variation in taxonomic opinion on diversification analyses. *Methods in Ecology and Evolution.*

Fenton, I.S., Pearson, P.N., Jones, T.D., Farnsworth, A., Lunt, D.J. *et al.* (2016) The impact of cenozoic cooling on assemblage diversity in planktonic foraminifera. *Phil. Trans. R. Soc. B*, **371**, 1691, 20150224.

Fick, S. & Hijmans, R. (2016) Worldclim 2: New 1-km spatial resolution climate surfaces for global land areas. *International Journal of Climatology.*

Findell, K.L., Shevliakova, E., Milly, P. & Stouffer, R.J. (2007) Modeled impact of anthropogenic land cover change on climate. *Journal of Climate*, **20**, 14, 3621–3634.

FitzJohn, R.G. (2010) Quantitative traits and diversification. *Systematic Biology*, syq053.

FitzJohn, R.G., Maddison, W.P. & Otto, S.P. (2009) Estimating trait-dependent speciation and extinction rates from incompletely resolved phylogenies. *Systematic biology*, **58**, 6, 595–611.

Foley, J.A., DeFries, R., Asner, G.P., Barford, C., Bonan, G. *et al.* (2005) Global consequences of land use. *Science*, **309**, 5734, 570–574.

Friis, E.M., Crane, P.R. & Pedersen, K.R., *Early flowers and angiosperm evolution.* Cambridge University Press, 2011.

Frishkoff, L.O., Hadly, E.A. & Daily, G.C. (2015) Thermal niche predicts tolerance to habitat conversion in tropical amphibians and reptiles. *Global change biology*, **21**, 11, 3901–3916.

Frishkoff, L.O., Karp, D.S., Flanders, J.R., Zook, J., Hadly, E.A. *et al.* (2016) Climate change and habitat conversion favour the same species. *Ecology letters*, **19**, 9, 1081–1090.

Fritz, S.A. & Purvis, A. (2010) Selectivity in mammalian extinction risk and threat types: a new measure of phylogenetic signal strength in binary traits. *Conservation Biology*, **24**, 4, 1042–1051.

Fritz, S.A., Schnitzler, J., Eronen, J.T., Hof, C., Böhning-Gaese, K. *et al.* (2013) Diversity in time and space: wanted dead and alive. *Trends in ecology & evolution*, **28**, 9, 509–516.

Frodin, D.G. (2004) History and concepts of big plant genera. *Taxon*, **53**, 3, 753–776.

Gallagher, R., Randall, R. & Leishman, M. (2015) Trait differences between naturalized and invasive plant species independent of residence time and phylogeny. *Conservation Biology*, **29**, 2, 360–369.

Gámez-Virués, S., Perović, D.J., Gossner, M.M., Börschig, C., Blüthgen, N. *et al.* (2015) Landscape simplification filters species traits and drives biotic homogenization. *Nature communications*, **6**.

García-Valdés, R., Svenning, J.C., Zavala, M.A., Purves, D.W. & Araújo, M.B. (2015) Evaluating the combined effects of climate and land-use change on tree species distributions. *Journal of Applied Ecology*, **52**, 4, 902–912.

Gardner, T.A., Barlow, J., Chazdon, R., Ewers, R.M., Harvey, C.A. *et al.* (2009) Prospects for tropical forest biodiversity in a human-modified world. *Ecology Letters*, **12**, 6, 561–582.

Garzón-Orduña, I.J., Silva-Brandão, K.L., Willmott, K.R., Freitas, A.V. & Brower, A.V. (2015) Incompatible ages for clearwing butterflies based on alternative secondary calibrations. *Systematic Biology*, **64**, 5, 752–767.

Gaston, K.J. (1991) Body size and probability of description: the beetle fauna of britain. *Ecological Entomology*, **16**, 4, 505–508.

Gaston, K.J. (1996) Species-range-size distributions: patterns, mechanisms and implications. *Trends in Ecology & Evolution*, **11**, 5, 197–201.

Gaston, K.J., Scoble, M.J. & Crook, A. (1995) Patterns in species description: a case study using the geometridae (lepidoptera). *Biological Journal of the Linnean Society*, **55**, 3, 225–237.

Gelman, A., Su, Y.S., Yajima, M., Hill, J., Pittau, M.G. *et al.* (2009) arm: Data analysis using regression and multilevel/hierarchical models (r package, version 9.01).

Gibson, L., Lee, T.M., Koh, L.P., Brook, B.W., Gardner, T.A. *et al.* (2011) Primary forests are irreplaceable for sustaining tropical biodiversity. *Nature*, **478**, 7369, 378–381.

Gillespie, R. (2004) Community assembly through adaptive radiation in hawaiian spiders. *Science*, **303**, 5656, 356–359.

Gilroy, J.J., Prescott, G.W., Cardenas, J.S., González del Pliego Castañeda, P., Sanchez, A. *et al.* (2015) Minimizing the biodiversity impact of neotropical oil palm development. *Global Change Biology*, **21**, 4, 1531âĂŞ–1540. doi:10.1111/gcb.12696.

Gilroy, J.J., Woodcock, P., Edwards, F.A., Wheeler, C., Baptiste, B.L. *et al.* (2014) Cheap carbon and biodiversity co-benefits from forest regeneration in a hotspot of endemism. *Nature Climate Change*, **4**, 6, 503–507.

Goldberg, E.E., Kohn, J.R., Lande, R., Robertson, K.A., Smith, S.A. *et al.* (2010) Species selection maintains self-incompatibility. *Science*, **330**, 6003, 493–495.

Goldberg, E.E., Lancaster, L.T. & Ree, R.H. (2011) Phylogenetic inference of reciprocal effects between geographic range evolution and diversification. *Systematic Biology*, **60**, 4, 451–465.

Green, R.E., Cornell, S.J., Scharlemann, J.P. & Balmford, A. (2005) Farming and the fate of wild nature. *Science*, **307**, 5709, 550–555.

Gutiérrez-Lemus, D.L., Serrano, V.H. & Ramírez-Pinilla, M.P. (2004) Composición y abundancia de anuros en dos tipos de bosque (natural y cultivado) en la coordillera oriental de Colombia. *Caldasia*, **1**, 245–264.

Haddad, N.M., Brudvig, L.A., Clobert, J., Davies, K.F., Gonzalez, A. *et al.* (2015) Habitat fragmentation and its lasting impact on earthâĂŹs ecosystems. *Science Advances*, **1**, 2, e1500052.

Hallinan, N. (2012) The generalized time variable reconstructed birth–death process. *Journal of theoretical biology*, **300**, 265–276.

Hansen, M., Potapov, P., Moore, R., Hancher, M., Turubanova, S. *et al.* (2013) High-resolution global maps of 21st-century forest cover change. *Science*, **342**, 6160, 850–853.

Harfoot, M., Tittensor, D.P., Newbold, T., McInerny, G., Smith, M.J. *et al.* (2014) Integrated assessment models for ecologists: the present and the future. *Global Ecology and Biogeography*, **23**, 124–143.

Harmon, L.J., Schulte, J.A., Larson, A. & Losos, J.B. (2003) Tempo and mode of evolutionary radiation in iguanian lizards. *Science*, **301**, 5635, 961–964.

Harmon, L.J., Weir, J.T., Brock, C.D., Glor, R.E. & Challenger, W. (2008) Geiger: investigating evolutionary radiations. *Bioinformatics*, **24**, 1, 129–131.

Harrington, M.G., Jackes, B.R., Barrett, M.D., Craven, L.A. & Barrett, R.L. (2012) Phylogenetic revision of backhousieae (myrtaceae): Neogene divergence, a revised circumscription of backhousia and two new species. *Australian Systematic Botany*, **25**, 6, 404–417.

Harvey, P.H., May, R.M. & Nee, S. (1994) Phylogenies without fossils. *Evolution*, 523–529.

Harwood, T.D., Mokany, K. & Paini, D.R. (2014) Microclimate is integral to the modeling of plant responses to macroclimate. *Proceedings of the National Academy of Sciences of the United States of America*, **111**, 13, E1164–5.

Hawkes, J. & Smith, P. (1965) Continental drift: continental drift and the age of angiosperm genera. *Nature*, **207**, 48–50.

Heard, S.B. & Hauser, D.L. (1995) Key evolutionary innovations and their ecological mechanisms. *Historical biology*, **10**, 2, 151–173.

Hermy, M. & Verheyen, K. (2007) Legacies of the past in the present-day forest biodiversity: a review of past land-use effects on forest plant species composition and diversity. *Ecological research*, **22**, 3, 361–371.

Hernández-Ordóñez, O., Urbina-Cardona, N. & Martínez-Ramos, M. (2015) Recovery of amphibian and reptile assemblages during old-field succession of tropical rain forests. *Biotropica*, **47**, 3, 377–388.

Hertweck, K.L., Kinney, M.S., Stuart, S.A., Maurin, O., Mathews, S. *et al.* (2015) Phylogenetics, divergence times and diversification from three genomic partitions in monocots. *Botanical Journal of the Linnean Society*.

Higuera, D. & Wolf, J.H. (2010) Vascular epiphytes in dry oak forests show resilience to anthropogenic disturbance, cordillera oriental, Colombia. *Caldasia*, **32**, 1, 161–174.

Hijmans, R.J. (2015) *raster: Geographic Data Analysis and Modeling*. R package version 2.3-40.

Hijmans, R.J., Cameron, S.E., Parra, J.L., Jones, P.G. & Jarvis, A. (2005) Very high resolution interpolated climate surfaces for global land areas. *International journal of climatology*, **25**, 15, 1965–1978.

Hijmans, R.J., Gavrilenko, T., Stephenson, S., Bamberg, J., Salas, A. *et al.* (2007) Geographical and environmental range expansion through polyploidy in wild potatoes (solanum section petota). *Global Ecology and Biogeography*, **16**, 4, 485–495.

Hobbs, R.J. & Mooney, H.A. (1998) Broadening the extinction debate: population deletions and additions in california and western australia. *Conservation Biology*, **12**, 2, 271–283.

Hodges, S.A. & Arnold, M.L. (1995) Spurring plant diversification: are floral nectar spurs a key innovation? *Proceedings of the Royal Society of London B: Biological Sciences*, **262**, 1365, 343–348.

Höhna, S. (2015) The time-dependent reconstructed evolutionary process with a key-role for mass-extinction events. *Journal of theoretical biology*, **380**, 321–331.

Höhna, S., Landis, M.J., Heath, T.A., Boussau, B., Lartillot, N. *et al.* (2016a) Revbayes: Bayesian phylogenetic inference using graphical models and an interactive model-specification language. *Systematic biology*, syw021.

Höhna, S., May, M.R. & Moore, B.R. (2016b) Tess: Bayesian inference of lineage diversification rates from (incompletely sampled) molecular phylogenies in r. *Bioinformatics*, **35**, 5, 789–791. doi:10.1093/bioinformatics/btv651.

Höhna, S., Stadler, T., Ronquist, F. & Britton, T. (2011) Inferring speciation and extinction rates under different sampling schemes. *Molecular biology and evolution*, **28**, 9, 2577–2589.

Honnay, O., Verheyen, K., Butaye, J., Jacquemyn, H., Bossuyt, B. *et al.* (2002) Possible effects of habitat fragmentation and climate change on the range of forest plant species. *Ecology Letters*, **5**, 4, 525–530.

Hooper, D.U., Adair, E.C., Cardinale, B.J., Byrnes, J.E., Hungate, B.A. *et al.* (2012) A global synthesis reveals biodiversity loss as a major driver of ecosystem change. *Nature*, **486**, 7401, 105–108.

Hudson, L.N., Newbold, T., Contu, S., Hill, S.L., Lysenko, I. *et al.* (2017) The database of the predicts (projecting responses of ecological diversity in changing terrestrial systems) project. *Ecology and Evolution*, **7**, 1, 145–188.

Hudson, L.N., Newbold, T., Contu, S., Hill, S.L.L., Lysenko, I. *et al.* (2014) The predicts database: a global database of how local terrestrial biodiversity responds to human impacts. *Ecology and Evolution*, **4**, 24, 4701–4735. doi:10.1002/ece3.1303.

Huelsenbeck, J.P., Nielsen, R. & Bollback, J.P. (2003) Stochastic mapping of morphological characters. *Systematic Biology*, **52**, 2, 131–158.

Hughes, C. & Eastwood, R. (2006) Island radiation on a continental scale: exceptional rates of plant diversification after uplift of the andes. *Proceedings of the National Academy of Sciences*, **103**, 27, 10334–10339.

Hunter, J.P. & Jernvall, J. (1995) The hypocone as a key innovation in mammalian evolution. *Proceedings of the National Academy of Sciences*, **92**, 23, 10718–10722.

Hurtt, G.C., Chini, L.P., Frolking, S., Betts, R.A., Feddema, J. *et al.* (2011) Harmonization of land-use scenarios for the period 1500–2100: 600 years of global gridded annual land-use transitions, wood harvest, and resulting secondary lands. *Climatic Change*, **109**, 1, 117–161. doi:10.1007/s10584-011-0153-2.

IPCC, *Climate Change 2007: Synthesis Report. Contribution of Working Groups I, II and III to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change.* Geneva, Switzerland, 2007.

IPCC, *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change.* Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2013.

Isaac, N.J., AGAPOW, P.M., Harvey, P.H. & Purvis, A. (2003) Phylogenetically nested comparisons for testing correlates of species richness: a simulation study of continuous variables. *Evolution*, **57**, 1, 18–26.

Isaacs, P. & Urbina-Cardona, J.N. (2011) Anthropogenic disturbance and edge effects on anuran assemblages inhabiting cloud forest fragments in colombia. *Natureza a Conservação: Brazilian Journal of Nature Conservation*, **9**, 1, 39–46.

Isbell, F., Calcagno, V., Hector, A., Connolly, J., Harpole, W.S. *et al.* (2011) High plant diversity is needed to maintain ecosystem services. *Nature*, **477**, 7363, 199.

Jablonski, D., Huang, S., Roy, K. & Valentine, J.W. (2017) Shaping the latitudinal diversity gradient: new perspectives from a synthesis of paleobiology and biogegraphy. *The American Naturalist*, **189**, 1–12.

Jantz, S.M., Barker, B., Brooks, T.M., Chini, L.P., Huang, Q. *et al.* (2015) Future habitat loss and extinctions driven by land-use change in biodiversity hotspots under four scenarios of climate-change mitigation. *Conservation Biology*, **29**, 4, 1122–1131.

Jetz, W., Thomas, G., Joy, J., Hartmann, K. & Mooers, A. (2012) The global diversity of birds in space and time. *Nature*, **491**, 7424, 444–448.

Jetz, W., Wilcove, D.S. & Dobson, A.P. (2007) Projected impacts of climate and land-use change on the global diversity of birds. *PLoS Biol*, **5**, 6, e157.

Kattge, J., Diaz, S., Lavorel, S., Prentice, I.C., Leadley, P. *et al.* (2011) Try–a global database of plant traits. *Global change biology*, **17**, 9, 2905–2935.

Kearney, M. & Porter, W. (2009) Mechanistic niche modelling: combining physiological and spatial data to predict species' ranges. *Ecology letters*, **12**, 4, 334–350.

Klein Goldewijk, K., Beusen, A., Van Drecht, G. & De Vos, M. (2011) The HYDE 3.1 spatially explicit database of human-induced global land-use change over the past 12,000 years. *Global Ecology and Biogeography*, **20**, 1, 73–86.

Knapp, S. (2013) A revision of the dulcamaroid clade of solanum l.(solanaceae). *PhytoKeys*, , 22, 1.

Knapp, S. & Vorontsova, M.S. (2016) A revision of the "african non-spiny" clade of solanum l.(solanum sections afrosolanum bitter, benderianum bitter, lemurisolanum bitter, lyciosolanum bitter, macronesiotes bitter, and quadrangulare bitter: Solanaceae). *PhytoKeys*, **66**, 1.

Köster, N., Kreft, H., Nieder, J. & Barthlott, W. (2013) Range size and climatic niche correlate with the vulnerability of epiphytes to human land use in the tropics. *Journal of Biogeography*, **40**, 5, 963–976.

Kuhn, T.S., Mooers, A.Ø. & Thomas, G.H. (2011) A simple polytomy resolver for dated phylogenies. *Methods in Ecology and Evolution*, **2**, 5, 427–436.

Landis, M.J., Matzke, N.J., Moore, B.R. & Huelsenbeck, J.P. (2013) Bayesian analysis of biogeography when the number of areas is large. *Systematic biology*, syt040.

Lawrence, D. & Vandecar, K. (2015) Effects of tropical deforestation on climate and agriculture. *Nature Climate Change*, **5**, 1, 27–36.

Lawton, J.H., Bignell, D., Bolton, B., Bloemers, G., Eggleton, P. *et al.* (1998) Biodiversity inventories, indicator taxa and effects of habitat modification in tropical forest. *Nature*, **391**, 6662, 72–76.

Lawver, L.A. & Gahagan, L.M. (2003) Evolution of cenozoic seaways in the circum-antarctic region. *Palaeogeography, Palaeoclimatology, Palaeoecology*, **198**, 1, 11–37.

Levin, R.A., Myers, N.R. & Bohs, L. (2006) Phylogenetic relationships among the âĂIJspiny solanumsâĂİ(solanum subgenus leptostemonum, solanaceae). *American Journal of Botany*, **93**, 1, 157–169.

Levin, R.A., Watson, K. & Bohs, L. (2005) A four-gene study of evolutionary relationships in solanum section acanthophora. *American Journal of Botany*, **92**, 4, 603–612.

Lidgard, S. & Crane, P.R. (1988) Quantitative analyses of the early angiosperm radiation. *Nature*, **331**, 6154, 344–346.

Lieberman, B.S. (2005) Geobiology and paleobiogeography: tracking the coevolution of the earth and its biota. *Palaeogeography, Palaeoclimatology, Palaeoecology*, **219**, 1, 23–33.

Linder, H.P. (2014) The evolution of african plant diversity. *Frontiers in Ecology and Evolution*, **2**, 38.

Liow, L.H., Quental, T.B. & Marshall, C.R. (2010) When can decreasing diversification rates be detected with molecular phylogenies and the fossil record? *Systematic Biology*, syq052.

Liu, Y.Y., Dijk, A.I., McCabe, M.F., Evans, J.P. & Jeu, R.A. (2013) Global vegetation biomass change (1988–2008) and attribution to environmental and human drivers. *Global ecology and biogeography*, **22**, 6, 692–705.

Lomolino, M.V., Riddle, B.R., Whittaker, R.J. & Brown, J.H., *Biogeography*. Sunderland, 2010.

Longino, J.T., Coddington, J. & Colwell, R.K. (2002) The ant fauna of a tropical rain forest: estimating species richness three different ways. *Ecology*, **83**, 3, 689–702.

López-Quintero, C.A., Straatsma, G., Franco-Molano, A.E. & Boekhout, T. (2012) Macrofungal diversity in Colombian amazon forests varies with regions and regimes of disturbance. *Biodiversity and Conservation*, **21**, 9, 2221–2243.

Losos, J.B. (2010) Adaptive radiation, ecological opportunity, and evolutionary determinism. *The American Naturalist*, **175**, 6, 623–639.

Losos, J.B. & Schluter, D. (2000) Analysis of an evolutionary species-area relationship. *Nature*, **408**, 6814, 847.

Lozada, T., de Koning, G., Kessler, M., Klein, A.M. & Tscharntke, T. (2008) Geographical range size of tropical plants influences their response to anthropogenic activities. *Diversity and Distributions*, **14**, 1, 59–68.

Luck, G.W. (2007) A review of the relationships between human population density and biodiversity. *Biological Reviews*, **82**, 4, 607–645.

Luo, Z.X. (2007) Transformation and diversification in early mammal evolution. *Nature*, **450**, 7172, 1011–1019.

Mac Nally, R. (2007) Use of the abundance spectrum and relative-abundance distributions to analyze assemblage change in massively altered landscapes. *The American Naturalist*, **170**, 3, 319–330.

Maddison, W. (2014) Papers using diversitree.

Maddison, W.P. & FitzJohn, R.G. (2015) The unsolved challenge to phylogenetic correlation tests for categorical characters. *Systematic biology*, **64**, 1, 127–136.

Maddison, W.P., Midford, P.E. & Otto, S.P. (2007) Estimating a binary character's effect on speciation and extinction. *Systematic Biology*, **56**, 5, 701–710.

Madriñán, S., Cortés, A.J. & Richardson, J.E. (2013) Páramo is the world's fastest evolving and coolest biodiversity hotspot. *Frontiers in Genetics*, **4**, 192. doi:10.3389/fgene.2013.00192.

Magallón, S., Gómez-Acevedo, S., Sánchez-Reyes, L.L. & Hernández-Hernández, T. (2015) A metacalibrated time-tree documents the early rise of flowering plant phylogenetic diversity. *New Phytologist*.

Magallón, S. & Sanderson, M.J. (2001) Absolute diversification rates in angiosperm clades. *Evolution*, **55**, 9, 1762–1780.

Magurran, A.E., *Measuring biological diversity*. Blackwell, 2004.

Magurran, A.E. & Henderson, P.A. (2010) Temporal turnover and the maintenance of diversity in ecological assemblages. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **365**, 1558, 3611–3620.

Martin, P.A., Newton, A.C. & Bullock, J.M. (2013) Carbon pools recover more quickly than plant biodiversity in tropical secondary forests. *Proceedings of the Royal Society of London B: Biological Sciences*, **280**, 1773, 20132236.

Martine, C.T., Vanderpool, D., Anderson, G.J. & Les, D.H. (2006) Phylogenetic relationships of andromonoecious and dioecious australian species of solanum subgenus leptostemonum section melongena: inferences from its sequence data. *Systematic botany*, **31**, 2, 410–420.

Matzke, N.J. (2012) Founder-event speciation in biogeobears package dramatically improves likelihoods and alters parameter inference in dispersal–extinction–cladogenesis (dec) analyses. *Front. Biogeogr*, **4**, 210.

Matzke, N.J. (2014) Model selection in historical biogeography reveals that founder-event speciation is a crucial process in island clades. *Systematic biology*, syu056.

Matzke, N.J. (2016) Stochastic mapping under biogeographical models. PhyloWiki BioGeoBEARS Available at `http://phylo.wikidot.com/biogeobears#stochastic_mapping`.

May, M.R. & Moore, B.R. (2016) How well can we detect lineage-specific diversification-rate shifts? a simulation study of sequential aic methods. *Systematic biology*, syw026.

McCune, J.L. & Vellend, M. (2013) Gains in native species promote biotic homogenization over four decades in a human-dominated landscape. *Journal of Ecology*, **101**, 6, 1542–1551.

McKinney, M.L. & Lockwood, J.L. (1999) Biotic homogenization: a few winners replacing many losers in the next mass extinction. *Trends in ecology & evolution*, **14**, 11, 450–453.

McPeek, M.A. & Brown, J.M. (2007) Clade age and not diversification rate explains species richness among animal taxa. *The American Naturalist*, **169**, 4, E97–E106.

Meseguer, A.S., Lobo, J.M., Ree, R., Beerling, D.J. & Sanmartín, I. (2014) Integrating fossils, phylogenies, and niche models into biogeography to reveal ancient evolutionary history: the case of hypericum (hypericaceae). *Systematic biology*, **64**, 2, 215–232.

Meyer, C., Kreft, H., Guralnick, R. & Jetz, W. (2015) Global priorities for an effective information basis of biodiversity distributions. *Nature communications*, **6**, 8221.

Millennium Ecosystem Assessment, *Ecosystems and human well-being*. Island Press Washington, DC, 2005.

Miller, M.A., Pfeiffer, W. & Schwartz, T., Creating the cipres science gateway for inference of large phylogenetic trees. In *Gateway Computing Environments Workshop (GCE), 2010*. IEEE, 2010, 1–8.

Mitchell, J.S. & Rabosky, D.L. (2016) Bayesian model selection with bamm: effects of the model prior on the inferred number of diversification shifts. *Methods in Ecology and Evolution*.

Mitter, C., Farrell, B. & Wiegmann, B. (1988) The phylogenetic study of adaptive zones: has phytophagy promoted insect diversification? *American Naturalist*, 107–128.

Mittermeier, R., Mittermeier, C., Brooks, T., Pilgrim, J., Konstant, W. *et al.* (2003) Wilderness and biodiversity conservation. *Proceedings of the National Academy of Sciences*, **100**, 18, 10309âĂŞ–10313.

Moles, A.T., Warton, D.I., Warman, L., Swenson, N.G., Laffan, S.W. *et al.* (2009) Global patterns in plant height. *Journal of Ecology*, **97**, 5, 923–932.

Moore, B.R. & Donoghue, M.J. (2007) Correlates of diversification in the plant clade dipsacales: geographic movement and evolutionary innovations. *the american naturalist*, **170**, S2, S28–S55.

Moore, B.R., Höhna, S., May, M.R., Rannala, B. & Huelsenbeck, J.P. (2016) Critically evaluating the theory and performance of bayesian analysis of macroevolutionary mixtures. *Proceedings of the National Academy of Sciences*, **113**, 34, 9569–9574.

Moore, B.R. & May, M.R. (2014) Phylogenetic approaches for exploring differential rates of diversification. Bodega Workshop.

Moore, M.J., Bell, C.D., Soltis, P.S. & Soltis, D.E. (2007) Using plastid genome-scale data to resolve enigmatic relationships among basal angiosperms. *Proceedings of the National Academy of Sciences*, **104**, 49, 19363–19368.

Morin, X. & Lechowicz, M.J. (2013) Niche breadth and range area in north american trees. *Ecography*, **36**, 3, 300–312.

Morlon, H. (2014) Phylogenetic approaches for studying diversification. *Ecology Letters*, n/a–n/a. doi:10.1111/ele.12251.

Morlon, H., Parsons, T.L. & Plotkin, J.B. (2011) Reconciling molecular phylogenies with the fossil record. *Proceedings of the National Academy of Sciences*, **108**, 39, 16327–16332.

Moss, R.H., Edmonds, J.A., Hibbard, K.A., Manning, M.R., Rose, S.K. *et al.* (2010) The next generation of scenarios for climate change research and assessment. *Nature*, **463**, 7282, 747–756.

Murtaugh, P.A. (2002) On rejection rates of paired intervention analysis. *Ecology*, **83**, 6, 1752–1761.

Myers, N., Mittermeier, R., Mittermeier, C., da Fonseca, G. & Kent, J. (2000) Biodiversity hotspots for conservation priorities. *Nature*, **403**, 6772, 853–858.

Navarro, I.L., Roman, A.K., Gómez, F.H. & Pérez, A. (2011) Variación estacional en escarabajos coprófagos (Coleoptera: Scarabaeidae: Scarabaeinae) de la serranía de coraza, Sucre (Colombia). *Revista Colombiana de Ciencia Animal*, **3**, 1, 102–110.

Nee, S., Holmes, E.C., May, R.M. & Harvey, P.H. (1994a) Extinction rates can be estimated from molecular phylogenies. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **344**, 1307, 77–82.

Nee, S., May, R.M. & Harvey, P.H. (1994b) The reconstructed evolutionary process. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **344**, 1309, 305–311.

Nee, S., Mooers, A.O. & Harvey, P.H. (1992) Tempo and mode of evolution revealed from molecular phylogenies. *Proceedings of the National Academy of Sciences*, **89**, 17, 8322–8326.

Newbold, T., Hudson, L.N., Arnell, A.P., Contu, S., De Palma, A. *et al.* (2016a) Has land use pushed terrestrial biodiversity beyond the planetary boundary? a global assessment. *Science*, **353**, 6296, 288–291.

Newbold, T., Hudson, L.N., Contu, S., Hill, S.L.L., Beck, J. *et al.* (in Review) Widespread winners and narrowly distributed losers: land use homogenizes biodiversity worldwide.

Newbold, T., Hudson, L.N., Hill, S.L., Contu, S., Gray, C.L. *et al.* (2016b) Global patterns of terrestrial assemblage turnover within and among land uses. *Ecography*, **10.1111/ecog.01932**. doi:10.1111/ecog.01932.

Newbold, T., Hudson, L.N., Hill, S.L.L., Contu, S., Lysenko, I. *et al.* (2015) Global effects of land use on local terrestrial biodiversity. *Nature*, **520**, 7545, 45–50.

Newbold, T., Hudson, L.N., Phillips, H.R.P., Hill, S.L.L., Contu, S. *et al.* (2014) A global model of the response of tropical and sub-tropical forest biodiversity to anthropogenic pressures. *Proceedings of the Royal Society B: Biological Sciences*, **281**, 1792, 20141371.

Newbold, T., Scharlemann, J.P., Butchart, S.H., Şekercioğlu, Ç.H., Alkemade, R. *et al.* (2013) Ecological traits affect the response of tropical forest bird species to land-use intensity. *Proceedings of the Royal Society B: Biological Sciences*, **280**, 1750. doi: 10.1098/rspb.2012.2131.

Niklas, K.J., Tiffney, B.H. & Knoll, A.H. (1983) Patterns in vascular land plant diversification.

Nogués-Bravo, D., Pulido, F., Araújo, M.B., Diniz-Filho, J.A.F., García-Valdés, R. *et al.* (2014) Phenotypic correlates of potential range size and range filling in european trees. *Perspectives in Plant Ecology, Evolution and Systematics*, **16**, 5, 219–227.

Norden, N., Chazdon, R.L., Chao, A., Jiang, Y.H. & Vílchez-Alvarado, B. (2009) Resilience of tropical rain forests: tree community reassembly in secondary forests. *Ecology Letters*, **12**, 5, 385–394.

Noriega, J., Realpe, E. & Fagua, G. (2007) Diversidad de escarabajos coprófagos (Coleoptera: Scarabaeidae) en un bosque de galería con tres estadios de alteración. *Universitas Scientiarum*, **12**, 51–63.

Noriega, J.A., Palacio, J.M., Monroy-G, J.D. & Valencia, E. (2012) Estructura de un ensamblaje de escarabajos coprófagos (Coleoptera:scarabaeinae) en tres sitios con diferente uso del suelo en Antioquia, Colombia. *Actualidades Biológicas*, **34**, 96, 43–54.

Oksanen, J., Blanchet, F.G., Friendly, M., Kindt, R., Legendre, P. *et al.* (2016) *vegan: Community Ecology Package.* R package version 2.4-1.

Olden, J.D., Poff, N.L., Douglas, M.R., Douglas, M.E. & Fausch, K.D. (2004) Ecological and evolutionary consequences of biotic homogenization. *Trends in ecology & evolution*, **19**, 1, 18–24.

Olmstead, R.G. (2013) Phylogeny and biogeography in solanaceae, verbenaceae and bignoniaceae: a comparison of continental and intercontinental diversification patterns. *Botanical Journal of the Linnean Society*, **171**, 1, 80–102.

Olmstead, R.G., Bohs, L., Migid, H.A., Santiago-Valentin, E., Garcia, V.F. *et al.* (2008) A molecular phylogeny of the solanaceae. *Taxon*, **57**, 4, 1159–1181.

Olmstead, R.G. & Palmer, J.D. (1997) Implications for the phylogeny, classification, and biogeography of solanum from cpdna restriction site variation. *Systematic Botany*, 19–29.

Olson, D. & Dinerstein, E. (2002) The global 200: Priority ecoregions for global conservation. *Annals of the Missouri Botanical Garden*, **89**, 199–224.

Olson, D.M., Dinerstein, E., Wikramanayake, E.D., Burgess, N.D., Powell, G.V. *et al.* (2001) Terrestrial ecoregions of the world: A new map of life on earth a new global map of terrestrial ecoregions provides an innovative tool for conserving biodiversity. *BioScience*, **51**, 11, 933–938.

Orme, C.D.L., Davies, R.G., Burgess, M., Eigenbrod, F., Pickup, N. *et al.* (2005) Global hotspots of species richness are not congruent with endemism or threat. *Nature*, **436**, 7053, 1016–1019.

Otavo, S.E., Parrado-Rosselli, Á. & Noriega, J.A. (2013) Superfamilia Scarabaeoidea (Insecta: Coleoptera) como elemento bioindicador de perturbación antropogénica en un parque nacional amazónico. *Revista de Biología Tropical*, **61**, 2, 735–752.

Owens, I.P. & Bennett, P.M. (2000) Ecological basis of extinction risk in birds: habitat loss versus human persecution and introduced predators. *Proceedings of the National Academy of Sciences*, **97**, 22, 12144–12148.

Pagel, M. (1999) The maximum likelihood approach to reconstructing ancestral character states of discrete characters on phylogenies. *Systematic biology*, **48**, 3, 612–622.

Palmer, J.D. & Zamir, D. (1982) Chloroplast dna evolution and phylogenetic relationships in lycopersicon. *Proceedings of the National Academy of Sciences*, **79**, 16, 5006–5010.

Parra, A. & Nates-Parra, G. (2007) Variación de la comunidad de abejas de las orquídeas (Hymenoptera: Apidae) en tres ambientes perturbados del piedemonte llanero colombiano. *Revista de Biología Tropical*, **55**, 3-4, 931–941.

Patterson, B. (1994) Accumulating knowledge on the dimensions of biodiversity: systematic perspectives on neotropical mammals. *Biodiversity Letters*, 79–86.

Pebesma, E.J. & Bivand, R.S. (2005) Classes and methods for spatial data in R. *R News*, **5**, 2, 9–13.

Pecl, G.T., Araújo, M.B., Bell, J.D., Blanchard, J., Bonebrake, T.C. *et al.* (2017) Biodiversity redistribution under climate change: Impacts on ecosystems and human well-being. *Science*, **355**, 6332, eaai9214.

Peralta, I.E. & Spooner, D.M. (2001) Granule-bound starch synthase (gbssi) gene phylogeny of wild tomatoes (solanum l. section lycopersicon [mill.] wettst. subsection lycopersicon). *American Journal of Botany*, **88**, 10, 1888–1902.

Pereira, H.M., Leadley, P.W., Proença, V., Alkemade, R., Scharlemann, J.P.W. *et al.* (2010) Scenarios for global biodiversity in the 21st century. *Science*, **330**, 6010, 1496–1501. doi:10.1126/science.1196624.

Perring, M.P., De Frenne, P., Baeten, L., Maes, S.L., Depauw, L. *et al.* (2016) Global environmental change effects on ecosystems: the importance of land-use legacies. *Global change biology*, **22**, 4, 1361–1371.

Phalan, B., Onial, M., Balmford, A. & Green, R.E. (2011) Reconciling food production and biodiversity conservation: land sharing and land sparing compared. *Science*, **333**, 6047, 1289–1291.

Phillimore, A.B., Freckleton, R.P., Orme, C.D.L. & Owens, I.P. (2006) Ecology predicts large-scale patterns of phylogenetic diversification in birds. *The American Naturalist*, **168**, 2, 220–229.

Phillimore, A.B. & Price, T.D. (2008) Density-dependent cladogenesis in birds. *PLoS Biol*, **6**, 3, e71.

Phillips, H.R.P., Halley, J.M., Urbina-Cordona, J.N. & Purvis, A. (2017) The effect of fragment area on site-level biodiversity. *Ecography*, n/a–n/a. doi:10.1111/ecog.02956.

Pielke, R.A., Pitman, A., Niyogi, D., Mahmood, R., McAlpine, C. *et al.* (2011) Land use/land cover changes and climate: modeling analysis and observational evidence. *Wiley Interdisciplinary Reviews: Climate Change*, **2**, 6, 828–850.

Pimm, S.L., Jenkins, C.N., Abell, R., Brooks, T.M., Gittleman, J.L. *et al.* (2014) The biodiversity of species and their rates of extinction, distribution, and protection. *Science*, **344**, 6187, 1246752.

Plummer, M., Best, N., Cowles, K. & Vines, K. (2006) Coda: Convergence diagnosis and output analysis for mcmc. *R News*, **6**, 1, 7–11.

Poczai, P., Hyvönen, J. & Symon, D.E. (2011) Phylogeny of kangaroo apples (solanum subg. archaesolanum, solanaceae). *Molecular biology reports*, **38**, 8, 5243–5259.

Poorter, L., Bongers, F., Aide, T.M., Almeyda Zambrano, A.M., Balvanera, P. *et al.* (2016) Biomass resilience of neotropical secondary forests. *Nature*, **530**, 7589, 211–214.

Poveda, K., Martínez, E., Kersch-Becker, M.F., Bonilla, M.A. & Tscharntke, T. (2012) Landscape simplification and altitude affect biodiversity, herbivory and andean potato yield. *Journal of Applied Ecology*, **49**, 2, 513–522.

Pulliam, H.R. & Danielson, B.J. (1991) Sources, sinks, and habitat selection: a landscape perspective on population dynamics. *American Naturalist*, **137**, S50–S66.

Purvis, A., *New uses for new phylogenies*, vol. 153, chap. Using interspecies phylogenies to test macroevolutionary hypotheses. Oxford University Press, Oxford, 1996a.

Purvis, A. (1996b) Using interspecies phylogenies to test macroevolutionary hypotheses. *New uses for new phylogenies. Oxford University Press, Oxford*, 153–168.

Purvis, A. (2008) Phylogenetic approaches to the study of extinction. *Annual Review of Ecology, Evolution, and Systematics*, 301–319.

Purvis, A., Fritz, S.A., Rodríguez, J., Harvey, P.H. & Grenyer, R. (2011) The shape of mammalian phylogeny: patterns, processes and scales. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, **366**, 1577, 2462–2477.

Purvis, A., Gittleman, J.L., Cowlishaw, G. & Mace, G.M. (2000) Predicting extinction risk in declining species. *Proceedings of the Royal Society of London B: Biological Sciences*, **267**, 1456, 1947–1952.

Purvis, A., Gittleman, J.L. & Luh, H.K. (1994) Truth or consequences: effects of phylogenetic accuracy on two comparative methods. *Journal of theoretical biology*, **167**, 3, 293–300.

Purvis, A., Nee, S. & Harvey, P.H. (1995) Macroevolutionary inferences from primate phylogeny. *Proceedings of the Royal Society of London B: Biological Sciences*, **260**, 1359, 329–333.

Purvis, A., Newbold, T., Palma, A.D., Sara Contu1, S.L.L.H., Sanchez-Ortiz, K. *et al.* (in press) Modelling and projecting the response of local terrestrial biodiversity worldwide to land use and related pressures: the predicts project. *Advances in Ecological Research*.

Pybus, O.G. & Harvey, P.H. (2000) Testing macro–evolutionary models using incomplete molecular phylogenies. *Proceedings of the Royal Society of London B: Biological Sciences*, **267**, 1459, 2267–2272.

Quental, T.B. & Marshall, C.R. (2010) Diversity dynamics: molecular phylogenies need the fossil record. *Trends in Ecology & Evolution*, **25**, 8, 434–441.

R Core Team (2013) R: A language and environment for statistical computing. R foundation for statistical computing, Vienna, Austria.

Rabosky, A.D., Cox, C., Rabosky, D., Title, P., Holmes, I. *et al.* (2016) Coral snakes predict the evolution of mimicry across new world snakes. *Nature Communications*, **7**, 11484.

Rabosky, D. (2006) Likelihood methods for detecting temporal shifts in diversification rates. *Evolution*, **60**, 6, 1152–1164.

Rabosky, D.L. (2010) Extinction rates should not be estimated from molecular phylogenies. *Evolution*, **64**, 6, 1816–1824.

Rabosky, D.L. (2014) Automatic detection of key innovations, rate shifts, and diversity-dependence on phylogenetic trees. *PloS one*, **9**, 2, e89543.

Rabosky, D.L. (2015a) Bamm 2.5.0 documentation.

Rabosky, D.L. (2015b) No substitute for real data: A cautionary note on the use of phylogenies from birth–death polytomy resolvers for downstream comparative analyses. *Evolution*, **69**, 12, 3207–3216.

Rabosky, D.L., Donnellan, S.C., Grundler, M. & Lovette, I.J. (2014a) Analysis and visualization of complex macroevolutionary dynamics: an example from australian scincid lizards. *Systematic biology*, **63**, 4, 610–627.

Rabosky, D.L., Donnellan, S.C., Talaba, A.L. & Lovette, I.J. (2007) Exceptional among-lineage variation in diversification rates during the radiation of australia's most diverse vertebrate clade. *Proceedings of the Royal Society of London B: Biological Sciences*, **274**, 1628, 2915–2923.

Rabosky, D.L. & Glor, R.E. (2010) Equilibrium speciation dynamics in a model adaptive radiation of island lizards. *Proceedings of the National Academy of Sciences*, **107**, 51, 22178–22183.

Rabosky, D.L. & Goldberg, E.E. (2015) Model inadequacy and mistaken inferences of trait-dependent speciation. *Systematic biology*, **64**, 2, 340–355.

Rabosky, D.L. & Goldberg, E.E. (2017) Fisse: A simple nonparametric test for the effects of a binary character on lineage diversification rates. *Evolution*, **71**, 6, 1432–1442.

Rabosky, D.L., Grundler, M., Anderson, C., Shi, J.J., Brown, J.W. *et al.* (2014b) Bammtools: an r package for the analysis of evolutionary dynamics on phylogenetic trees. *Methods in Ecology and Evolution*.

Rabosky, D.L. & Huang, H. (2015) A robust semi-parametric test for detecting trait-dependent diversification. *Systematic biology*, syv066.

Rabosky, D.L. & Lovette, I.J. (2008) Density-dependent diversification in north american wood warblers. *Proceedings of the Royal Society of London B: Biological Sciences*, **275**, 1649, 2363–2371.

Rabosky, D.L., Mitchell, J.S. & Chang, J. (2017) Is bamm flawed? theoretical and practical concerns in the analysis of multi-rate diversification models. *Systematic biology*. doi:https://doi.org/10.1093/sysbio/syx037.

Rabosky, D.L., Santini, F., Eastman, J., Smith, S.A., Sidlauskas, B. *et al.* (2013) Rates of speciation and morphological evolution are correlated across the largest vertebrate radiation. *Nature communications*, **4**.

Rambaut, A., Suchard, M., Xie, D. & Drummond, A. (2009) Mcmc trace analysis tool. version v1. 6.0. *Institute of Evolutionary Biology, University of Edinburgh. Available: a. rambaut@ ed. ac. uk. Accessed Jan*, **16**, 2011.

Raup, D.M., Gould, S.J., Schopf, T.J. & Simberloff, D.S. (1973) Stochastic models of phylogeny and the evolution of diversity. *The Journal of Geology*, **81**, 5, 525–542.

Reddy, S. (2014) What's missing from avian global diversification analyses? *Molecular phylogenetics and evolution*, **77**, 159–165.

Ree, R.H., Moore, B.R., Webb, C.O., Donoghue, M.J. & Crandall, K. (2005) A likelihood framework for inferring the evolution of geographic range on phylogenetic trees. *Evolution*, **59**, 11, 2299–2311.

Ree, R.H. & Smith, S.A. (2008) Maximum likelihood inference of geographic range evolution by dispersal, local extinction, and cladogenesis. *Systematic Biology*, **57**, 1, 4–14.

Rey-Velasco, J.C. & Miranda-Esquivel, D.R. (2013) Habitat modification in andean forest: the response of ground beetles (coleoptera = carabidae) on the northeastern colombian andes. Thesis in the Biology Department, Universidad Industrial de Santander.

Ricklefs, R.E. (2006) Evolutionary diversification and the origin of the diversity–environment relationship. *Ecology*, **87**, sp7.

Ricklefs, R.E. (2007) Estimating diversification rates from phylogenetic information. *Trends in Ecology & Evolution*, **22**, 11, 601–610.

Ricklefs, R.E., Latham, R.E. & Qian, H. (1999) Global patterns of tree species richness in moist forests: distinguishing ecological influences and historical contingency. *Oikos*, 369–373.

Ricklefs, R.E. & Renner, S.S. (1994) Species richness within families of flowering plants. *Evolution*, **48**, 5, 1619–1636.

Rodriguez, F., Wu, F., Ané, C., Tanksley, S. & Spooner, D.M. (2009) Do potatoes and tomatoes have a single evolutionary history, and what proportion of the genome supports this history? *BMC Evolutionary Biology*, **9**, 1, 1.

Ronquist, F. (1997) Dispersal-vicariance analysis: a new approach to the quantification of historical biogeography. *Systematic Biology*, **46**, 1, 195–203.

Ronquist, F. & Huelsenbeck, J.P. (2003) Mrbayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics*, **19**, 12, 1572–1574.

Ronquist, F. & Sanmartín, I. (2011) Phylogenetic methods in biogeography. *Annual Review of Ecology, Evolution, and Systematics*, **42**, 1, 441.

Rooney, T.P., Olden, J.D., Leach, M.K. & Rogers, D.A. (2007) Biotic homogenization and conservation prioritization. *Biological Conservation*, **134**, 3, 447–450.

Rosselli, L. & Stiles, F.G. (2012a) Local and landscape environmental factors are important for the conservation of endangered wetland birds in a high andean plateau. *Waterbirds*, **35**, 3, 453–469.

Rosselli, L. & Stiles, F.G. (2012b) Wetland habitats of the sabana de Bogotá andean highland plateau and their birds. *Aquatic Conservation: Marine and Freshwater Ecosystems*, **22**, 3, 303–317.

Sala, O.E., Chapin III, S.F., Armesto, J.J., Berlow, E., Bloomfield, J. *et al.* (2000) Global biodiversity scenarios for the year 2100. *Science*, **287**, 5459, 1770–1774.

Särkinen, T., Barboza, G.E. & Knapp, S. (2015) True black nightshades: Phylogeny and delimitation of the morelloid clade of solanum. *Taxon*, **64**, 5, 945–958.

Särkinen, T., Bohs, L., Olmstead, R.G. & Knapp, S. (2013) A phylogenetic framework for evolutionary study of the nightshades (Solanaceae): a dated 1000-tip tree. *BMC evolutionary biology*, **13**, 1, 214. doi:10.1186/1471-2148-13-214.

Scheffers, B.R., Edwards, D.P., Diesmos, A., Williams, S.E. & Evans, T.A. (2014) Microhabitats reduce animal's exposure to climate extremes. *Global Change Biology*, **20**, 2, 495–503.

Schluter, D., *The ecology of adaptive radiation.* Oxford University Press, 2000.

Schluter, D. (2016) Speciation, ecological opportunity, and latitude: (american society of naturalists address). *The American Naturalist*, **187**, 1, 1–18.

Schluter, D. & Pennell, M.W. (2017) Speciation gradients and the distribution of biodiversity. *Nature*, **546**, 7656, 48–55.

Schrodt, F., Kattge, J., Shan, H., Fazayeli, F., Joswig, J. *et al.* (2015) Bhpmf – a hierarchical bayesian approach to gap-filling and trait prediction for macroecology and functional biogeography. *Global Ecology and Biogeography*, **24**, 12, 1510–1521.

Scotland, R.W. & Sanderson, M.J. (2004) The significance of few versus many in the tree of life. *Science*, **303**, 5658, 643–643.

Senior, R.A., Hill, J.K., del Pliego, P.G., Goode, L.K. & Edwards, D.P. (2017) A pantropical analysis of the impacts of forest degradation and conversion on local temperature. *Ecology and Evolution.*

Shi, J.J. & Rabosky, D.L. (2015) Speciation dynamics during the global radiation of extant bats. *Evolution*, **69**, 6, 1528–1545.

Silvestro, D., Cascales-Miñana, B., Bacon, C.D. & Antonelli, A. (2015) Revisiting the origin and diversification of vascular plants through a comprehensive bayesian analysis of the fossil record. *New Phytologist.*

Silvestro, D., Schnitzler, J. & Zizka, G. (2011) A bayesian framework to estimate diversification rates and their variation through time and space. *BMC evolutionary biology*, **11**, 1, 311.

Simões, M., Breitkreuz, L., Alvarado, M., Baca, S., Cooper, J. *et al.* (2016) The evolving theory of evolutionary radiations. *Trends in ecology & evolution*, **31**, 1, 27–34.

Simpson, G.G., *Major features of evolution.* Columbia University Press: New York, 1955.

Slatyer, R.A., Hirst, M. & Sexton, J.P. (2013) Niche breadth predicts geographical range size: a general ecological pattern. *Ecology letters*, **16**, 8, 1104–1114.

Slowinski, J.B. & Guyer, C. (1993) Testing whether certain traits have caused amplified diversification: an improved method based on a model of random speciation and extinction. *American Naturalist*, 1019–1024.

Smith-Pardo, A. & Gonzales, V.H. (2007) Diversidad de abejas (Himenoptera: Apoidea) en estados sucesionales del bosque humedo tropical. *Acta Biológica Colombiana*, **12**, 1, 43–56.

Solow, A.R., Mound, L.A. & Gaston, K.J. (1995) Estimating the rate of synonymy. *Systematic Biology*, **44**, 1, 93–96.

Soltis, P.S. & Soltis, D.E. (2004) The origin and diversification of angiosperms. *American Journal of Botany*, **91**, 10, 1614–1626.

South, A. (2011) rworldmap: A new r package for mapping global data. *The R Journal*, **3**, 1, 35–43.

Spooner, D.M., Anderson, G.J. & Jansen, R.K. (1993) Chloroplast dna evidence for the interrelationships of tomtoes, potatoes, and pepinos (solanaceae). *American Journal of Botany*, 676–688.

Spooner, D.M., Berg, v.d.R., Rodrigues, A., Bamberg, J., Hijmans, R.J. *et al.* (2004) Wild potatoes (solanum section petota; solanaceae) of north and central america. *The American Society of Plant Taxonomists (Systematic botany monographs 68).*

Spooner, D.M., Ghislain, M., Simon, R., Jansky, S.H. & Gavrilenko, T. (2014) Systematics, diversity, genetics, and evolution of wild and cultivated potatoes. *The Botanical Review*, **80**, 4, 283–383.

Spooner, D.M., Peralta, I.E. & Knapp, S. (2005) Comparison of aflps with other markers for phylogenetic inference in wild tomatoes [solanum l. section lycopersicon (mill.) wettst.]. *Taxon*, 43–61.

Srivastava, D.S. & Vellend, M. (2005) Biodiversity-ecosystem function research: is it relevant to conservation? *Annual Review of Ecology, Evolution, and Systematics*, **36**, 267–294.

Stadler, T. (2011) Mammalian phylogeny reveals recent diversification rate shifts. *Proceedings of the National Academy of Sciences*, **108**, 15, 6187–6192.

Stadler, T. (2013) Recovering speciation and extinction dynamics based on phylogenies. *Journal of evolutionary biology*, **26**, 6, 1203–1219.

Stadler, T. (2014) *TreeSim: Simulating trees under the birth-death model.* R package version 2.1.

Stanley, S.M., Signor III, P.W., Lidgard, S. & Karr, A.F. (1981) Natural clades differ from “random” clades: simulations and analyses. *Paleobiology*, 115–127.

Steffen, W., Richardson, K., Rockström, J., Cornell, S.E., Fetzer, I. *et al.* (2015) Planetary boundaries: Guiding human development on a changing planet. *Science*, **347**, 6223, 1259855.

Stern, S., Agra, M.d.F. & Bohs, L. (2011) Molecular delimitation of clades within new world species of the “spiny solanums” (solanum subg. leptostemonum). *Taxon*, **60**, 5, 1429–1441.

Stern, S., Bohs, L., Giacomin, L., Stehmann, J. & Knapp, S. (2013) A revision of solanum section gonatotrichum. *Systematic Botany*, **38**, 2, 471–496.

Stevens, J.T., Safford, H.D., Harrison, S. & Latimer, A.M. (2015) Forest disturbance accelerates thermophilization of understory plant communities. *Journal of Ecology*, **103**, 5, 1253–1263.

Stillman, J.H. (2003) Acclimation capacity underlies susceptibility to climate change. *Science*, **301**, 5629, 65–65.

Strathmann, R.R. & Slatkin, M. (1983) The improbability of animal phyla with few species. *Paleobiology*, 97–106.

Swihart, R.K., Gehring, T.M., Kolozsvary, M.B. & Nupp, T.E. (2003) Responses of 'resistant' vertebrates to habitat loss and fragmentation: the importance of niche breadth and range boundaries. *Diversity and Distributions*, **9**, 1, 1–18.

Symon, D.E. *et al.* (1981) A revision of the genus solanum in australia. *Journal of the Adelaide Botanic Gardens*, **4**, 1–367.

Takhtajan, A., *Flowering plants*. Springer Science & Business Media, 2009.

Tank, D.C., Eastman, J.M., Pennell, M.W., Soltis, P.S., Soltis, D.E. *et al.* (2015) Nested radiations and the pulse of angiosperm diversification: increased diversification rates often follow whole genome duplications. *New Phytologist*.

Tepe, E.J. & Bohs, L. (2010) A molecular phylogeny of solanum sect. pteroidea (solanaceae) and the utility of cosii markers in resolving relationships among closely related species. *Taxon*, 733–743.

Tepe, E.J. & Bohs, L. (2011) A revision of solanum section herpystichum. *Systematic Botany*, **36**, 4, 1068–1087.

The World Bank (2015) Agriculture, value added (percentage of GDP) in Colombia. Available at `http://data.worldbank.org/indicator/NV.AGR.TOTL.ZS/countries/CO?display=graph` (accessed: 2016.03.03).

Thomas, C.D. (2013) Local diversity stays about the same, regional diversity increases, and global diversity declines. *Proceedings of the National Academy of Sciences*, **110**, 48, 19187–19188.

Thomas, G.H., Hartmann, K., Jetz, W., Joy, J.B., Mimoto, A. *et al.* (2013) Pastis: an r package to facilitate phylogenetic assembly with soft taxonomic inferences. *Methods in Ecology and Evolution*, **4**, 11, 1011–1017.

Thuiller, W., Brotons, L., Araújo, M.B. & Lavorel, S. (2004) Effects of restricting environmental range of data to project current and future species distributions. *Ecography*, **27**, 2, 165–172.

Tilman, D., May, R.M., Lehman, C.L. & Nowak, M.A. (1994) Habitat destruction and the extinction debt. *Nature*, **371**, 65–66.

Tittensor, D.P., Walpole, M., Hill, S.L., Boyce, D.G., Britten, G.L. *et al.* (2014) A mid-term analysis of progress toward international biodiversity targets. *Science*, **346**, 6206, 241–244.

Toussaint, E.F. & Condamine, F.L. (2015) To what extent do new fossil discoveries change our understanding of clade evolution? a cautionary tale from burying beetles (coleoptera: Nicrophorus). *Biological Journal of the Linnean Society.*

Trabucco, A. & Zomer, R. (2009) Global aridity index (global-aridity) and global potential evapo-transpiration (global-pet) geospatial database. *CGIAR-CSI GeoPortal.*

Travis, J. (2003) Climate change and habitat destruction: a deadly anthropogenic cocktail. *Proceedings of the Royal Society of London B: Biological Sciences*, **270**, 1514, 467–473.

Van Kleunen, M., Dawson, W., Essl, F., Pergl, J., Winter, M. *et al.* (2015) Global exchange and accumulation of non-native plants. *Nature.*, **525**, 7567, 100–103.

Van Valen, L. (1973) A new evolutionary law. *Evolutionary Theory*, **1**, 1, 1–30.

Van Vuuren, D.P., Edmonds, J., Kainuma, M., Riahi, K., Thomson, A. *et al.* (2011) The representative concentration pathways: an overview. *Climatic Change*, **109**, 1-2, 5–31.

van Vuuren, D.P., Sala, O.E., Pereira, H.M. *et al.* (2006) The future of vascular plant diversity under four global scenarios. *Ecology and Society*, **11**, 2, 25.

Vilela, B. & Villalobos, F. (2015a) letsr: a new r package for data handling and analysis in macroecology. *Methods in Ecology and Evolution*, **6**, 10, 1229–1234.

Vilela, B. & Villalobos, F. (2015b) *letsR: Tools for Data Handling and Analysis in Macroecology.* R package version 2.1.

von Hagen, K.B. & Kadereit, J.W. (2001) The phylogeny of gentianella (gentianaceae) and its colonization of the southern hemisphere as revealed by nuclear and chloroplast dna sequence variation. *Organisms Diversity & Evolution*, **1**, 1, 61–79.

Vorontsova, M., Stern, S., Bohs, L., Knapp, S. *et al.* (2013) African spiny solanum (subgenus leptostemonum, solanaceae): a thorny phylogenetic tangle. *Botanical Journal of the Linnean Society*, **173**, 2, 176–193.

Vrba, E.S. (1992) Mammals as a key to evolutionary theory. *Journal of Mammalogy*, **73**, 1, 1–28.

Wahlert, G.A., Chiarini, F. & Bohs, L. (2014) Phylogeny of the carolinense clade of solanum (solanaceae) inferred from nuclear and plastid dna sequences. *Systematic Botany*, **39**, 4, 1208–1216.

Wahlert, G.A., Chiarini, F.E. & Bohs, L. (2015) A revision of solanum section lathyrocarpum (the carolinense clade, solanaceae). *Systematic Botany*, **40**, 3, 853–887.

Wardle, D.A., Bardgett, R.D., Callaway, R.M. & Van der Putten, W.H. (2011) Terrestrial ecosystem responses to species gains and losses. *Science*, **332**, 6035, 1273–1277.

Wearn, O.R., Reuman, D.C. & Ewers, R.M. (2012) Extinction debt and windows of conservation opportunity in the Brazilian amazon. *Science*, **337**, 6091, 228–232.

Weese, T.L. & Bohs, L. (2007) A three-gene phylogeny of the genus solanum (solanaceae). *Systematic Botany*, **32**, 2, 445–463.

Weese, T.L. & Bohs, L. (2010) Eggplant origins: out of africa, into the orient. *Taxon*, **59**, 1, 49–56.

Weir, J.T. & Schluter, D. (2007) The latitudinal gradient in recent speciation and extinction rates of birds and mammals. *Science*, **315**, 5818, 1574–1576.

Whalen, M.D. (1979) Allozyme variation and evolution in solanum section androceras. *Systematic Botany*, 203–222.

Whalen, M.D. (1984) Conspectus of species groups in solanum subgenus leptostemonum. *Gentes Herb.(Ithaca)*, **12**, 4, 179–282.

Whalen, M.D. & Caruso, E.E. (1983) Phylogeny in solanum sect. lasiocarpa (solanaceae): congruence of morphological and molecular data. *Systematic Botany*, 369–380.

White, F. (1983) The vegetation of africa, a descriptive memoir to accompany the unesco/aetfat/unso vegetation map of africa (3 plates, northwestern africa, northeastern africa, and southern africa, 1: 5,000,000).

Whitton, F.J., Purvis, A., Orme, C.D.L. & Olalla-Tárraga, M.Á. (2012) Understanding global patterns in amphibian geographic range size: does rapoport rule? *Global Ecology and Biogeography*, **21**, 2, 179–190.

Wiens, J.J. (2011) The causes of species richness patterns across space, time, and clades and the role of âĂIJecological limitsâĂİ. *The Quarterly Review of Biology*, **86**, 2, 75–96.

Wiens, J.J. (2015) Faster diversification on land than sea helps explain global biodiversity patterns among habitats and animal phyla. *Ecology letters*, **18**, 11, 1234–1241.

Wiens, J.J. & Donoghue, M.J. (2004) Historical biogeography, ecology and species richness. *Trends in ecology & evolution*, **19**, 12, 639–644.

Wiens, J.J. & Tiu, J. (2012) Highly incomplete taxa can rescue phylogenetic analyses from the negative impacts of limited taxon sampling. *PloS one*, **7**, 8, e42925.

Wilf, P., Carvalho, M.R., Gandolfo, M.A. & Cúneo, N.R. (2017) Eocene lantern fruits from gondwanan patagonia and the early origins of solanaceae. *Science*, **355**, 6320, 71–75.

Williams, J.W., Jackson, S.T. & Kutzbach, J.E. (2007) Projected distributions of novel and disappearing climates by 2100 ad. *Proceedings of the National Academy of Sciences*, **104**, 14, 5738–5742.

Willis, J.C. & Yule, G.U. (1922) Some statistics of evolution and geographical distribution in plants and animals, and their significance. *Nature*, **109**, 2728, 177–179.

Wood, H.M., Matzke, N.J., Gillespie, R.G. & Griswold, C.E. (2012) Treating fossils as terminal taxa in divergence time estimation reveals ancient vicariance patterns in the palpimanoid spiders. *Systematic Biology*, **62**, 2, 264–284.

Wood, S. (2012) mgcv: Mixed gam computation vehicle with gcv/aic/reml smoothness estimation.

Wright, J.S. & Muller-Landau, H.C. (2006) The uncertain future of tropical forest species. *Biotropica*, **38**, 4, 443–445.

Yang, Z. & Donoghue, P.C.J. (2016) Dating species divergences using rocks and clocks. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, **371**, 1699. doi:10.1098/rstb.2015.0126.

Yoder, J., Clancey, E., Des Roches, S., Eastman, J., Gentry, L. *et al.* (2010) Ecological opportunity and the origin of adaptive radiations. *Journal of evolutionary biology*, **23**, 8, 1581–1596.

Zanne, A.E., Tank, D.C., Cornwell, W.K., Eastman, J.M., Smith, S.A. *et al.* (2014) Three keys to the radiation of angiosperms into freezing environments.

Zomer, R.J., Bossio, D.A., Trabucco, A., Yuanjie, L., Gupta, D.C. *et al.*, *Trees and water: smallholder agroforestry on irrigated lands in Northern India*, vol. 122. IWMI, 2007.

Zomer, R.J., Trabucco, A., Bossio, D.A. & Verchot, L.V. (2008) Climate change mitigation: A spatial analysis of global land suitability for clean development mechanism afforestation and reforestation. *Agriculture, ecosystems & environment*, **126**, 1, 67–80.

Zuur, A., Ieno, E.N., Walker, N., Saveliev, A.A. & Smith, G.M., *Mixed effects models and extensions in ecology with R*. Springer, 2009.

APPENDIX A

Supplement to Chapter 1

(1) what are the speciation and extinction rates of particular taxa? (Nee *et al.*, 1992, 1994b) derived a maximum likelihood method to estimate lineage speciation and extinction rates based on phylogenies of extant species. This maximum likelihood function depends only on the distribution of waiting times between speciation events, not on the tree symmetry. The accumulation of lineages through time is commonly visualised using lineage-through-time (LTT) plots (Figure A.1). In the simple branching process known as the Yule process (i.e., speciation without extinction, cite here), the number of lineages increases exponentially through time at a rate λ (see Figure A.1). When extinction is included in the branching model (i.e., birth-death model), an apparent acceleration of lineages towards the present known as the pull-of-the-present effect is created (Nee *et al.*, 1994b). This change in the accumulation of lineages is due to the decrease of the influence of extinction towards the present (i.e., lineages that appeared closer to the present are less likely to have already become extinct than those that appeared closer to the past). Thus, the change in slope part of the curve provides an estimate of speciation rates (λ) and the straight line an estimate of diversification (λ - mu) (Harvey *et al.*, 1994; Purvis, 1996a). Intuitively, this method determines parameters in a way that the LTT plots of reconstructed trees are as similar as possible to the LTT of empirical trees (Stadler, 2013; Morlon, 2014).

(2) Are diversification rates constant through time? The method of Nee *et al.* (1994a) assumes that λ and μ remain constant through time. Using this assumption as a null model, it is possible to test whether rates are constant or vary through time. There are two ways to test this hypothesis: (1) summary statistics that describe the distribution of branching times such as γ statistics or (2) maximum-likelihood model-fitting and comparison with the constant rates null hypothesis. In the first approach, the γ statistics detect departure from the null model (pure-death model) based on tree symmetry and

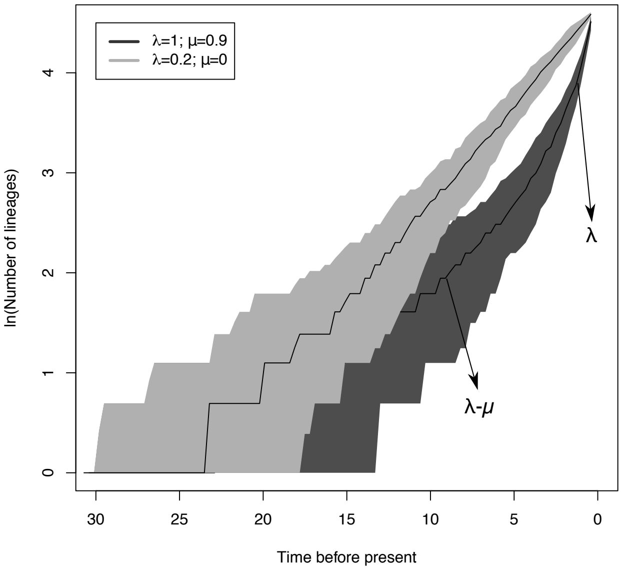


Figure A.1: Expected lineage-through-time (LTT) plots under two different speciation-extinction models. When extinction is taking into account, the slope of the LTT plot increases from $lambda - mu$ to $lambda$ towards the present (pull-of-the-present effect). Region around the median curve corresponds to the two-tailed 95 % lower and upper quantiles of number of species of 100 simulated trees. The R package TreeSim v 2.1 (Stadler, 2014) was used to perform the simulations.

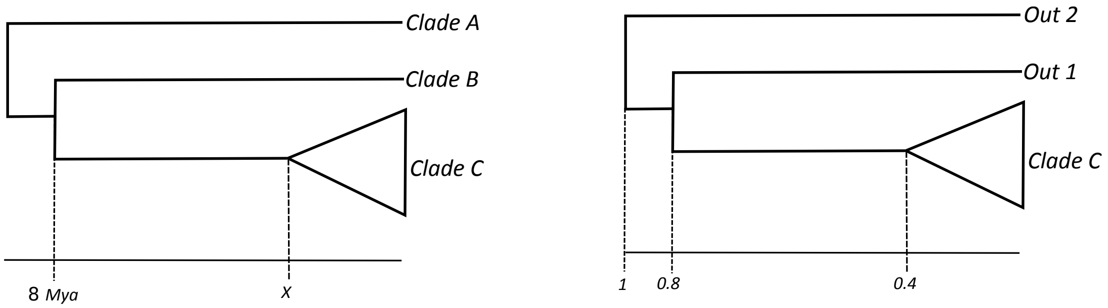
distribution of internode distance (Pybus & Harvey, 2000). This index is expected to be positive under a constant birth-death model with a non-zero extinction (*i.e.,* increase in the lineage accumulation rate towards the present); and negative when the branching tempo departs from the constant rate birth-death model (*i.e.,* slowdown in lineage accumulation rate towards the present). The γ index provides a useful tool to detect variation of rates through time; however, alternative diversification scenarios can produce phylogenies with similar tree shape which limits the inference of evolutionary dynamics more precisely (Morlon, 2014). Maximum-likelihood model-fitting is the second approach to detect variation of rates through time. Instead of considering a single summary statistics; maximum-likelihood methods use the information of the whole reconstructed tree. Several time-varying diversification models have been proposed (e.g., Purvis *et al.* 1995; Bininda-Emonds *et al.* 2007; Rabosky 2006; Morlon *et al.* 2011; Stadler 2011; Hallinan 2012); however, additional models are needed to understand which processes are behind these changes in rates through time such as environmental-dependence models (e.g., Condamine *et al.* 2013) or diversity-dependence models (e.g., Etienne *et al.* 2011; Rabosky & Lovette 2008; Etienne & Haegeman 2012).

(4) Is there a significant association between diversification rate change and trait states? The detection of regions in a phylogeny that has experienced significant changes in rates can help us to test hypotheses about the influence of specific characters on shifts of diversification rates (*e.g.,* geographical distribution, morphological traits or life-history). Given a phylogeny and trait values for the species or clades at the tips, it is possible to infer significant associations between changes in rates and values of a trait (Slowinski & Guyer, 1993). Initially, sister-group comparisons were used to understand significant correlations between trait states and diversity (usually measured as number of species) or net diversification rates (e.g., Mitter *et al.* 1988; Barraclough *et al.* 1995; Agapow & Isaac 2002; Isaac *et al.* 2003; Phillimore *et al.* 2006). However, these approaches limit the use of phylogenetic information and so do not consider directly changes in speciation and extinction rates associated with trait evolution (Barraclough & Nee, 2001; Maddison *et al.*, 2007). To overcome these limitations, Maddison *et al.* (2007) suggested the first set of models to estimate the effect of a character on speciation and extinction rates: Binary State Speciation Extinction (BiSSE) model for discrete traits. Since then other similar methods have been proposed: Quantitative State Speciation and Extinction (QuaSSE) approach for continuous characters (FitzJohn, 2010) and Geographic State Speciation and Extinction (GeoSSE) for geographic traits assuming cladogenetic character evolution (Goldberg *et al.*, 2011).

These methods have been widely used in different taxonomic groups and regions (more than 110 papers published between 2012-2014 according to Maddison 2014). However, Rabosky & Goldberg (2015) recently showed how BiSSE-like approaches have a high Type I error rate, making their conclusions unreliable. One of the main problems is that these models do not require repeated associations between states and diversification changes for statistical significance. This issue has been highlighted not just recently by Maddison & FitzJohn (2015), but also in the past by Slowinski & Guyer (1993, p.1020) who stated: "the existence of a cause-and-effect relationship between some trait and increased diversity can be tested if several to many groups possessing the trait are considered, not just one". Recently, a new approach for detecting character-dependent diversification using a semi-parametric test has been proposed by Rabosky & Huang (2015) and Rabosky & Goldberg (2017), that requires repeated associations between trait states and changes in diversification. However, no empirical studies have yet applied this new approach.

APPENDIX B

Supplement to Chapter 2



Backbone

PASTIS

Figure B.1: *Solanum* phylogeny backbone-clade grafting The depth of each PASTIS subtree, which represents the PASTIS runs for each clade/section of *Solanum*, was scaled to 1.0 in order to substitute the ingroup into the backbone tree to replace the single branch that the clade represents. The two outgroups of the clade were dropped and then the clade was grafted into the backbone inferring the depth of the crown group in the backbone from the depth of the stem group of the PASTIS subtree and the depth of the node in the backbone that supports the clade and its outgroup. For instance, if the depth of the node linking the clade C and the clade containing clade C's outgroup is 8 Ma, the crown age of Clade C in the backbone is set to $8 * 0.4/0.8 = 4Ma$.

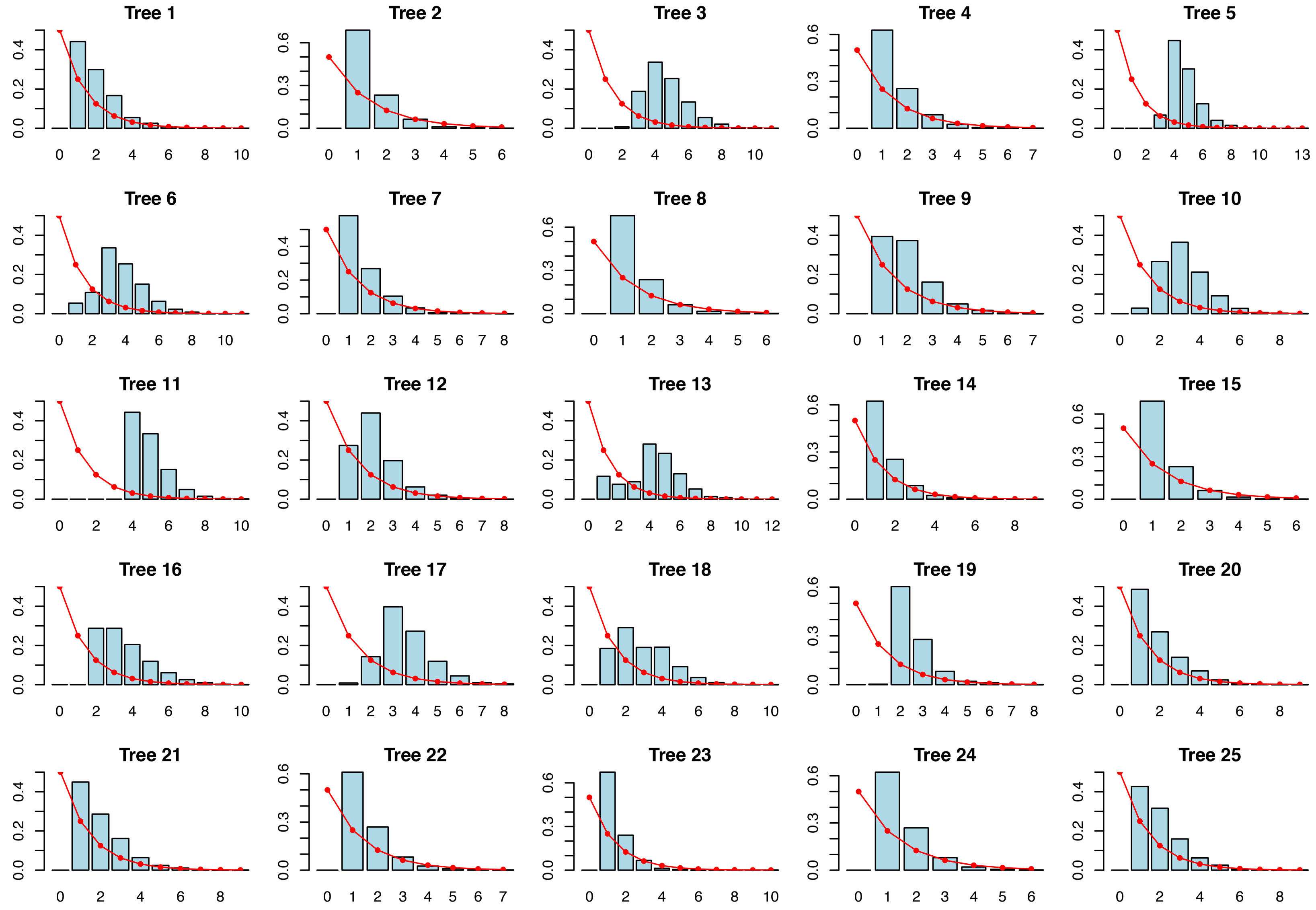


Figure B.2: Marginal posterior probability distributions on the number of shifts for 1-25 trees used in the BAMM analysis (filled histograms). The prior distribution on the number of shifts for each tree is illustrated in red.

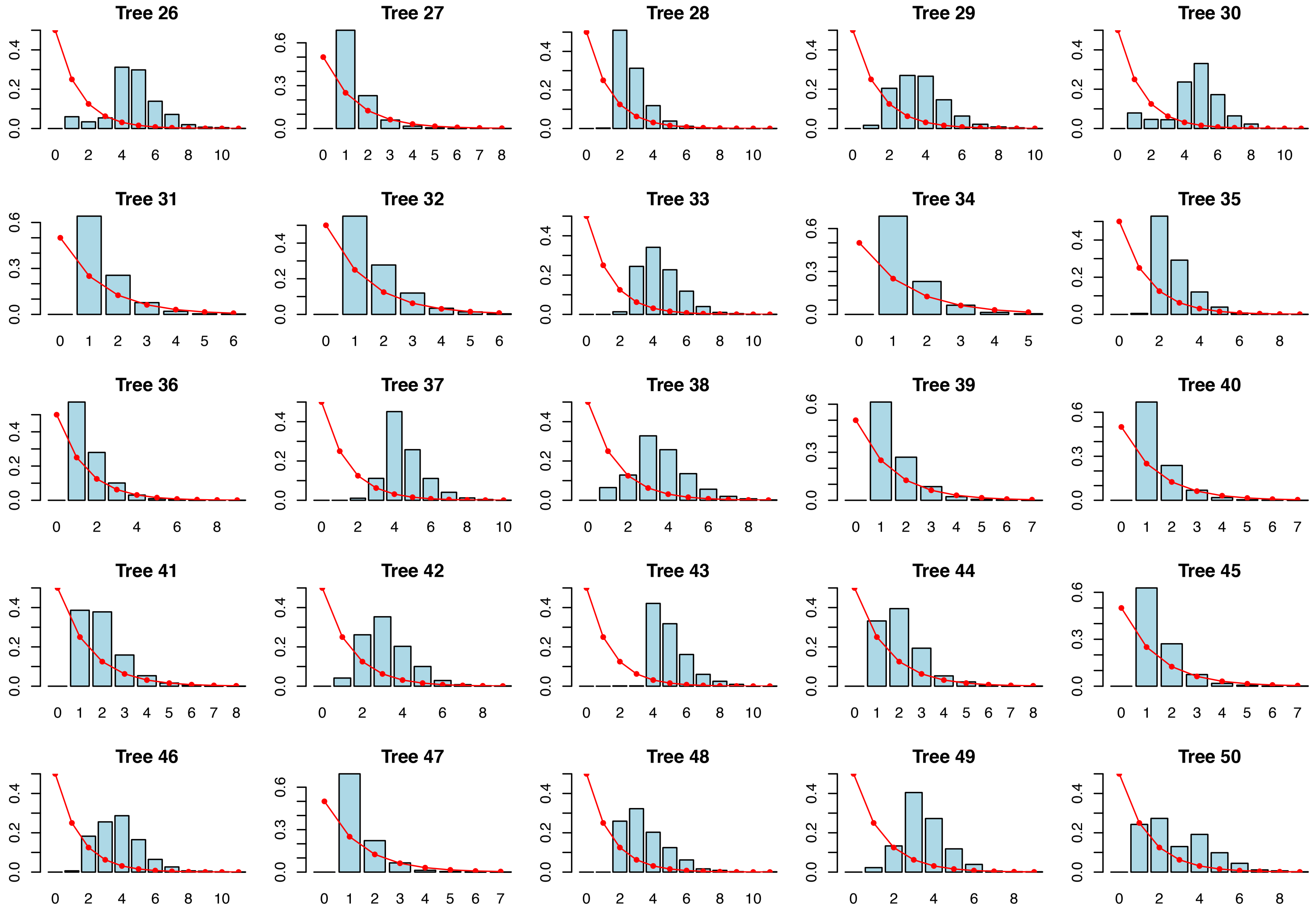


Figure B.3: Marginal posterior probability distributions on the number of shifts for 26-50 trees used in the BAMM analysis (filled histograms). The prior distribution on the number of shifts for each tree is illustrated in red.

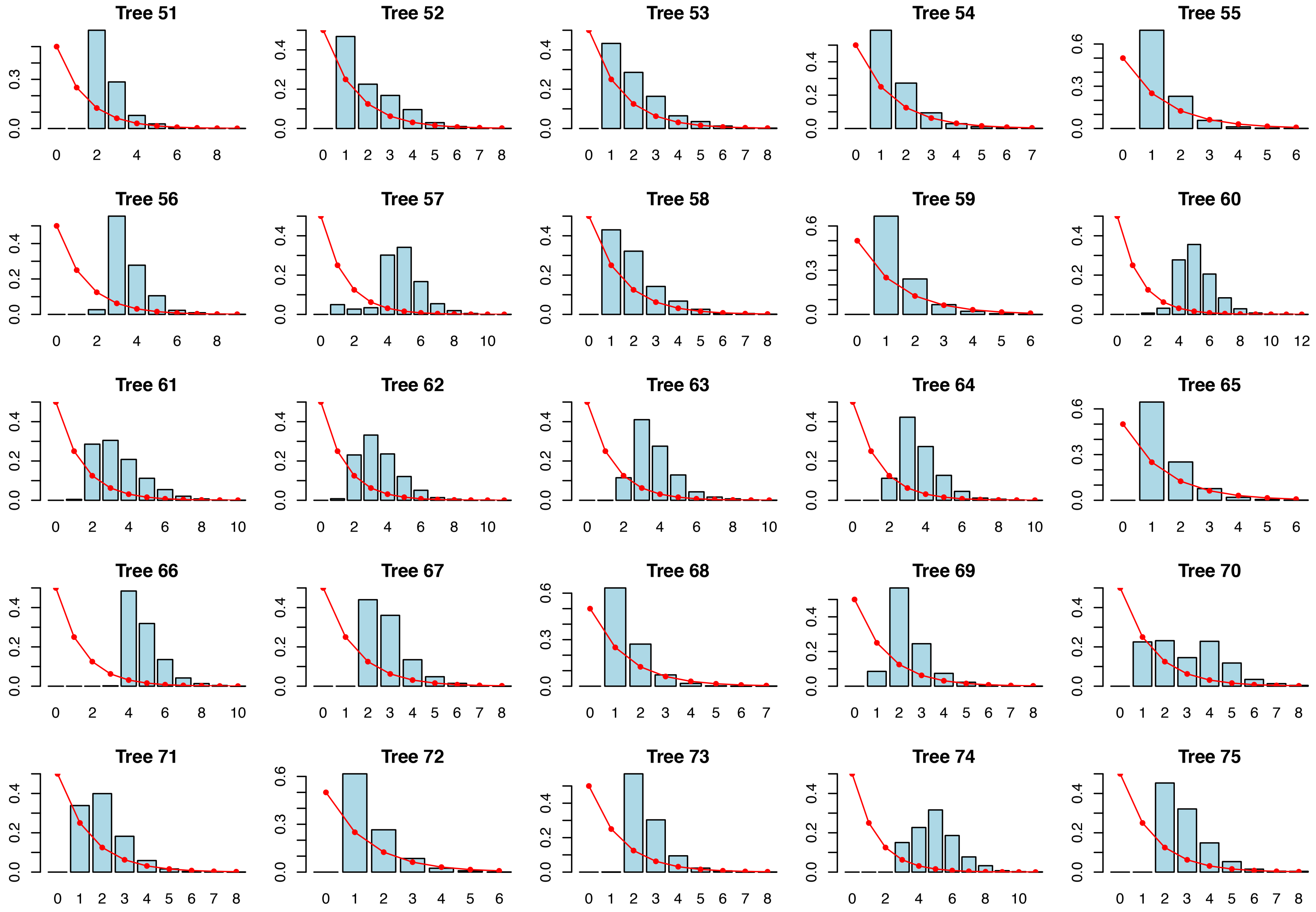


Figure B.4: Marginal posterior probability distributions on the number of shifts or 51-75 trees used in the BAMM analysis (filled histograms). The prior distribution on the number of shifts for each tree is illustrated in red.

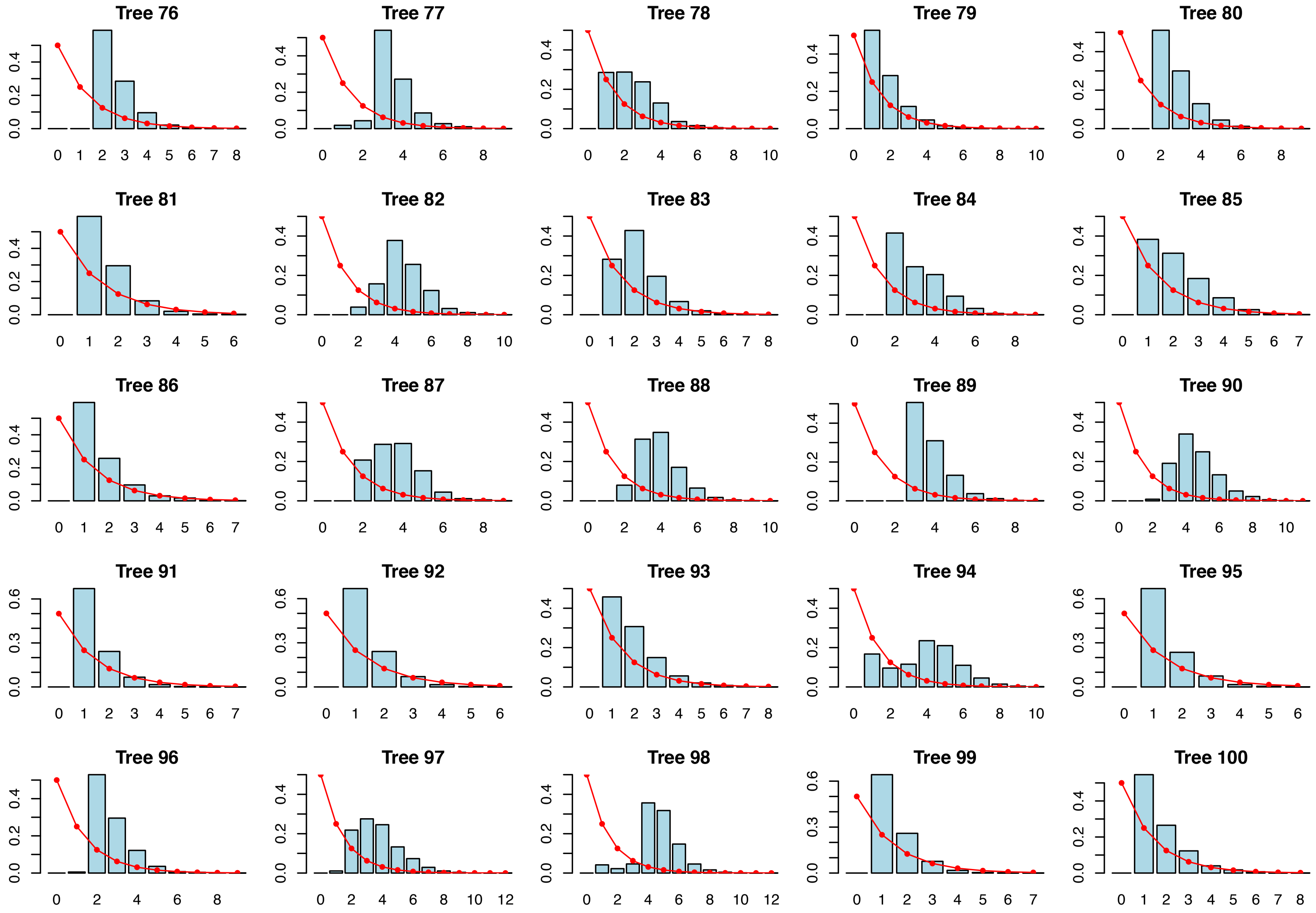


Figure B.5: Marginal posterior probability distributions on the number of shifts for 76-100 trees used in the BAMM analysis (filled histograms). The prior distribution on the number of shifts for each tree is illustrated in red.

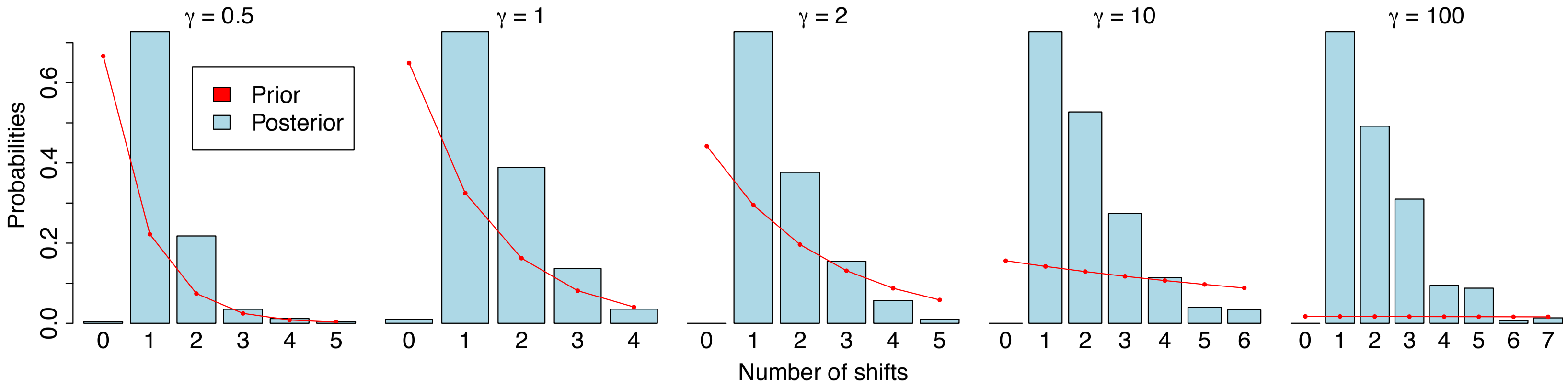


Figure B.6: Effect of the prior on the marginal posterior distribution of rate shifts in *Solanum* using the Särkinen *et al.* (2013) phylogeny. Histograms represent the marginal posterior probability distribution on the number of shifts and lines correspond to the prior distribution on the expected number of shifts. Although there is a slight change in the distribution on the probabilities with different priors, there is no evidence that the results are unusually sensitive to the priors used. Note that some models were not observed in the different treatments (e.g., models with 0 shifts were usually not observed showing an overwhelming evidence of rate heterogeneity in both trees).

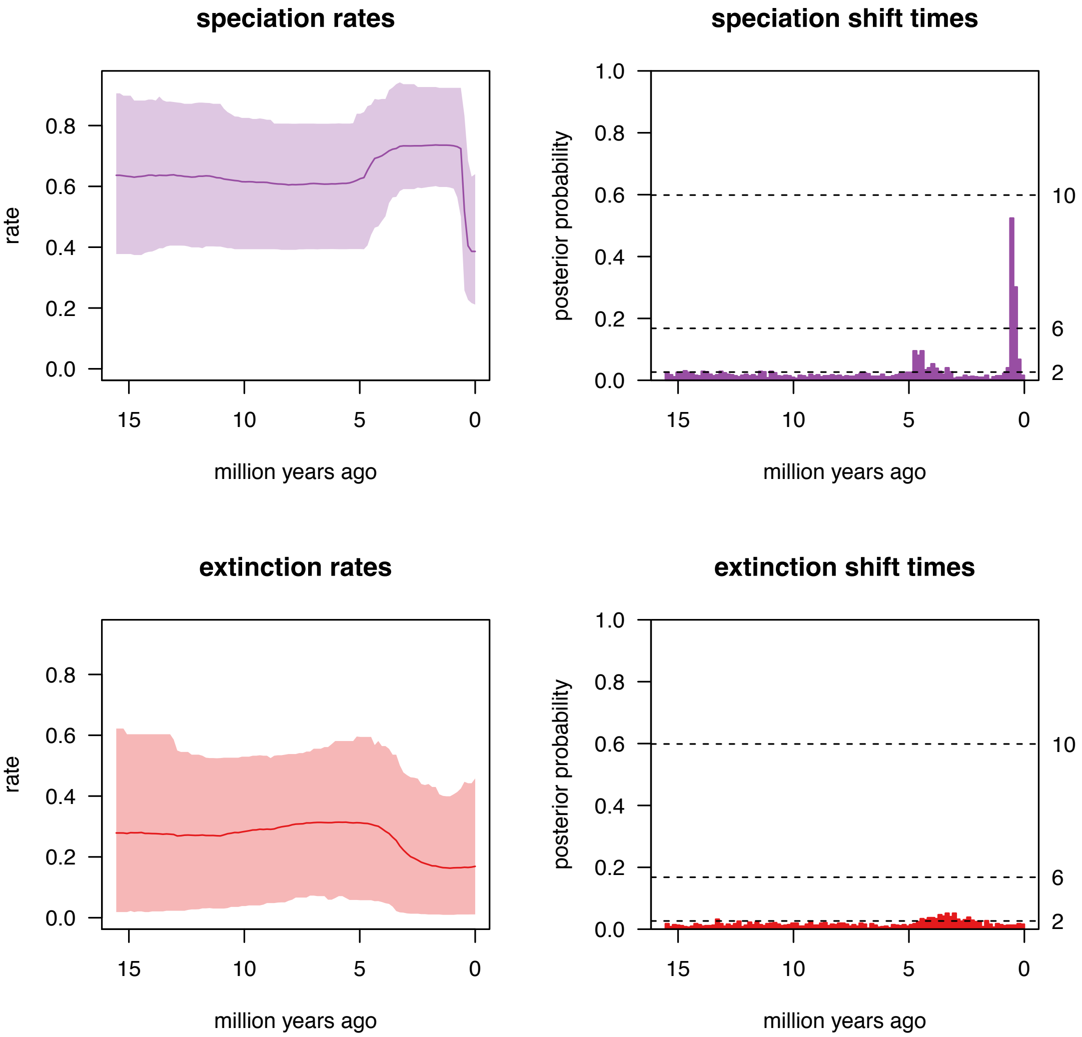


Figure B.7: Estimates of the rates and shifts in lineages diversification through time under the TESS approach using the Särkinen *et al.* (2013) phylogeny. Plots in the left show the posterior mean and 95% confidence intervals for speciation and extinction rates. Plots in the right show the temporal significant shifts estimated by Bayes factors (lnBF, numbers in the right axes). Bars indicate the posterior probability of shifts in the time slide. Significant shifts exceed the specified significant threshold (2 ln BF > 6).

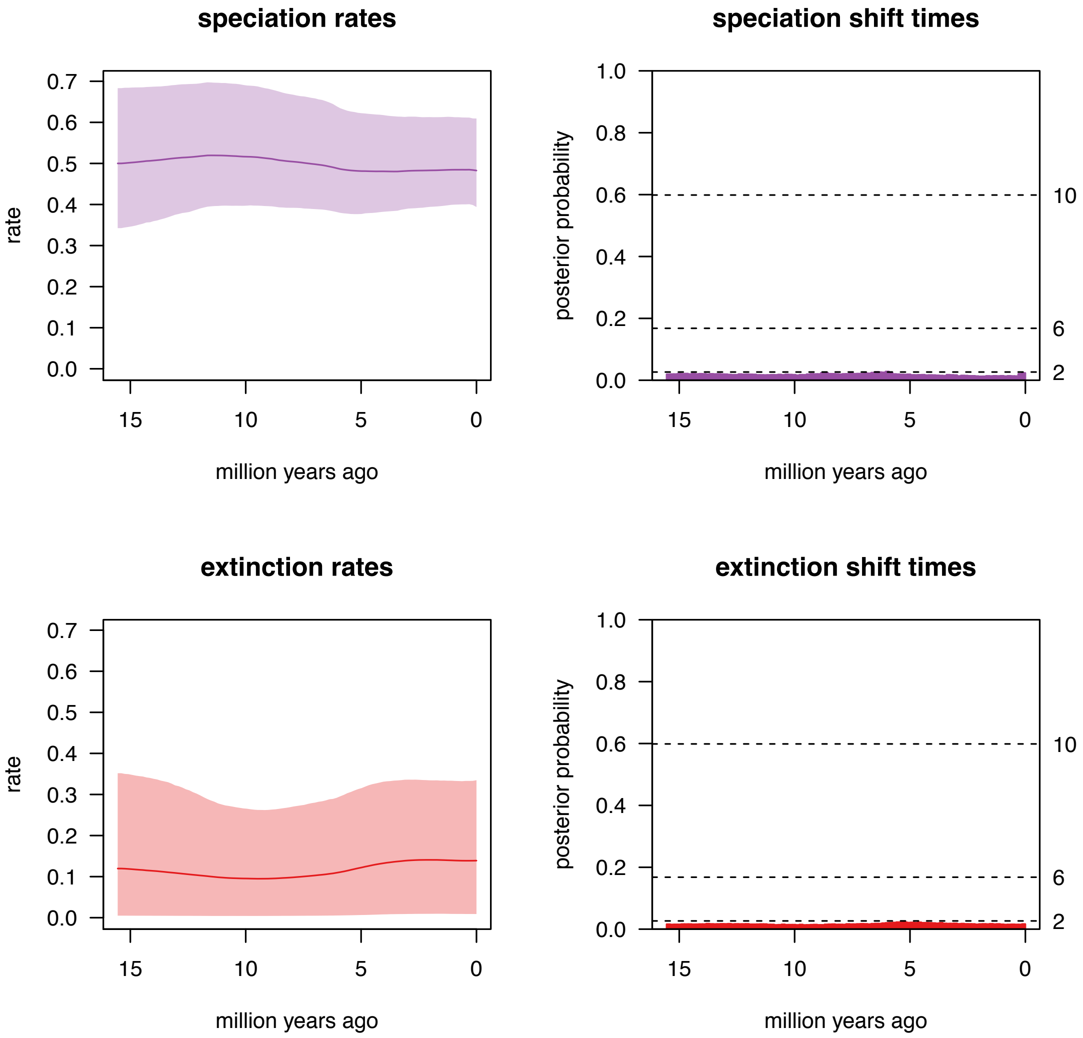


Figure B.8: Estimates of the rates and shifts in lineages diversification through time under the TESS approach using a distribution of 100 trees created by the polytomy resolver PASTIS. Plots in the left show the posterior mean and 95% confidence intervals for speciation and extinction rates. Plots in the right show the temporal significant shifts estimated by Bayes factors (lnBF, numbers in the right axes). Bars indicate the posterior probability of shifts in the time slide. Significant shifts exceed the specified significant threshold (2 ln BF $>$ 6).

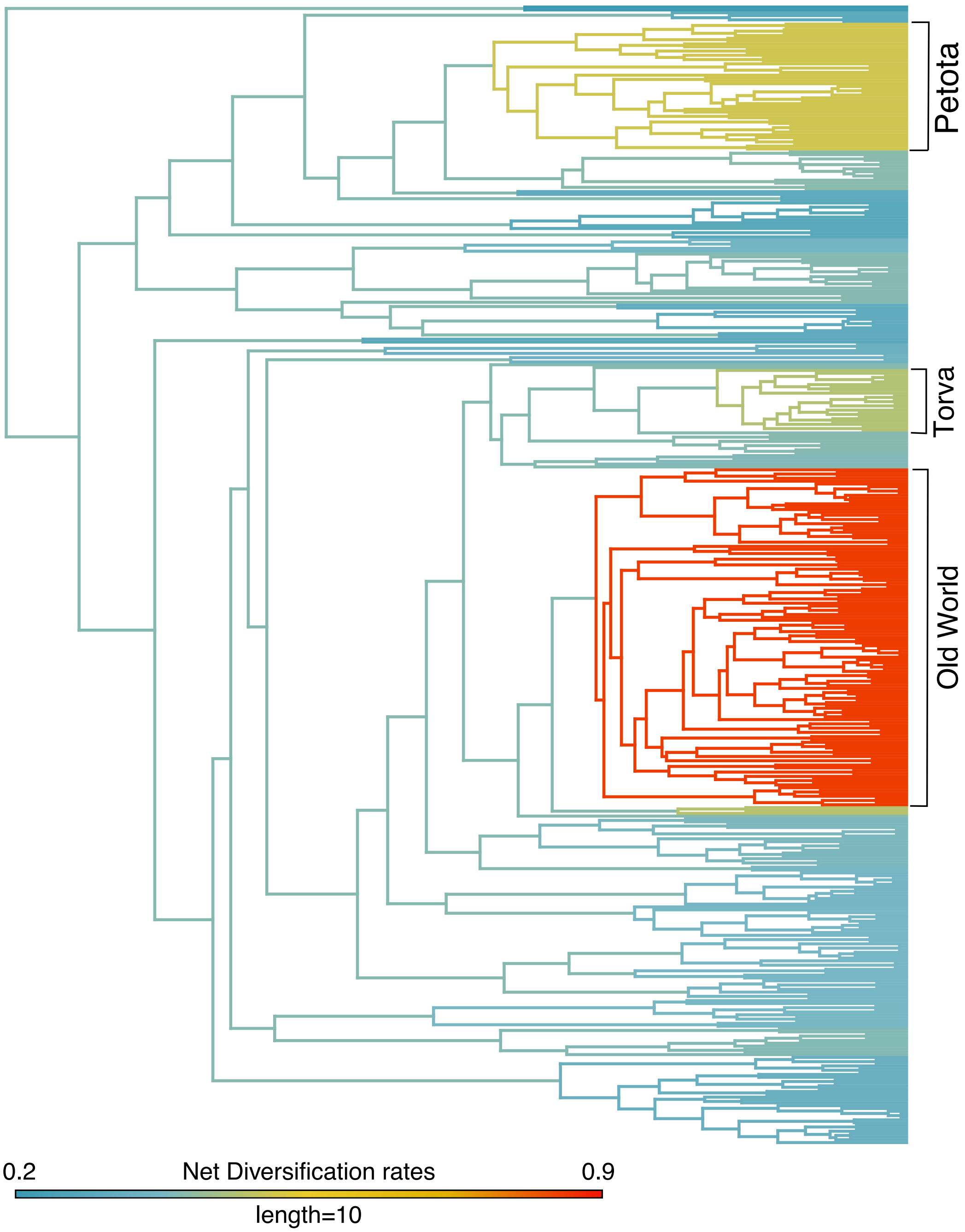


Figure B.9: Net diversification rates of *Solanum* estimated by RevBayes. RevBayes results (using 20 out of the 100 trees created by PASTIS) supporting the distinctive radiation of the Old world clade and the signal of other potential radiations such as the Petota clade within the non-spiny solanums and Torva clade within the spiny solanums.

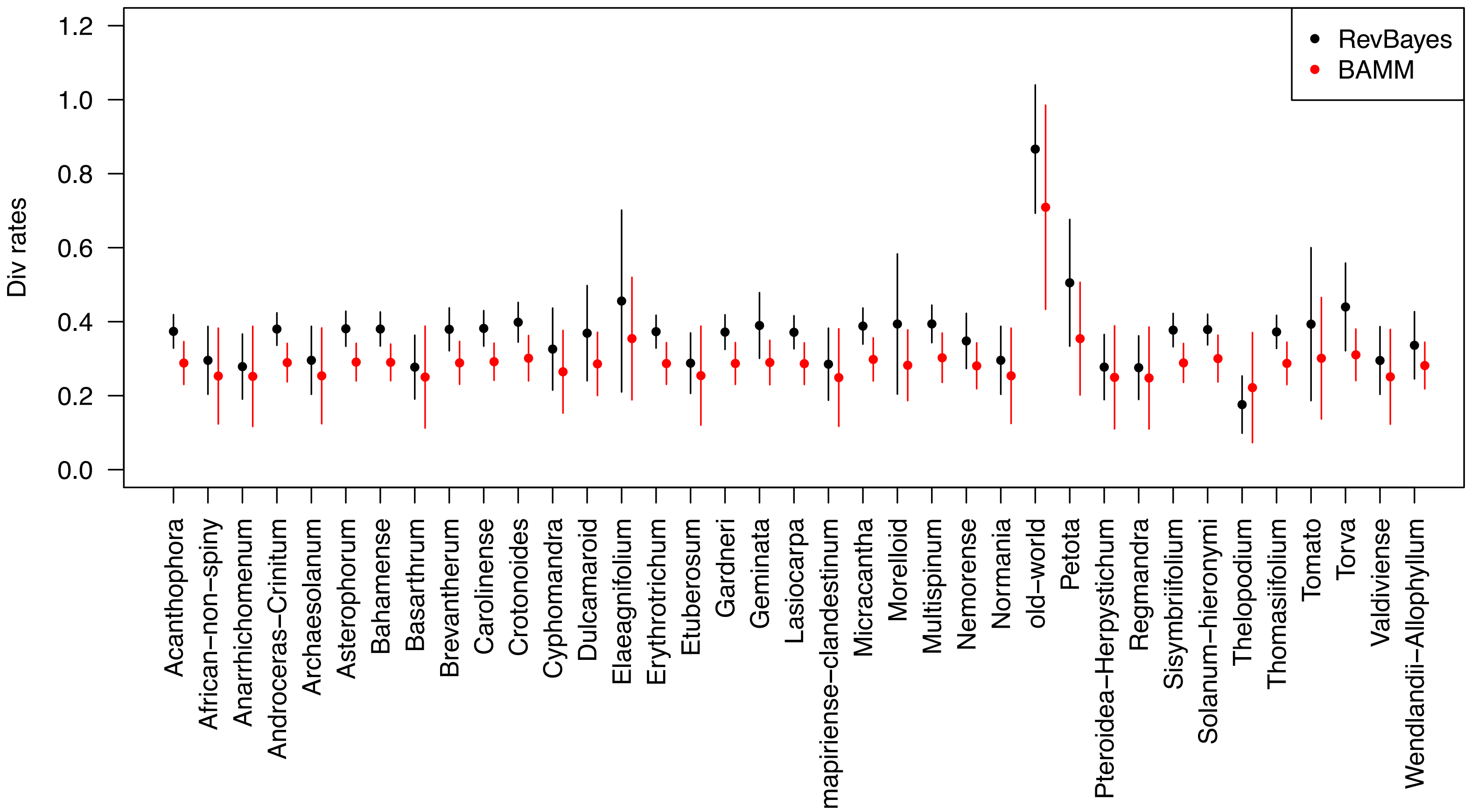


Figure B.10: Mean net diversification rates in each of the *Solanum* sections estimated using BAMM and RevBayes. Average net diversification rates under BAMM and RevBayes analyses using a distribution of 20 trees obtained by the polytomy resolver PASTIS. Bars represent the 95% confidence intervals for the diversification estimates.

APPENDIX C

Supplement to Chapter 3

Species	Source	Link
Solanum aculeatissimum	Solanaceae source	http://solanaceaesource.org/taxonomy/term/105386/descriptions
Solanum aethiopicum	Solanaceae source	http://solanaceaesource.org/taxonomy/term/105425/descriptions
Solanum americanum	Edmonds & Chweya (1997)	
Solanum atropurpureum	Solanaceae source	http://solanaceaesource.org/taxonomy/term/105749/descriptions
Solanum bahamense	Solanaceae source	http://solanaceaesource.org/taxonomy/term/105804/descriptions
Solanum betaceum	Solanaceae source	http://solanaceaesource.org/taxonomy/term/105869/descriptions
Solanum capsicoides	Solanaceae source	http://solanaceaesource.org/taxonomy/term/106111/descriptions
Solanum chrysotrichum	Solanaceae source	http://solanaceaesource.org/taxonomy/term/106264/descriptions
Solanum diphyllum	Solanaceae source	http://solanaceaesource.org/taxonomy/term/106688/descriptions
Solanum dulcamara	Solanaceae source	http://solanaceaesource.org/taxonomy/term/106742/descriptions
Solanum elaeagnifolium	Solanaceae source	http://solanaceaesource.org/taxonomy/term/106820/descriptions
Solanum jamaicense	Solanaceae source	http://solanaceaesource.org/taxonomy/term/107611/descriptions
Solanum laxum	Solanaceae source	http://solanaceaesource.org/taxonomy/term/107817/descriptions
Solanum linnaeanum	Solanaceae source	http://solanaceaesource.org/taxonomy/term/107893/descriptions
Solanum lycopersicum	Solanaceae source	http://solanaceaesource.org/taxonomy/term/107988/descriptions
Solanum macrocarpon	Solanaceae source	http://solanaceaesource.org/taxonomy/term/108022/descriptions
Solanum mammosum	Solanaceae source	http://solanaceaesource.org/taxonomy/term/108080/descriptions
Solanum mauritianum	Solanaceae source	http://solanaceaesource.org/taxonomy/term/108132/descriptions
Solanum melongena	Solanaceae source	http://solanaceaesource.org/taxonomy/term/108174/descriptions
Solanum nigrum	Edmonds & Chweya (1997)	
Solanum pimpinellifolium	Solanaceae source	http://solanaceaesource.org/taxonomy/term/108944/descriptions
Solanum pseudocapsicum	Solanaceae source	http://solanaceaesource.org/taxonomy/term/109079/descriptions
Solanum robustum	Solanaceae source	http://solanaceaesource.org/taxonomy/term/109318/descriptions
Solanum scabrum	Edmonds & Chweya (1997)	
Solanum sisymbriifolium	Solanaceae source	http://solanaceaesource.org/taxonomy/term/109613/descriptions
Solanum torvum	Solanaceae source	http://solanaceaesource.org/taxonomy/term/110004/descriptions
Solanum triflorum	GRIN database	http://www.ars-grin.gov/cgi-bin/npgs/html/taxon.pl?101526
Solanum tuberosum	Solanaceae source	http://solanaceaesource.org/taxonomy/term/110092/descriptions
Solanum viarum	Solanaceae source	http://solanaceaesource.org/taxonomy/term/110343/descriptions
Solanum villosum	Edmonds & Chweya (1997)	
Solanum wendlandii	Solanaceae source	http://solanaceaesource.org/taxonomy/term/110454/descriptions
Solanum wrightii	Solanaceae source	http://solanaceaesource.org/taxonomy/term/110479/descriptions

Table C.1: List of *Solanum* species with widespread distribution which are considered as cultivated or naturalised and whose distribution could be obscured by current human activities. For each species, the descriptions of their native distribution were used to keep or drop their location into the analysis

	F	**U**	**I**	**N**	**T**	**P**
F	1	0.5	1	0.001	0.5	1
U	0.5	1	1	0.001	0.5	0.001
I	1	1	1	0.001	0.5	0.5
N	0.001	0.001	0.001	1	1	0.5
T	0.5	0.5	0.5	1	1	0.001
P	1	0.001	0.5	0.5	0.001	1

Table C.2: Dispersal matrix used to constrain the dispersal probabilities among areas in the biogeographic analysis. This probabilities were based on general distances among regions. 1, 0.5 and 0.001 represent easy, medium, and hard dispersal, respectively. Abbreviations: F, Africa; U, Australia; I, Indo-Pacific; N, Neartic; T, Neotropics; P, Paleartic.

	Des	FGS	M	Med	Mnt	TeF	TeC	TeSav	TSC	SDTF	Tsv	TrF	Total
Des	-	0.01 (0.1)	0.02 (0.14)	3.12 (1.16)	1.65 (0.54)	0.85 (0.77)	0.85 (0.36)	2.06 (1.08)	0.77 (0.65)	2.83 (1.3)	2.51 (1.2)	1.53 (1.11)	16.2
FGS	0 (0)	-	0 (0)	0.01 (0.1)	0 (0)	0.01 (0.1)	0 (0)	0 (0)	0 (0)	0 (0)	0.03 (0.17)	0 (0)	0.05
M	0 (0)	0 (0)	-	0 (0)	0 (0)	0.23 (0.42)	0 (0)	0 (0)	0 (0)	0.19 (0.39)	0.21 (0.41)	0.09 (0.29)	0.72
Med	2.4 (1.1)	0 (0)	0 (0)	-	0.49 (0.63)	1.94 (1.03)	0 (0)	0.22 (0.42)	0 (0)	0.12 (0.33)	0.53 (0.7)	0.2 (0.4)	5.9
Mnt	0.11 (0.31)	0 (0)	0 (0)	0.14 (0.35)	-	0.64 (0.52)	0 (0)	0 (0)	0.08 (0.27)	1.1 (0.61)	0.95 (0.81)	1.68 (1.11)	4.7
TeF	1.28 (0.93)	0 (0)	0.22 (0.42)	4.19 (1.16)	0.52 (0.59)	-	0.48 (0.5)	0.83 (0.68)	0 (0)	1.28 (1.04)	4.51 (1.37)	2.61 (1.01)	15.92
TeC	0.15 (0.36)	0 (0)	0 (0)	0 (0)	0 (0)	0.51 (0.5)	-	0 (0)	0.06 (0.24)	0 (0)	0.03 (0.17)	0 (0)	0.75
TeSav	1.12 (0.89)	0 (0)	0 (0)	0.22 (0.42)	0 (0)	0.39 (0.58)	0 (0)	-	0 (0)	0.04 (0.2)	0.13 (0.34)	0.23 (0.42)	2.13
TSC	1.71 (0.73)	0 (0)	0 (0)	0 (0)	0.01 (0.1)	0 (0)	1.11 (0.42)	0 (0)	-	1.45 (1.1)	0.04 (0.2)	0.31 (0.53)	4.63
SDTF	2.09 (1.26)	0 (0)	0.24 (0.43)	0.17 (0.4)	0.3 (0.48)	1.44 (0.92)	0 (0)	0.07 (0.26)	1.97 (1.08)	-	2.03 (1.02)	2.47 (1.52)	10.78
Tsv	3.48 (1.23)	0.96 (0.2)	0.37 (0.49)	0.91 (0.73)	3.13 (0.97)	2.17 (1.13)	0.03 (0.17)	0.43 (0.5)	0.07 (0.26)	1.61 (1.01)	-	5.74 (1.86)	18.9
TrF	8.56 (1.83)	0 (0)	2.32 (0.78)	0.55 (0.77)	5.59 (1.3)	5.48 (1.23)	0.04 (0.2)	0.29 (0.46)	6.14 (1.22)	16.77 (2.56)	16.12 (2.4)	-	61.86
Total	20.9	0.97	3.17	9.31	11.69	13.66	2.51	3.9	9.09	25.39	27.09	14.86	142.54

Table C.3: Mean number and standard deviations of dispersal events among biomes estimated with biogeographical stochastical mapping (BSM). Counts were averaged across 100 BSMs. Events with more than one dispersal are colored in grey. The rows represent the sources and the columns the sinks. Des, Deserts and Xeric Shrublands; FGS, Flooded Grasslands and Savannas; M, Mangroves; Med, Mediterranean Forests, Woodlands and Scrub; Mnt, Montane Grasslands and Shrublands; TeF, Temperate Broadleaf and Mixed Forests; TeC, Temperate Conifer Forests; TeSav, Temperate Grasslands, Savannas and Shrublands; TSC, Tropical and Subtropical Coniferous Forests; SDTF, Tropical and Subtropical Dry Broadleaf Forests; Tsv, Tropical and Subtropical Grasslands, Savannas and Shrublands; TrF, Tropical and Subtropical Moist Broadleaf Forests.

APPENDIX D

Supplement to Chapter 4

Table D.1: List of species with significant negative log-response ratios – ln(RR)– for the Hum/Nat land-use treatment analysis, and considered mesic and aseasonal according to their climatic affiliations (see Figure 4.2 as reference). se, standard errors of ln(RR); N, number of studies used to calculate the log-response ratios; IUCN RL, the IUCN Red List of threatened species status for each species.

Species	Family	ln(RR)	se	sig.	N	IUCN RL
Abrus precatorius	Fabaceae	-3.671	1.512	0.02	2	
Acalypha diversifolia	Euphorbiaceae	-1.353	0.665	0.04	1	
Acineta barkeri	Orchidaceae	-3.648	1.503	0.02	1	
Aciotis rubricaulis	Melastomataceae	-1.857	0.256	0.00	1	
Adenanthera pavonina	Fabaceae	-3.651	1.445	0.01	1	
Afrostyrax lepidophyllus	Huaceae	-2.916	1.466	0.05	1	VU
Aglaia argentea	Meliaceae	-4.634	1.568	0.00	1	LR/LC
Aglaia korthalsii	Meliaceae	-3.285	1.483	0.03	1	LR/NT
Aidia genipiflora	Rubiaceae	-3.409	1.498	0.02	1	
Alphitonia incana	Rhamnaceae	-3.262	1.500	0.03	1	
Alpinia oceanica	Zingiberaceae	-1.444	0.541	0.01	1	
Alpinia purpurata	Zingiberaceae	-4.619	1.646	0.01	1	
Amaranthus spinosus	Amaranthaceae	-3.076	1.463	0.04	1	
Amphilophium pannosum	Bignoniaceae	-3.084	1.333	0.02	1	
Anisophyllea cabole	Anisophylleaceae	-5.294	1.565	0.00	1	VU
Annickia polycarpa	Annonaceae	-3.677	1.521	0.02	1	
Anthonotha macrophylla	Fabaceae	-3.701	1.499	0.01	2	
Ardisia celebica	Primulaceae	-3.541	1.492	0.02	1	
Arenga undulatifolia	Arecaceae	-3.592	1.537	0.02	2	
Arthrostemma ciliatum	Melastomataceae	-3.515	1.498	0.02	1	

Artocarpus integer	Moraceae	-3.541	1.492	0.02	1	
Asplundia alata	Cyclanthaceae	-2.360	0.822	0.00	1	
Asplundia isabellina	Cyclanthaceae	-3.347	1.483	0.02	1	
Astronium graveolens	Anacardiaceae	-2.190	1.003	0.03	1	
Aulacocalyx pallens	Rubiaceae	-4.964	1.537	0.00	1	
Baccaurea tetrandra	Phyllanthaceae	-3.541	1.492	0.02	1	
Baccharis inamoena	Asteraceae	-5.041	1.660	0.00	1	
Baphia pubescens	Fabaceae	-3.060	1.474	0.04	1	
Begonia convallariodora	Begoniaceae	-3.219	1.473	0.03	1	
Begonia urophylla	Begoniaceae	-2.280	0.823	0.01	1	
Bridelia micrantha	Phyllanthaceae	-2.570	0.426	0.00	2	
Brosimum alicastrum	Moraceae	-5.052	1.639	0.00	1	
Buchanania arborescens	Anacardiaceae	-1.586	0.744	0.03	2	
Bursera simaruba	Burseraceae	-2.176	0.558	0.00	1	
Bussea occidentalis	Fabaceae	-3.188	1.482	0.03	1	
Calathea macrosepala	Marantaceae	-1.699	0.409	0.00	1	
Calpocalyx brevibracteatus	Fabaceae	-4.230	1.318	0.01	1	
Calycophyllum candidissimum	Rubiaceae	-4.152	1.543	0.01	1	
Campylospermum vogelii	Ochnaceae	-4.280	0.815	0.00	1	
Carapa procera	Meliaceae	-5.640	1.735	0.00	1	
Carludovica drudei	Cyclanthaceae	-3.946	1.539	0.01	1	
Caryota rumphiana	Arecaceae	-3.396	1.512	0.02	1	
Casearia barteri	Salicaceae	-2.204	0.400	0.00	1	LR/LC
Cassipourea hiotou	Rhizophoraceae	-3.895	1.541	0.01	1	VU
Castanopsis acuminatissima	Fagaceae	-4.225	0.848	0.00	1	
Cecropia obtusifolia	Urticaceae	-1.483	0.673	0.03	1	LR/LC
Centradenia paradoxa	Melastomataceae	-3.150	1.468	0.03	1	
Centrosema pubescens	Fabaceae	-2.063	0.824	0.01	2	
Cestrum laevigatum	Solanaceae	-1.631	0.613	0.01	1	
Chamaedorea matae	Arecaceae	-3.150	1.468	0.03	1	
Chamaedorea pumila	Arecaceae	-5.322	1.693	0.00	1	
Chamissoa altissima	Amaranthaceae	-5.069	1.663	0.00	1	
Chionanthus mildbraedii	Oleaceae	-1.423	0.409	0.00	1	
Chrysophyllum africanum	Sapotaceae	-3.899	1.459	0.01	1	
Chrysophyllum pruniforme	Sapotaceae	-2.916	1.466	0.05	1	
Cinchona calisaya	Rubiaceae	-1.601	0.309	0.00	1	
Cleistanthus libericus	Phyllanthaceae	-5.497	1.583	0.00	1	

Cleistopholis patens	Annonaceae	-3.188	1.482	0.03	1	
Clematis smilacifolia	Ranunculaceae	-4.201	1.597	0.01	1	
Coccocypselum herbaceum	Rubiaceae	-3.832	1.528	0.01	1	
Cochlospermum vitifolium	Bixaceae	-3.505	1.304	0.01	1	
Cola chlamydantha	Malvaceae	-2.926	0.826	0.00	2	
Cola nitida	Malvaceae	-4.920	1.650	0.00	1	
Cola verticillata	Malvaceae	-2.916	1.466	0.05	1	
Commelina diffusa	Commelinaceae	-3.797	1.525	0.01	2	LC
Cordyline fruticosa	Asparagaceae	-3.238	1.497	0.03	1	
Corynanthe pachyceras	Rubiaceae	-3.505	1.506	0.02	1	
Costus pulverulentus	Costaceae	-3.981	1.542	0.01	1	
Costus ricus	Costaceae	-3.347	1.483	0.02	1	LC
Coula edulis	Olacaceae	-3.060	1.474	0.04	1	
Craterispermum montanum	Rubiaceae	-4.646	1.109	0.00	1	VU
Croton stellulifer	Euphorbiaceae	-4.291	1.133	0.00	1	VU
Cryptocarya crassinerviopsis	Lauraceae	-4.335	1.543	0.00	1	
Cuphea hyssopifolia	Lythraceae	-4.526	1.601	0.00	1	
Cyclanthus bipartitus	Cyclanthaceae	-3.545	0.830	0.00	1	
Cynometra ananta	Fabaceae	-5.024	1.662	0.00	1	
Dacryodes klaineana	Burseraceae	-5.364	1.702	0.00	1	
Daemonorops sabut	Arecaceae	-2.891	1.463	0.05	1	
Dalbergia retusa	Fabaceae	-2.009	0.835	0.02	1	VU
Dalechampia cissifolia	Euphorbiaceae	-1.932	0.647	0.00	1	
Daniellia thurifera	Fabaceae	-4.236	1.575	0.01	1	
Dendropanax arboreus	Araliaceae	-3.342	1.472	0.02	1	
Desmodium incanum	Fabaceae	-1.107	0.342	0.00	2	
Dialium aubrevillei	Fabaceae	-3.650	1.484	0.03	1	
Diastema affine	Gesneriaceae	-3.076	1.463	0.04	1	
Dichorisandra amabilis	Commelinaceae	-3.566	1.502	0.02	1	
Dioscorea alata	Dioscoreaceae	-0.902	0.425	0.03	2	
Diospyros gabunensis	Ebenaceae	-3.658	1.427	0.02	1	
Diospyros kamerunensis	Ebenaceae	-5.307	1.695	0.00	1	
Diospyros sanza-minika	Ebenaceae	-3.325	1.345	0.03	1	
Dipteryx oleifera	Fabaceae	-1.978	0.299	0.00	1	
Discoglypremna caloneura	Euphorbiaceae	-3.959	1.547	0.01	1	
Dorstenia choconiana	Moraceae	-4.173	1.562	0.01	1	
Dracaena angustifolia	Asparagaceae	-2.145	0.821	0.01	3	

Drypetes aubrevillei	Putranjivaceae	-2.916	1.466	0.05	1	
Drypetes aylmeri	Putranjivaceae	-3.655	1.409	0.02	1	
Drypetes glabra	Putranjivaceae	-4.763	1.013	0.00	1	VU
Drypetes principum	Putranjivaceae	-3.402	1.434	0.03	1	
Dysopsis glechomoides	Euphorbiaceae	-2.298	0.549	0.00	2	
Dysoxylum densiflorum	Meliaceae	-4.335	1.543	0.00	1	
Elaeocarpus angustifolius	Elaeocarpaceae	-3.285	1.483	0.03	1	
Elaeocarpus musseri	Elaeocarpaceae	-3.362	0.848	0.00	1	
Entandrophragma angolense	Meliaceae	-3.188	1.482	0.03	1	VU
Entandrophragma cylindricum	Meliaceae	-2.916	1.466	0.05	1	VU
Epidendrum polyanthum	Orchidaceae	-3.576	1.498	0.02	1	
Epiphyllum phyllanthus	Cactaceae	-0.557	0.262	0.03	1	LC
Eremospatha wendlandiana	Arecaceae	-2.525	1.016	0.02	1	
Erigeron irazuensis	Asteraceae	-3.461	1.493	0.02	1	
Erythrina fusca	Fabaceae	-3.026	1.480	0.04	1	
Eugenia acapulcensis	Myrtaceae	-3.569	1.490	0.02	1	
Fagraea racemosa	Gentianaceae	-3.310	1.526	0.03	2	
Ficus chlamydocarpa	Moraceae	-1.184	0.530	0.03	1	
Ficus maxima	Moraceae	-3.972	1.526	0.01	1	
Fimbristylis complanata	Cyperaceae	-1.250	0.553	0.02	1	LC
Funtumia africana	Apocynaceae	-2.310	0.622	0.00	2	
Funtumia elastica	Apocynaceae	-2.899	1.450	0.05	2	
Gasteranthus acropodus	Gesneriaceae	-4.504	1.598	0.00	1	
Genipa americana	Rubiaceae	-3.342	1.472	0.02	1	
Geonoma cuneata	Arecaceae	-3.021	1.051	0.01	1	
Goniothalamus brevicuspis	Annonaceae	-4.079	1.524	0.01	1	
Goniothalamus majestatis	Annonaceae	-2.956	1.478	0.05	1	VU
Greenwayodendron oliveri	Annonaceae	-4.418	1.121	0.00	2	
Guazuma ulmifolia	Malvaceae	-3.221	1.116	0.00	1	
Harungana madagascariensis	Clusiaceae	-2.008	0.566	0.00	1	
Helichrysum panduratum	Asteraceae	-2.863	1.457	0.05	1	
Heliocarpus appendiculatus	Malvaceae	-2.946	1.446	0.04	1	
Heritiera utilis	Malvaceae	-4.893	1.647	0.00	1	VU
Heteropterys panamensis	Malpighiaceae	-4.015	1.546	0.01	1	
Hexalobus crispiflorus	Annonaceae	-3.050	1.166	0.02	1	
Homalium henriquesii	Salicaceae	-4.863	1.101	0.00	1	LR/NT
Homalomena erythropus	Araceae	-1.750	0.833	0.04	1	

Homalomena pendula	Araceae	-3.590	1.531	0.02	1	
Ipomoea batatoides	Convolvulaceae	-3.707	1.515	0.01	1	
Jatropha curcas	Euphorbiaceae	-3.342	1.472	0.02	1	
Klainedoxa gabonensis	Irvingiaceae	-3.060	1.474	0.04	1	
Koilodepas brevipes	Euphorbiaceae	-2.922	1.466	0.05	1	
Lasiodiscus mildbraedii	Rhamnaceae	-3.651	1.445	0.01	1	
Leandra mexicana	Melastomataceae	-3.832	1.528	0.01	1	
Leptaulus daphnoides	Icacinaceae	-3.827	1.534	0.01	1	
Lithocarpus celebicus	Fagaceae	-2.401	0.693	0.00	1	
Litsea diversifolia	Lauraceae	-3.541	1.492	0.02	1	
Litsea ferruginea	Lauraceae	-3.750	1.503	0.01	1	
Litsea formanii	Lauraceae	-3.750	1.503	0.01	1	
Litsea timoriana	Lauraceae	-3.541	1.492	0.02	1	
Lonchocarpus guatemalensis	Fabaceae	-2.683	1.109	0.02	1	LC
Lophira alata	Ochnaceae	-1.715	0.665	0.01	2	VU
Macaranga barteri	Euphorbiaceae	-3.188	1.482	0.03	1	
Macrocnemum roseum	Rubiaceae	-3.347	1.483	0.02	1	
Maesobotrya glabrata	Phyllanthaceae	-4.949	1.183	0.00	1	
Maesopsis eminii	Rhamnaceae	-3.782	1.452	0.01	1	
Mammea africana	Calophyllaceae	-2.516	0.822	0.00	2	
Mapania cuspidata	Cyperaceae	-1.909	0.812	0.02	1	
Mareya micrantha	Euphorbiaceae	-3.188	1.482	0.03	1	
Mendoncia gracilis	Acanthaceae	-3.406	1.488	0.02	1	
Miconia donaeana	Melastomataceae	-1.411	0.662	0.03	1	
Microdesmis puberula	Pandaceae	-3.026	0.643	0.00	2	
Mikania guaco	Asteraceae	-4.143	1.559	0.01	1	
Mikania sylvatica	Asteraceae	-4.113	1.556	0.01	1	
Mitrephora celebica	Annonaceae	-2.956	1.478	0.05	1	
Monodora myristica	Annonaceae	-3.754	1.527	0.01	1	
Musanga cecropioides	Urticaceae	-3.188	1.482	0.03	1	
Myrianthus arboreus	Urticaceae	-3.595	1.513	0.02	1	
Myrianthus libericus	Urticaceae	-3.190	1.310	0.03	1	
Napoleonaea vogelii	Napoleonaceae	-3.304	1.490	0.03	1	
Nectandra salicifolia	Lauraceae	-2.490	0.827	0.00	1	
Neonauclea intercontinentalis	Rubiaceae	-1.917	0.887	0.05	1	
Neurolaena lobata	Asteraceae	-2.961	0.823	0.00	1	
Ochna membranacea	Ochnaceae	-2.888	1.421	0.04	1	

Oreopanax geminatus	Araliaceae	-2.435	1.143	0.04	1	
Ornithocephalus inflexus	Orchidaceae	-4.158	1.539	0.01	1	
Palaquium quercifolium	Sapotaceae	-5.545	1.653	0.00	1	
Palmorchis trilobulata	Orchidaceae	-3.832	1.528	0.01	1	LC
Peperomia glabella	Piperaceae	-4.286	1.574	0.01	1	
Peperomia pseudoalpina	Piperaceae	-5.622	1.680	0.00	1	
Peperomia quadrifolia	Piperaceae	-1.953	0.923	0.04	1	
Petersianthus macrocarpus	Lecythidaceae	-3.304	1.490	0.03	1	
Philodendron aurantiifolium	Araceae	-1.661	0.653	0.01	1	
Philodendron burgeri	Araceae	-4.113	1.556	0.01	1	
Philodendron hederaceum	Araceae	-3.076	1.463	0.04	1	
Philodendron inaequilaterum	Araceae	-4.173	1.562	0.01	1	
Philodendron rhodoaxis	Araceae	-3.566	1.502	0.02	1	
Philodendron sulcatum	Araceae	-3.459	1.450	0.02	1	
Philodendron verrucosum	Araceae	-4.778	1.630	0.00	1	
Phyllanthus limmuensis	Phyllanthaceae	-2.754	0.438	0.00	1	
Phyllocosmus africanus	Ixonanthaceae	-3.677	1.521	0.02	1	
Pilea acuminata	Urticaceae	-5.244	1.684	0.00	1	
Pinus oocarpa	Pinaceae	-1.076	0.527	0.04	1	LC
Piper aequale	Piperaceae	-4.630	1.613	0.00	1	
Piper augustum	Piperaceae	-3.515	1.498	0.02	1	
Piper cenocladum	Piperaceae	-3.750	1.520	0.01	1	
Piper deductum	Piperaceae	-1.518	0.658	0.02	1	
Piper garagaranum	Piperaceae	-4.143	1.559	0.01	1	
Piper nudifolium	Piperaceae	-3.757	0.833	0.00	1	
Piper phytolaccifolium	Piperaceae	-4.894	1.643	0.00	1	
Piper sagittifolium	Piperaceae	-3.707	1.515	0.01	1	
Piptadeniastrum africanum	Fabaceae	-4.133	1.564	0.01	1	
Platymiscium dimorphandrum	Fabaceae	-3.760	1.506	0.01	1	LC
Poikilacanthus macranthus	Acanthaceae	-3.981	1.542	0.01	1	
Polyscias quintasii	Araliaceae	-2.602	0.846	0.00	1	EN
Prosthechea cochleata	Orchidaceae	-1.472	0.402	0.00	1	
Prosthechea radiata	Orchidaceae	-3.636	1.502	0.02	1	
Protomegabaria stapfiana	Phyllanthaceae	-5.227	1.686	0.00	1	
Pseudagrostistachys africana	Euphorbiaceae	-6.016	1.628	0.00	1	
Psidium guajava	Myrtaceae	-1.291	0.268	0.00	5	
Psychotria cuspidata	Rubiaceae	-5.109	1.668	0.00	1	

Psychotria elata	Rubiaceae	-3.071	0.639	0.00	1	
Psychotria solitudinum	Rubiaceae	-3.150	1.468	0.03	1	
Psychotria venosa	Rubiaceae	-2.000	0.589	0.00	1	
Psydrax subcordata	Rubiaceae	-4.671	1.514	0.00	1	
Pycnanthus angolensis	Myristicaceae	-2.121	0.337	0.00	2	
Randia armata	Rubiaceae	-2.946	1.446	0.04	1	
Rauvolfia vomitoria	Apocynaceae	-4.585	1.508	0.00	2	
Razisea spicata	Acanthaceae	-3.515	1.498	0.02	1	
Restrepiella ophiocephala	Orchidaceae	-3.904	1.520	0.01	1	
Rhodospatha osaensis	Araceae	-4.706	1.621	0.00	1	
Rhodospatha wendlandii	Araceae	-3.461	1.406	0.02	1	
Richardia scabra	Rubiaceae	-3.076	1.463	0.04	1	
Ricinodendron heudelotii	Euphorbiaceae	-3.595	1.513	0.02	1	
Rothmannia hispida	Rubiaceae	-1.811	0.839	0.03	2	
Ruellia pittieri	Acanthaceae	-4.813	1.633	0.00	1	
Santiria trimera	Burseraceae	-4.488	0.905	0.00	1	
Schefflera mannii	Araliaceae	-3.333	1.430	0.02	1	VU
Schistocarpha eupatorioides	Asteraceae	-3.615	1.507	0.02	1	
Schwackaea cupheoides	Melastomataceae	-4.242	0.841	0.00	1	
Scottellia klaineana	Achariaceae	-4.776	1.634	0.00	1	
Scytopetalum klaineanum	Lecythidaceae	-4.188	1.111	0.00	1	
Scytopetalum tieghemii	Lecythidaceae	-3.792	1.177	0.01	1	
Securidaca diversifolia	Polygalaceae	-3.150	1.468	0.03	1	
Selenicereus hamatus	Cactaceae	-3.093	1.478	0.04	1	LC
Semecarpus forstenii	Anacardiaceae	-4.079	1.524	0.01	1	
Shirakiopsis elliptica	Euphorbiaceae	-1.661	0.516	0.00	2	
Sinclairia polyantha	Asteraceae	-0.932	0.176	0.00	1	
Spatholobus ferrugineus	Fabaceae	-1.685	0.814	0.04	1	
Spatholobus gyrocarpus	Fabaceae	-0.955	0.440	0.03	1	LC
Spermacoce ovalifolia	Rubiaceae	-1.790	0.488	0.00	1	
Spondias purpurea	Anacardiaceae	-2.311	0.647	0.00	1	
Stanhopea oculata	Orchidaceae	-4.366	1.557	0.01	1	
Stenospermation angustifolium	Araceae	-2.998	1.457	0.04	1	
Sterculia longifolia	Malvaceae	-3.285	1.483	0.03	1	
Sterculia oblonga	Malvaceae	-3.188	1.482	0.03	1	VU
Sterculia oblongata	Malvaceae	-3.401	1.493	0.02	2	
Sterculia rhinopetala	Malvaceae	-4.459	1.599	0.01	1	

Strephonema pseudocola	Combretaceae	-3.505	1.506	0.02	1	
Strychnos ignatii	Loganiaceae	-1.396	0.638	0.03	1	
Symphonia globulifera	Clusiaceae	-4.111	0.939	0.00	3	
Syngonium hastiferum	Araceae	-4.173	1.562	0.01	1	
Syzygium acuminatissimum	Myrtaceae	-4.335	1.543	0.00	1	
Syzygium jambos	Myrtaceae	-3.619	1.484	0.01	1	
Tabernaemontana stenosiphon	Apocynaceae	-4.879	1.210	0.00	1	LR/NT
Tassadia obovata	Apocynaceae	-1.995	0.826	0.02	1	
Terminalia oblonga	Combretaceae	-3.151	1.156	0.01	1	
Tillandsia fasciculata	Bromeliaceae	-0.334	0.146	0.02	1	LC
Tillandsia ghiesbreghtii	Bromeliaceae	-5.020	1.618	0.00	1	
Tillandsia gymnobotrya	Bromeliaceae	-3.881	1.518	0.01	1	
Tillandsia usneoides	Bromeliaceae	-0.646	0.120	0.00	1	
Tillandsia viridiflora	Bromeliaceae	-6.227	1.745	0.00	1	
Trichilia grandifolia	Meliaceae	-2.371	0.855	0.01	1	LR/NT
Trichilia monadelpha	Meliaceae	-4.611	1.615	0.00	1	
Trichilia prieuriana	Meliaceae	-3.102	1.378	0.04	1	
Trichocentrum stramineum	Orchidaceae	-5.365	1.653	0.00	1	
Trilepisium madagascariense	Moraceae	-0.789	0.273	0.01	2	
Turpinia occidentalis	Staphyleaceae	-3.498	1.484	0.02	1	
Turraeanthus africanus	Meliaceae	-2.916	1.466	0.05	1	
Uncaria lanosa	Rubiaceae	-1.674	0.814	0.04	2	
Valeriana scandens	Valerianaceae	-1.548	0.493	0.00	1	
Ventilago oblongifolia	Rhamnaceae	-1.757	0.813	0.03	1	
Vepris dainellii	Rutaceae	-0.699	0.238	0.01	1	
Vitis tiliifolia	Vitaceae	-1.704	0.652	0.01	1	
Xylopia quintasii	Annonaceae	-3.677	1.521	0.02	1	
Zuelania guidonia	Salicaceae	-1.405	0.575	0.01	1	

Table D.2: List of species with significant positive log-response ratios – ln(RR)– for the Hum/Nat land-use treatment analysis, and considered arid and seasonal according to their climatic affiliations (see Figure 4.2). se, standard errors of ln(RR); N, number of studies used to calculate the log-response ratios; IUCN RL, the IUCN Red List of threatened species status for each species.

Species	Family	ln(RR)	se	sig.	N	IUCN RL
Acacia mearnsii	Fabaceae	2.042	0.67	0.003	1	
Acer pseudoplatanus	Sapindaceae	2.06	1.04	0.048	2	
Achillea millefolium	Asteraceae	1.197	0.161	0	5	LC

Agrostis canina	Poaceae	1.662	0.615	0.008	1	LC
Agrostis castellana	Poaceae	3.41	1.001	0.001	2	
Agrostis curtisii	Poaceae	2.005	0.444	0	2	
Agrostis glabra	Poaceae	3.12	1.477	0.035	1	
Agrostis inconspicua	Poaceae	2.661	0.451	0	2	
Agrostis pourretii	Poaceae	3.317	0.792	0	1	
Agrostis uliginosa	Poaceae	2.879	0.742	0	1	
Alchemilla erythropoda	Rosaceae	1.035	0.34	0.002	1	
Allium cepa	Amaryllidaceae	3.441	1.518	0.023	1	
Alopecurus geniculatus	Poaceae	0.767	0.305	0.012	1	
Alternanthera pungens	Amaranthaceae	4.982	1.572	0.002	1	
Alternanthera sessilis	Amaranthaceae	2.524	0.875	0.004	4	LC
Amaranthus hybridus	Amaranthaceae	2.484	0.403	0	2	
Amaranthus viridis	Amaranthaceae	4.763	1.554	0.002	1	
Angelica anomala	Apiaceae	1.342	0.662	0.043	1	
Anthyllis lotoides	Fabaceae	1.375	0.603	0.023	1	
Antimima solida	Aizoaceae	1.327	0.414	0.001	1	
Antimima watermeyeri	Aizoaceae	2.162	0.435	0	1	
Arrhenatherum album	Poaceae	4.364	1.603	0.006	1	
Asparagus aphyllus	Asparagaceae	2.486	0.917	0.007	1	LC
Aster squamatus	Asteraceae	1.661	0.77	0.034	1	
Astragalus pelecinus	Fabaceae	2.298	0.863	0.008	1	
Atriplex semibaccata	Chenopodiaceae	3.336	1.469	0.023	1	
Aulacomnium palustre	Aulacomniaceae	3.661	1.513	0.016	1	
Azorella lycopodioides	Apiaceae	5.131	1.651	0.002	2	
Azorella trifurcata	Apiaceae	2.067	0.794	0.009	2	
Berberis empetrifolia	Berberidaceae	0.461	0.125	0	2	
Berkheya rhapontica	Asteraceae	1.614	0.654	0.014	1	
Beta vulgaris	Chenopodiaceae	3.984	1.575	0.011	1	
Bidens bipinnata	Asteraceae	0.898	0.268	0.001	1	
Bothriochloa ischaemum	Poaceae	2.999	0.836	0	1	
Brachypodium phoenicoides	Poaceae	1.054	0.318	0.001	1	
Brachythecium rivulare	Brachytheciaceae	2.509	1.083	0.02	1	
Bromus squarrosus	Poaceae	2.456	0.607	0	1	
Bromus tectorum	Poaceae	0.771	0.351	0.028	3	
Calceolaria uniflora	Calceolariaceae	2.592	0.872	0.003	2	
Calluna vulgaris	Ericaceae	2.438	0.323	0	2	

Campanula rotundifolia	Campanulaceae	1.999	0.809	0.013	1	
Campylium stellatum	Amblystegiaceae	1.854	0.569	0.001	1	
Capsella bursa-pastoris	Brassicaceae	1.782	0.425	0	2	
Carex andina	Cyperaceae	1.985	0.989	0.046	1	
Carex argentina	Cyperaceae	3.043	1.394	0.03	1	
Carex caryophyllea	Cyperaceae	0.77	0.267	0.004	1	
Carex decidua	Cyperaceae	0.683	0.34	0.045	2	
Carex magellanica	Cyperaceae	1.284	0.285	0	2	LC
Celtis africana	Cannabaceae	1.908	0.573	0.001	2	
Centaurea stoebe	Asteraceae	1.376	0.429	0.001	1	
Cephalaria oblongifolia	Dipsacaceae	1.556	0.459	0.001	1	
Chaerophyllum aureum	Apiaceae	1.026	0.499	0.04	1	
Cheilanthes hirta	Pteridaceae	3.591	1.477	0.015	1	
Chenopodium album	Chenopodiaceae	1.944	0.904	0.031	2	
Chenopodium mucronatum	Chenopodiaceae	3.006	1.465	0.04	1	
Chlorophytum comosum	Asparagaceae	3.591	1.477	0.015	1	
Chuquiraga aurea	Asteraceae	3.068	1.448	0.035	1	
Cirsium vulgare	Asteraceae	1.335	0.291	0	4	
Cladanthus mixtus	Asteraceae	4.277	1.019	0	1	
Cleome monophylla	Cleomaceae	2.32	0.832	0.006	1	
Clinopodium darwinii	Lamiaceae	2.12	0.813	0.009	1	
Clutia pulchella	Euphorbiaceae	3.336	1.469	0.023	1	
Conophytum subfenestratum	Aizoaceae	3.441	1.505	0.022	1	
Convolvulus arvensis	Convolvulaceae	1.042	0.384	0.007	1	
Crataegus monogyna	Rosaceae	2.236	0.436	0	3	
Crepis vesicaria	Asteraceae	2.409	1.162	0.038	2	
Crotalaria sphaerocarpa	Fabaceae	3.8	1.487	0.011	1	
Cruciata laevipes	Rubiaceae	1.481	0.708	0.036	2	
Crupina vulgaris	Asteraceae	0.766	0.358	0.033	1	
Cucurbita pepo	Cucurbitaceae	4.076	1.533	0.008	2	
Cuscuta epithymum	Convolvulaceae	2.945	1.471	0.045	1	
Cynara humilis	Asteraceae	3.587	1.522	0.018	1	
Cynodon incompletus	Poaceae	2.02	0.8	0.013	1	
Cynosurus echinatus	Poaceae	3.122	1.307	0.017	2	
Cyperus alternifolius	Cyperaceae	1.247	0.624	0.046	1	LC
Cyperus rotundus	Cyperaceae	6.613	1.72	0	1	LC
Cyphia oligotricha	Campanulaceae	2.878	1.461	0.049	1	

Daboecia cantabrica	Ericaceae	2.512	1.164	0.041	1	
Datura stramonium	Solanaceae	3.675	0.554	0	1	
Deschampsia kingii	Poaceae	1.58	0.393	0	2	
Dichondra micrantha	Convolvulaceae	3.8	1.487	0.011	1	
Diclis reptans	Scrophulariaceae	3.966	1.354	0.003	2	
Dicranoweisia cirrata	Dicranaceae	2.51	0.44	0	1	
Digitaria sanguinalis	Poaceae	3.972	1.061	0	1	
Digitaria tricholaenoides	Poaceae	1.566	0.644	0.017	2	
Dittrichia viscosa	Asteraceae	4.052	1.569	0.01	1	
Draba funiculosa	Brassicaceae	1.251	0.311	0	1	
Drosanthemum ramosissimum	Aizoaceae	1.722	0.356	0	1	
Drosanthemum schoenlandianum	Aizoaceae	1.04	0.509	0.046	1	
Dysphania carinata	Chenopodiaceae	3.8	1.487	0.011	1	
Elymus magellanicus	Poaceae	2.177	0.678	0.001	2	
Ephedra frustillata	Ephedraceae	1.979	0.814	0.015	1	LC
Epilobium australe	Onagraceae	0.289	0.102	0.005	2	
Equisetum arvense	Equisetaceae	3.999	1.586	0.012	1	LC
Eragrostis biflora	Poaceae	5.323	1.601	0.001	1	
Eragrostis lehmanniana	Poaceae	1.151	0.499	0.021	1	
Eragrostis plana	Poaceae	0.675	0.329	0.043	2	
Erigeron patagonicus	Asteraceae	1.441	0.451	0.001	2	
Erodium cicutarium	Geraniaceae	1.959	0.414	0	2	
Eruca vesicaria	Brassicaceae	3.457	0.651	0	2	
Eryngium campestre	Apiaceae	1.038	0.287	0	1	
Eucalyptus globulus	Myrtaceae	4.179	1.623	0.013	2	
Euphorbia epicyparissias	Euphorbiaceae	1.537	0.776	0.048	1	
Euphorbia prostrata	Euphorbiaceae	3.499	1.464	0.017	2	
Euphrasia stricta	Orobanchaceae	1.092	0.431	0.011	1	
Felicia filifolia	Asteraceae	2.476	0.434	0	2	
Festuca dalmatica	Poaceae	1.273	0.32	0	1	
Festuca magellanica	Poaceae	1.66	0.736	0.025	2	
Filago arvensis	Asteraceae	3.365	1.464	0.022	1	
Filago gallica	Asteraceae	1.791	0.874	0.04	1	
Fimbristylis dichotoma	Cyperaceae	3.338	0.676	0	2	LC
Fuirena pubescens	Cyperaceae	2.502	0.578	0	2	LC
Galenia sarcophylla	Aizoaceae	1.833	0.646	0.005	1	
Gazania tenuifolia	Asteraceae	2.279	0.64	0	1	

Genista falcata	Fabaceae	3.809	1.546	0.014	1	
Geranium magellanicum	Geraniaceae	2.739	1.151	0.017	2	
Gerbera aurantiaca	Asteraceae	1.951	0.729	0.007	1	
Gomphrena celosioides	Amaranthaceae	2.827	0.674	0	1	
Grimmia alpestris	Grimmiaceae	4.434	1.592	0.005	1	
Grimmia incurva	Grimmiaceae	4.138	1.561	0.008	1	
Guilleminea densa	Amaranthaceae	3.591	1.477	0.015	1	
Halimium lasianthum	Cistaceae	1.442	0.233	0	1	
Helichrysum acutatum	Asteraceae	4.679	1.547	0.002	1	
Helichrysum adenocarpum	Asteraceae	4.804	1.396	0.001	2	
Helictotrichon imberbe	Poaceae	1.322	0.318	0	1	
Hibiscus trionum	Malvaceae	1.775	0.566	0.002	1	
Hordeum murinum	Poaceae	3.248	1.054	0.002	3	LC
Hordeum secalinum	Poaceae	1.693	0.605	0.006	1	LC
Hymenoloma crispulum	Rhabdoweisiaceae	2.659	0.942	0.005	1	
Hyparrhenia dregeana	Poaceae	5.515	1.728	0.001	1	
Hyparrhenia hirta	Poaceae	0.809	0.379	0.035	3	
Hypochaeris glabra	Asteraceae	1.513	0.47	0.001	3	
Hypochaeris incana	Asteraceae	2.252	0.878	0.011	2	
Ifloga glomerata	Asteraceae	4.262	1.516	0.005	1	
Ipomoea purpurea	Convolvulaceae	2.404	0.591	0	1	
Koeleria macrantha	Poaceae	0.797	0.175	0	1	
Larix kaempferi	Pinaceae	2.275	0.825	0.006	1	LC
Leontodon crispus	Asteraceae	1.299	0.242	0	1	
Leskea polycarpa	Leskeaceae	3.595	1.507	0.017	1	
Leucanthemum vulgare	Asteraceae	0.859	0.267	0.001	4	
Ligustrum lucidum	Oleaceae	3.976	1.497	0.008	1	
Linum catharticum	Linaceae	1.019	0.431	0.018	2	
Lobelia oligophylla	Campanulaceae	1.22	0.35	0	2	
Lotononis falcata	Fabaceae	3.162	1.481	0.033	1	
Luzula chilensis	Juncaceae	0.68	0.336	0.043	2	
Lycium ameghinoi	Solanaceae	3.571	1.518	0.019	1	
Medicago falcata	Fabaceae	1.625	0.445	0	1	
Melinis repens	Poaceae	1.471	0.571	0.012	1	
Nassauvia aculeata	Asteraceae	3.357	1.498	0.025	1	
Nassauvia glomerulosa	Asteraceae	3.031	1.236	0.015	1	
Nassauvia ulicina	Asteraceae	2.238	1.034	0.031	1	

Noccaea magellanica	Brassicaceae	0.417	0.206	0.043	2	
Olea europaea	Oleaceae	2.215	0.829	0.008	4	
Ornithopus compressus	Fabaceae	2.136	0.687	0.002	1	
Orthotrichum macrocephalum	Orthotrichaceae	1.469	0.514	0.005	2	
Orthotrichum tenellum	Orthotrichaceae	1.874	0.374	0	2	
Oryzopsis miliacea	Poaceae	3.96	1.559	0.011	1	
Osmorhiza depauperata	Apiaceae	1.38	0.534	0.01	2	
Oxalis purpurea	Oxalidaceae	3.336	1.469	0.023	1	
Perezia recurvata	Asteraceae	1.53	0.365	0	2	
Persicaria hydropiper	Polygonaceae	2.916	1.485	0.05	1	LC
Peucedanum oreoselinum	Apiaceae	1.911	0.505	0	1	
Phalaris arundinacea	Poaceae	3.483	1.532	0.023	1	LC
Philonotis tomentella	Bartramiaceae	2.517	1.219	0.041	1	
Phleum pratense	Poaceae	0.803	0.246	0.001	6	
Physalis viscosa	Solanaceae	1.614	0.654	0.014	1	
Pilosella officinarum	Asteraceae	1.002	0.257	0	5	
Pimpinella saxifraga	Apiaceae	2.745	1.06	0.01	1	
Pinus pinaster	Pinaceae	2.527	0.932	0.009	2	LC
Plantago coronopus	Plantaginaceae	2.13	0.767	0.006	3	
Plantago major	Plantaginaceae	3.231	1.458	0.027	2	LC
Poa pratensis	Poaceae	1.063	0.394	0.007	6	LC
Poa spiciformis	Poaceae	2.21	0.72	0.002	1	
Polytrichum piliferum	Polytrichaceae	2.02	1.029	0.05	1	
Portulaca oleracea	Portulacaceae	2.338	0.385	0	2	
Potentilla cinerea	Rosaceae	0.977	0.418	0.021	1	
Prosopis denudans	Fabaceae	3.863	1.547	0.013	1	
Prunus persica	Rosaceae	4.128	1.507	0.006	1	
Pseudoleskea incurvata	Leskeaceae	3.617	1.509	0.017	1	
Pseudotsuga menziesii	Pinaceae	1.622	0.578	0.031	1	LC
Ranunculus fuegianus	Ranunculaceae	0.778	0.293	0.008	1	
Rhinanthus rumelicus	Orobanchaceae	1.464	0.66	0.043	1	
Rhynchopsidium pumilum	Asteraceae	0.899	0.085	0	1	
Richardia brasiliensis	Rubiaceae	2.167	0.461	0	2	
Rubia peregrina	Rubiaceae	1.047	0.467	0.025	1	
Rubus pungens	Rosaceae	1.426	0.559	0.011	1	
Rubus rigidus	Rosaceae	3.006	1.465	0.04	1	
Rumex crispus	Polygonaceae	0.797	0.294	0.007	2	

Rumex obtusifolius	Polygonaceae	3.532	1.537	0.022	1	
Ruschia burtoniae	Aizoaceae	0.885	0.27	0.001	1	
Ruschia spinosa	Aizoaceae	4.649	1.632	0.004	1	
Saelania glaucescens	Ditrichaceae	2.572	1.142	0.024	1	
Sarcocornia xerophila	Chenopodiaceae	2.956	1.466	0.044	1	
Sarmentypnum exannulatum	Amblystegiaceae	2.496	0.664	0	1	
Schinus marchandii	Anacardiaceae	3.089	1.475	0.036	1	
Scleranthus perennis	Caryophyllaceae	2.471	0.846	0.004	1	
Sebaea leiostyla	Gentianaceae	1.681	0.522	0.001	1	
Sedum hispanicum	Crassulaceae	1.949	0.546	0	1	
Senecio acanthifolius	Asteraceae	1.209	0.209	0	2	
Senecio filaginoides	Asteraceae	1.476	0.725	0.043	1	
Sida dregei	Malvaceae	2.476	0.771	0.002	1	
Sida rhombifolia	Malvaceae	1.316	0.274	0	4	
Sideritis montana	Lamiaceae	1.324	0.502	0.008	1	
Solanum lycopersicum	Solanaceae	3.71	1.481	0.012	3	
Solanum nigrum	Solanaceae	3.108	1.449	0.032	2	
Sorghum halepense	Poaceae	6.023	1.665	0	1	
Sphagnum denticulatum	Sphagnaceae	3.107	1.467	0.034	1	
Sphagnum palustre	Sphagnaceae	3.844	1.531	0.012	1	
Sporobolus africanus	Poaceae	1.32	0.298	0	2	
Stellaria alsine	Caryophyllaceae	0.593	0.293	0.043	1	
Stipa ibarii	Poaceae	2.721	1.353	0.045	1	
Straminergon stramineum	Amblystegiaceae	3.315	1.483	0.025	1	
Taraxacum campylodes	Asteraceae	1.562	0.463	0.001	6	
Tetragonia fruticosa	Aizoaceae	1.392	0.642	0.035	1	
Teucrium chamaedrys	Lamiaceae	1.144	0.493	0.022	1	
Thymus longicaulis	Lamiaceae	0.812	0.274	0.004	1	
Tiarella polyphylla	Saxifragaceae	3.202	1.507	0.034	1	
Tolpis barbata	Asteraceae	1.186	0.339	0	1	
Tolpis capensis	Asteraceae	1.366	0.608	0.025	1	
Tortella flavovirens	Pottiaceae	3.738	0.429	0	1	
Tribulus terrestris	Zygophyllaceae	3.809	0.948	0	1	
Trifolium arvense	Fabaceae	0.669	0.293	0.024	4	
Trifolium campestre	Fabaceae	1.125	0.481	0.019	2	
Trifolium dubium	Fabaceae	1.271	0.351	0	3	
Trifolium pannonicum	Fabaceae	3.75	1.511	0.013	1	

Trifolium pratense	Fabaceae	1.486	0.294	0	3	LC
Trifolium repens	Fabaceae	1.553	0.198	0	6	
Trifolium striatum	Fabaceae	2.08	0.656	0.002	2	
Trisetum caudulatum	Poaceae	2.948	1.464	0.044	1	
Ulex europaeus	Fabaceae	3.125	1.175	0.01	2	LC
Uncinia lechleriana	Cyperaceae	0.85	0.131	0	2	
Urochloa panicoides	Poaceae	6.049	1.667	0	1	LC
Ursinia nana	Asteraceae	0.928	0.273	0.001	2	
Urtica magellanica	Urticaceae	2.099	0.672	0.002	1	
Vicia bijuga	Fabaceae	1.304	0.307	0	2	
Vicia sativa	Fabaceae	1.659	0.545	0.002	2	
Vulpia muralis	Poaceae	3.712	1.534	0.016	1	
Zantedeschia albomaculata	Araceae	3.006	1.465	0.04	1	LC
Ziziphus mucronata	Rhamnaceae	4.763	1.554	0.002	1	

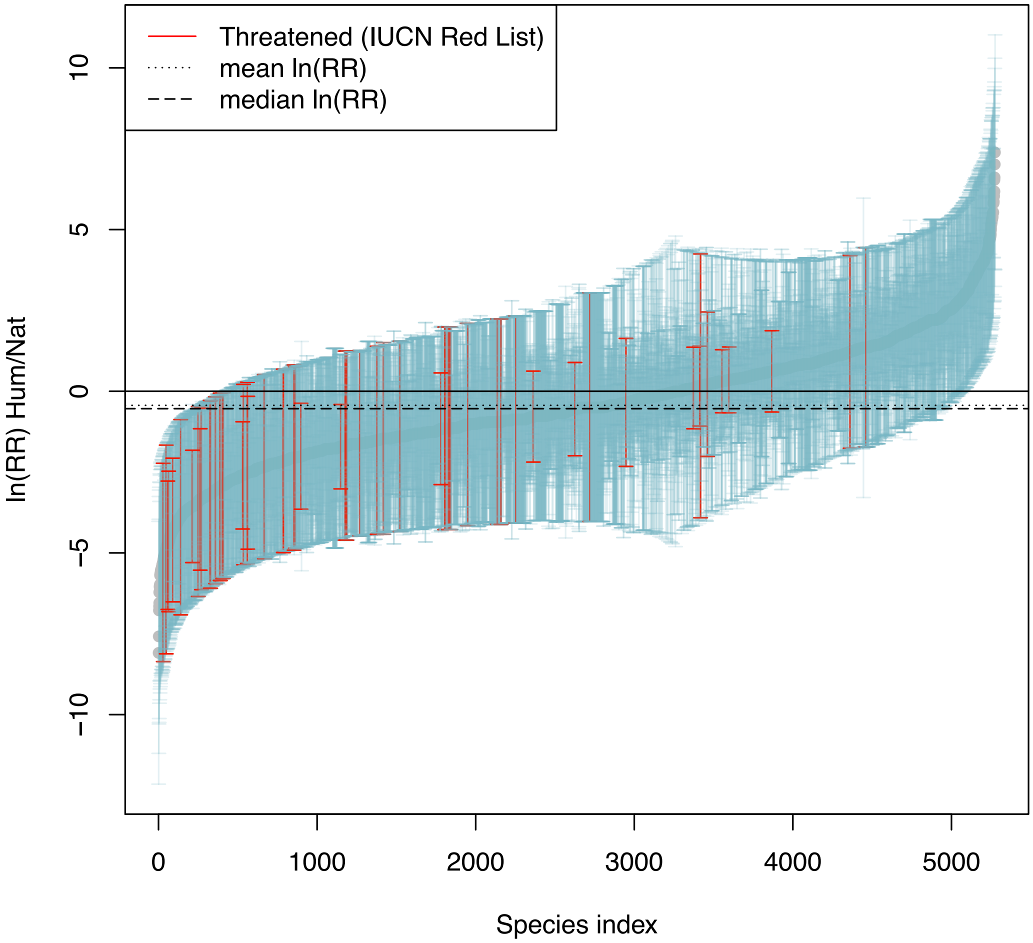


Figure D.1: Distribution of log-response ratios of the abundances of each of the species used under the Hum/Nat land-use treatment.

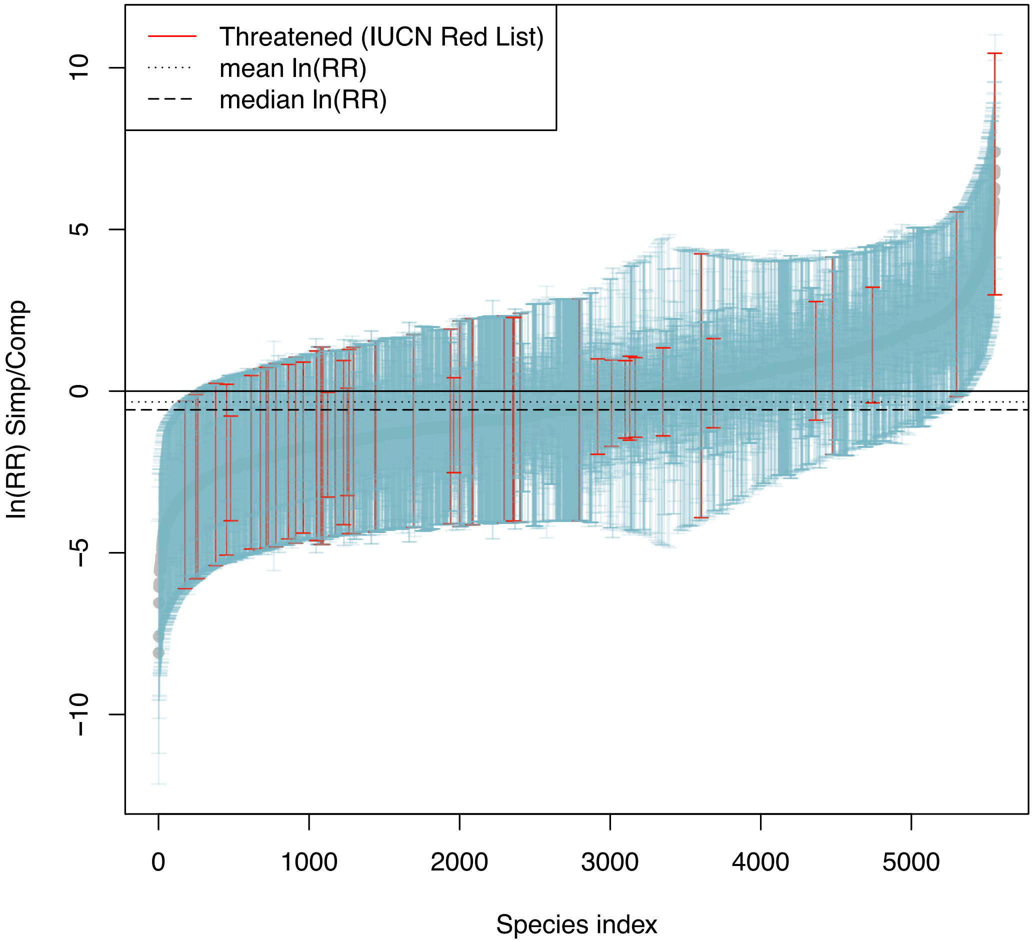


Figure D.2: Distribution of log-response ratio of the abundances of each of the species used under the Simp/Comp land-use treatment. The species marked as threatened with a significant positive log-response ratio (i.e., less sensitive to land use change) corresponds to *Afzelia xylocarpa* a tree species from South-East Asia. The abundance of this species increased significantly from Primary forest to Young secondary forests.

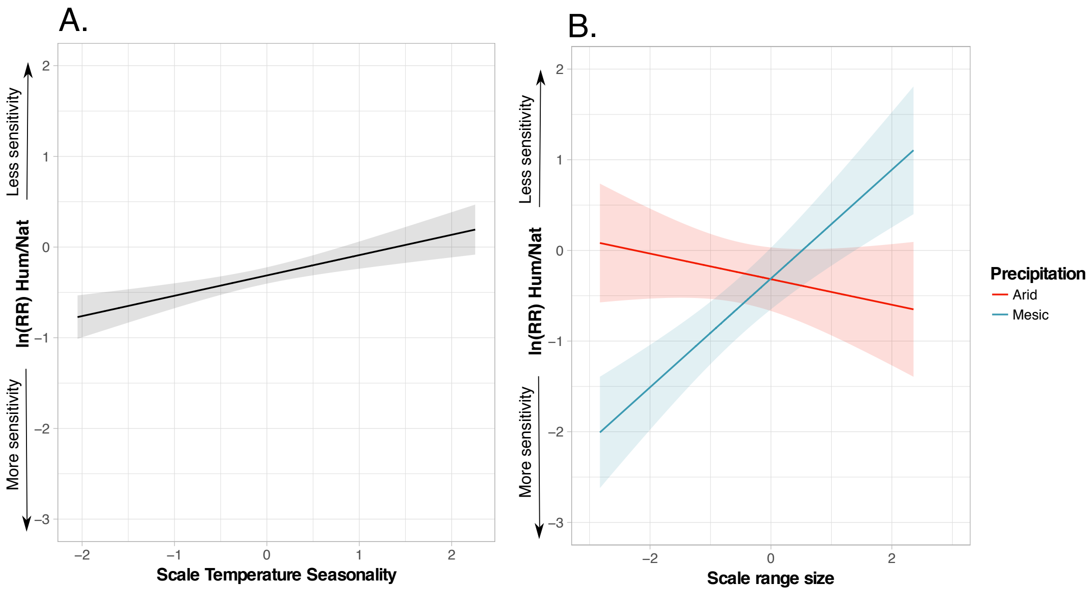


Figure D.3: Changes in the log-response ratios with (A) Temperature seasonality and the (B) interaction between range size and mean annual precipitation for the Hum/Nat land-use change treatment. The average value of seasonality was considered for the additive effect in (B). Error bars represent 95 % confidence intervals.

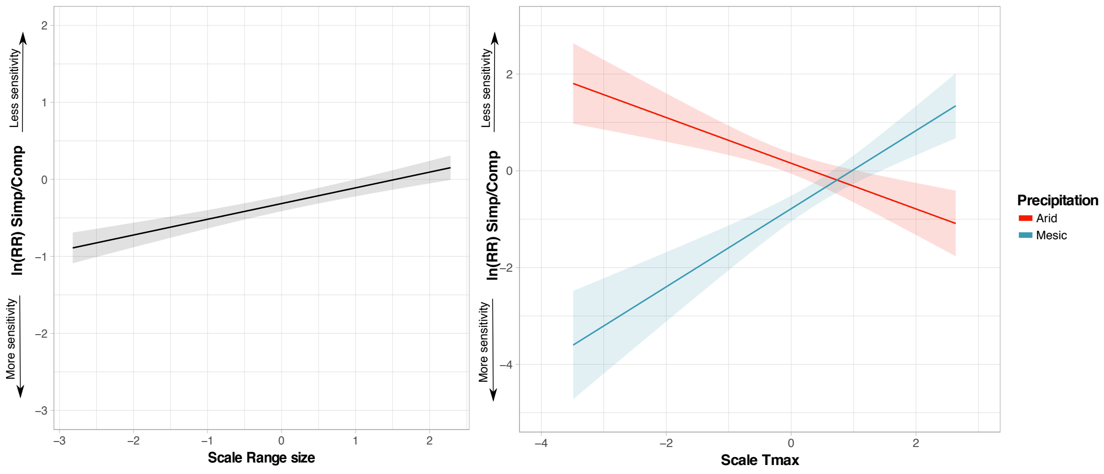


Figure D.4: Changes in log-response ratios with (A) Range size and the (B) interaction between mean annual precipitation and Tmax for the Simp/Comp land-use change treatment. The average value of range size was considered for the additive effect in (B). Error bars represent 95 % confidence intervals.

Term	d.f.	VIF	GVIF
Seasonality	1	1.49	1.22
Precipitation	1	1.79	1.34
Range size	1	1.07	1.03
Range size:Precipitation	1	1.31	1.15

Table D.3: Variance inflation factors (VIF) for the variables used to model the variation of log-response ratios under the Hum/Nat land use treatment. GVIF is the Generalized Variance Inflation Factor calculated as $VIF^{1/(2*d.f.)}$, which gives an indication of how much the standard errors are likely to be inflated due to collinearity between explanatory variables. VIF was calculated using the corvif function from (Zuur *et al.*, 2009)

Term	d.f.	VIF	GVIF
Tmax	1	1.25	1.12
Precipitation	1	1.19	1.09
Range size	1	1.00	1.00
Precipitation:Tmax	1	1.14	1.06

Table D.4: Variance inflation factors (VIF) for the variables used to model the variation of log-response ratios under the Simp/Comp land use treatment. GVIF is the Generalized Variance Inflation Factor calculated as $VIF^{1/(2*d.f.)}$, which gives an indication of how much the standard errors are likely to be inflated due to collinearity between explanatory variables. VIF was calculated using the corvif function from (Zuur *et al.*, 2009)

APPENDIX E

Supplement to Chapter 5

E.1 Data sources

Journals

- Biotropica
- Biodiversidad Neotropical
- Biota Colombiana
- Caldasia
- Acta Biológica Colombiana
- Revista de Biología Tropical
- Revista Colombiana de Ciencia Animal

Databases

- Universidad Nacional de Colombia library `http://www.sinab.unal.edu.co/`
- Universidad de Antioquia `http://biblioteca.udea.edu.co/`
- Universidad del Valle `http://biblioteca.univalle.edu.co/`
- Universidad Industrial de Santander `http://tangara.uis.edu.co/biblioweb/pags/cat/conbas.jsp`

E.2 Statistical Model Structure

We present the R formula for the models that best explain the data (i.e. after stepwise selection using AIC). Species richness (number of species) when compositional intactness using Sørensen similarity index (See Magurran, 2004) was used as response variables. Studies (SS), blocks within studies (SSB) were considered as random effects. Different factor levels of land use (LU) were considered as explanatory variables; see Figure 5.2 for LU classes used in compositional intactness.

Compositional Intactness $\leftarrow lmer(S\ddot{o}rensen \sim LU + (1|SS) + (1|SSB))$

Source	Species richness	Abundance	N sites	Taxa	MLE(m)	PV	MSV	ISV	YSV	Pas	Plan	Crop	Urb
				Andes									
Gutiérrez-Lemus *et al.* (2004)	8	179	2	Chordata	300	✓					✓		
	10	352	2	Chordata	300	✓					✓		
Armbrecht *et al.* (2006)	24	171	12	Arthropoda	94		✓					✓	
	29	136	12	Arthropoda	94		✓				✓		
	32	210	12	Arthropoda	94		✓				✓		
	19	133	12	Arthropoda	94		✓				✓		
Smith-Pardo & Gonzales (2007)	39	718	12	Arthropoda	424	✓			✓	✓			
	262	8,416	12	Arthropoda	424	✓			✓	✓			
	47	165	12	Arthropoda	424	✓			✓	✓			
	204	6,208	12	Arthropoda	424	✓			✓	✓			
Higuera & Wolf (2010)	17	(biomass)	6	Tracheophyta	42	✓			✓				
Arbeláez-Cortés *et al.* (2011)	75	671	3	Chordata	2,500		✓		✓				
Isaacs & Urbina-Cardona (2011)	8	251	6	Chordata		✓				✓			
Noriega *et al.* (2012)	13	779	3	Arthropoda	270	✓				✓	✓		
Poveda *et al.* (2012)	334	8,270	17	Arthropoda	23							✓	
	270	12,123	17	Arthropoda	23							✓	
Rey-Velasco & Miranda-Esquivel (2013)	21	997	7	Arthropoda	300	✓	✓			✓	✓		
Rosselli & Stiles (2012b,a)	115	(occurrence)	19	Chordata	774	✓							
Cabra-García *et al.* (2012)	290	6,765	39	Arthropoda	71	✓	✓			✓		✓	
				Orinoco									
Noriega *et al.* (2007)	22	2,358	3	Arthropoda	270		✓	✓	✓				
Parra & Nates-Parra (2007)	21	144	26	Arthropoda	193	✓		✓		✓		✓	✓
				Caribbean									
Dominguez & Armbrecht (2011)	82	4,219	13	Arthropoda	101		✓	✓	✓				
Navarro *et al.* (2011)	26	7,496	2	Arthropoda		✓				✓			
				Amazonian									
López-Quintero *et al.* (2012)	405	18,017	11	Basidiomycota Ascomycota	51	✓	✓	✓	✓			✓	
	619	836	10	Tracheophyta	52	✓	✓	✓	✓			✓	
Otavo *et al.* (2013)	92	593	3	Arthropoda	1,708		✓	✓	✓				
Total	2,582												

Table E.1: Description of the studies collated in the current research. Sources in grey were used in the compositional similarity analysis. Sources of information were divided where necessary into studies (defined as having sampled a list of taxa in a number of sites using the same sampling method). Species richness values correspond to the number of unique taxon names within each study; note that 2,582 is the number of unique taxon names overall data set, not the sum of the row values (which is 3,085) as some taxa are in multiple studies. Abundance values represent the summed of individuals of all species at a site (n.b. two studies reported measures other than counts of individuals, as indicated). Maximum linear extent (MLE) corresponds to the longest distance between any two sampling points in a site. PV= Primary vegetation, MSV= Mature secondary vegetation, ISV= Intermediate secondary vegetation, YSV= young secondary vegetation, Pas= Pastures, Plan= Plantation, Crop= Cropland, Urb= Urban.

Land Use	Minimum	Light	Intense	Cannot decide	Total
Primary vegetation	33	2	2	12	49
Mature secondary vegetation	31	1	1	0	33
Intermediate secondary vegetation	12	1	0	3	16
Young secondary vegetation	16	1	3	18	38
Pasture	5	3	9	10	27
Plantation forest	16	13	12	0	41
Cropland	0	34	8	8	50
Urban	0	0	0	11	11

Table E.2: Number of sites per combination of land-use LU and land-use intensity LUUI class for all the collated sources.

	d.f.	**VIF**	**GVIF**
LU	5	1.41	1.04
logHPD	1	1.46	1.21
logdistRd	1	1.75	1.32

Table E.3: Variance inflation factors (VIF) for the variables used to model site-level compositional intactness. GVIF is the generalized variance inflation factor calculated as VIF$^{1/(2*d.f.)}$, which gives an indication of how much the standard errors are likely to be in inflated due to collinearity between explanatory variables. VIF was calculated using the corvif function of Zuur *et al.* (2009).

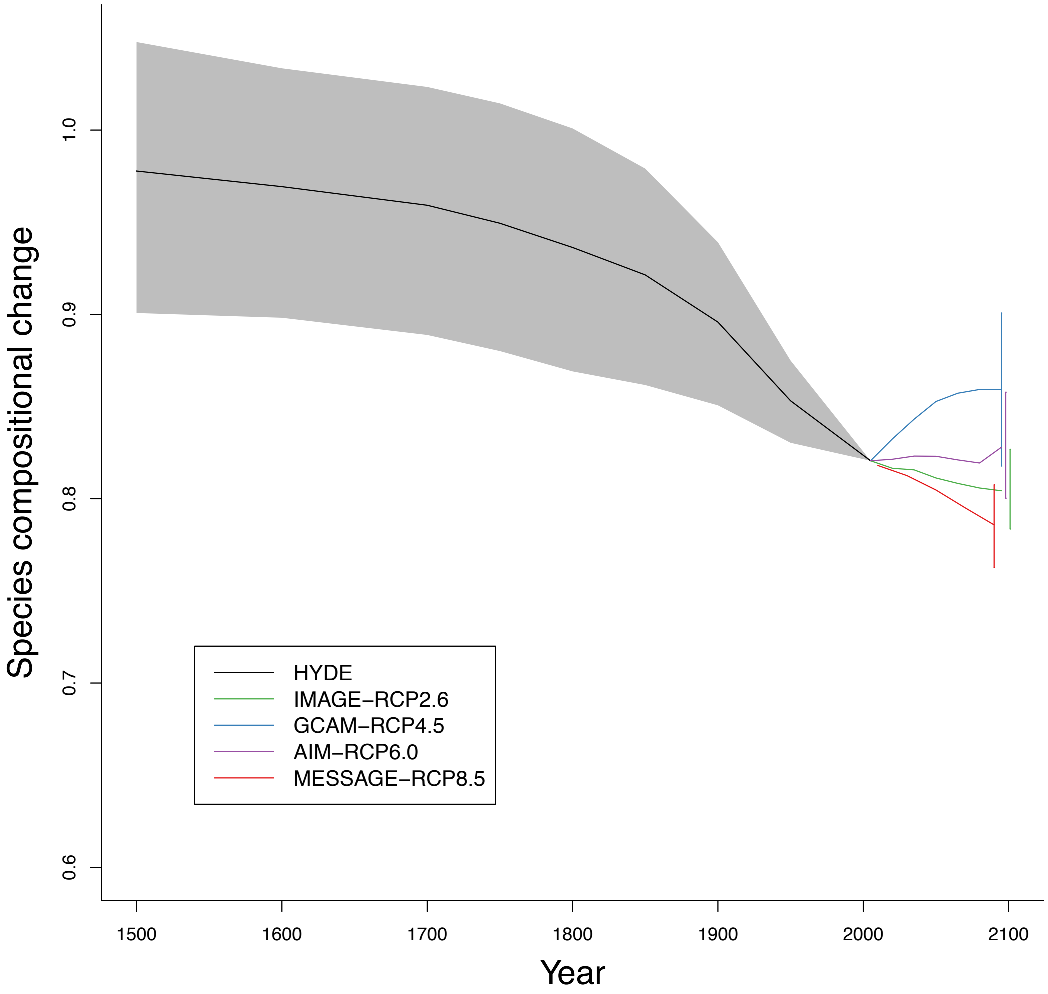


Figure E.1: Projected global net change in compositional intactness from 1500 to 2095 in Colombia. Grey shading (historical) and error bars (future) show ± 95% confidence intervals. Future projections are based on the four RCP scenarios (Table 5.2). Uncertainty estimates are based only on the modeled coefficients. Estimates of uncertainty were not available for the land-use projections.

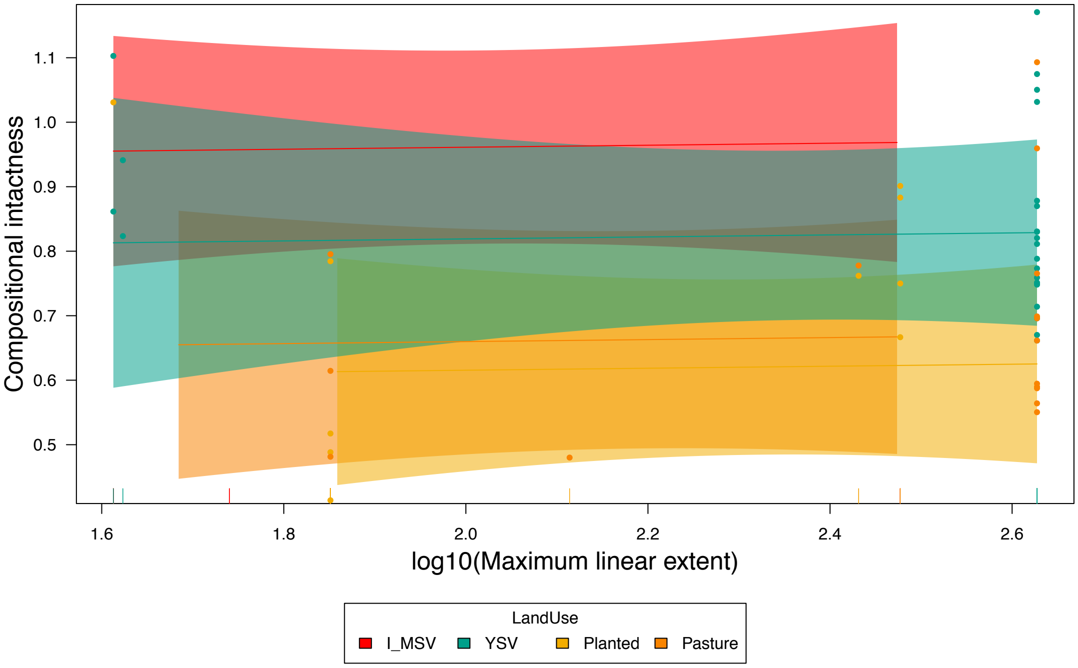


Figure E.2: Relationship between maximum linear extent (MLE) and study-level difference in compositional intactness. Shading indicates ± 95% confidence intervals. Rugs along the x axes in the line graphs show the values of the explanatory variables represented in the data set used for modelling. For clarity, data points are also included.

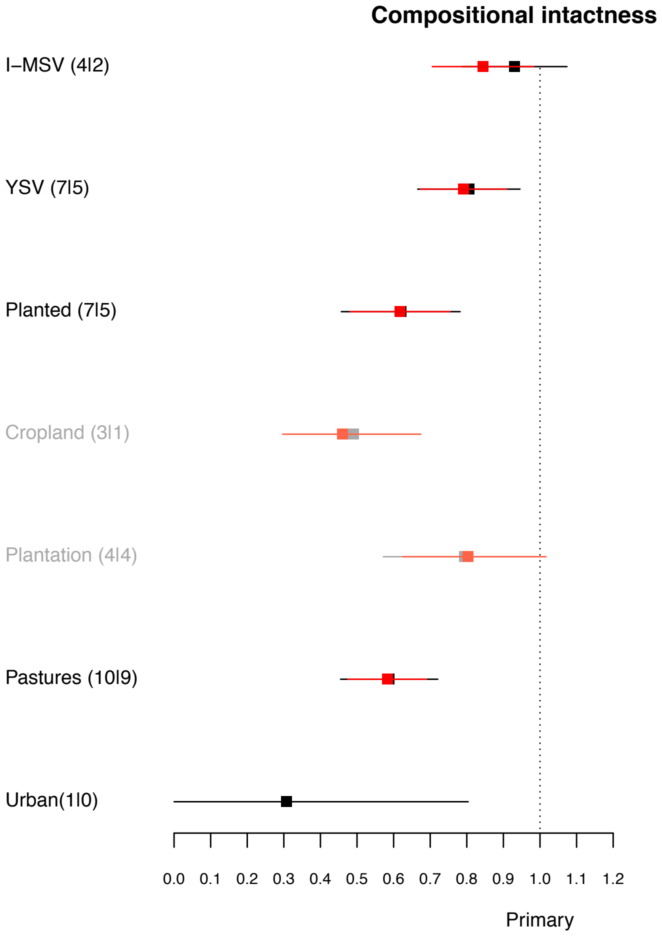


Figure E.3: Influence of studies with differences in site size on community intactness. Black and grey lines correspond to the original estimates (see Figure 5.2). Red lines represent the estimates when three studies (two in López-Quintero *et al.*, 2012, and one in Parra & Nates-Parra, 2007), with differences in maximum linear extent among sites, are removed. The main error bars show model coefficients estimates ± 95% CI. Values in brackets represent the number of studies used in the original vs the sensitivity analysis. YSV= Young secondary vegetation, I-MSV = Intermediate and mature secondary vegetation combined, Planted= Cropland and plantation.

APPENDIX F

Supplement to Chapter 6

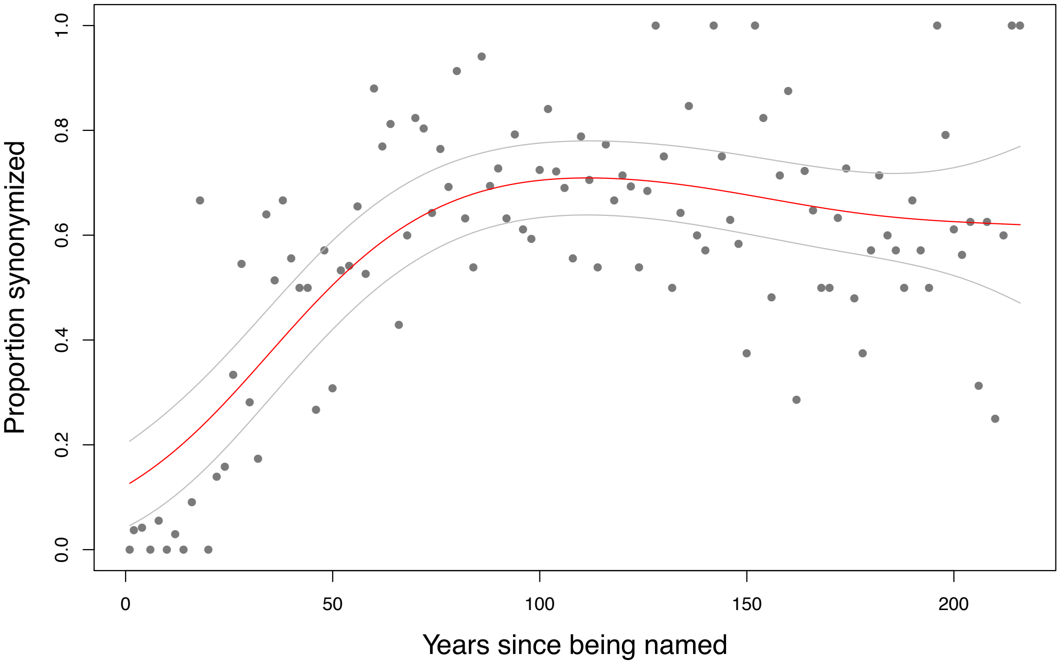


Figure F.1: Observed proportion of synonyms in the whole *Solanum* from Linnaeus first descriptions to 2016. This rate of synonyms provides a simplistic way to understand the level of taxonomic maturity in *Solanum* following the Solow *et al.* (1995) approach. In this plot, there is a tendency for the observed proportion to flatten out towards 0.6-0.7, which approximates to the current proportion of synonyms of *Solanums* 0.62 (acc=1211, n= 3133, syn=1923). The proportion of synonyms through time was modeled using Generalized Additive Model (gam) with binomial errors using the R package “mgcv” (Wood, 2012)

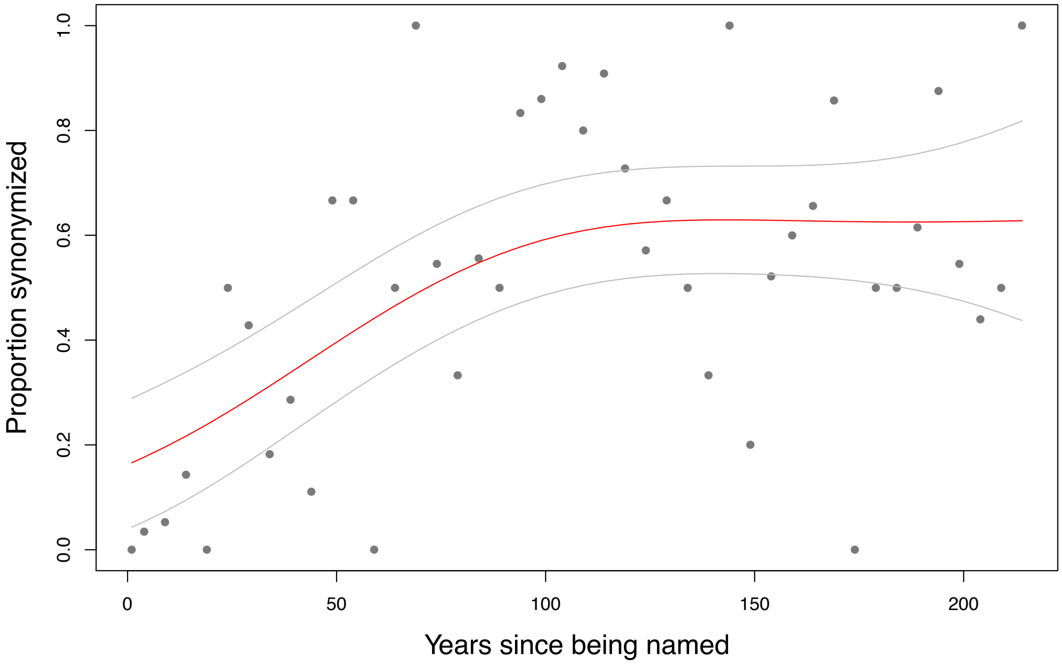


Figure F.2: Observed proportion of synonyms in the Old World clade of *Solanum* from Linnaeus first descriptions to 2016. As in Figure F.1, there is a tendency for the observed proportion to flatten out towards 0.6. The proportion of synonyms through time was modeled using Generalized Additive Model (gam) with binomial errors using the R package "mgcv" (Wood, 2012)