FUNCTIONAL AND PHYLOGENETIC DIMENSIONS OF TREE DIVERSITY ACROSS ENVIRONMENTAL GRADIENTS IN PUERTO RICO:

Insights to community assembly processes

Robert Muscarella

Submitted in partial fulfillment of the
requirements for the degree
of Doctor of Philosophy
in the Graduate School of Arts and Sciences

COLUMBIA UNIVERSITY

2015

ABSTRACT

FUNCTIONAL AND PHYLOGENETIC DIMENSIONS OF TREE DIVERSITY ACROSS ENVIRONMENTAL GRADIENTS IN PUERTO RICO:

Insights to community assembly processes

Robert Muscarella

One goal central to ecology is to understand how species interactions and biophysical processes interact over vastly different scales to govern past, current, and future patterns of diversity. Today, this goal is particularly critical given the degree to which rapid environmental change is affecting species distributions and community composition. Natural environmental gradients provide excellent opportunities to uncover possible mechanistic links between species distributions and environmental conditions – links that are invaluable for understanding how species may respond to environmental change. This dissertation builds on recent approaches that combine information on species' functional traits and evolutionary histories to refine our view of how contemporary and historical processes jointly govern the distribution of biodiversity. In the context of tropical tree communities of Puerto Rico, the following four chapters evaluate hypotheses about the distribution of different dimensions of diversity (i.e., species, functional, and phylogenetic) across regional abiotic gradients. In chapter 1, I develop an island-wide molecular phylogeny for the native and naturalized trees of Puerto Rico, and show preliminary evidence that dry forests comprise an evolutionarily clustered subset of the total island tree flora. In chapter 2, I examine functional and phylogenetic diversity across spatial resource gradients, and use these patterns to infer variation in community assembly processes along a gradient of water availability. In chapter 3, I use temporal shifts of functional and phylogenetic diversity during secondary succession to infer the shifts in the processes underlying successional change in wet forests of

Puerto Rico. Finally, in chapter 4, I evaluate the linkages between species functional traits and their geographic distributions, and test the hypothesis that community-weighted mean trait values reflect the ‘optimal’ strategy for a given set of abiotic conditions. A theme common to all chapters is the idea that functional and phylogenetic dimensions of diversity can shed light on the processes underlying patterns of diversity better than more traditional metrics of species diversity. I provide recommendations for future research directions at the end of each chapter and in the final conclusions.

Table of Contents

LIST OF TABLES AND FIGURES

Chapter 1

Chapter 2

Chapter 3

Chapter 4

Appendices

ACKNOWLEDGMENTS

Although difficult to quantify, characterize, and sometimes even recall, the contributions made by a huge number of people towards this project were essential. First, I thank my formidable committee whose friendly guidance has done much to shape my scientific thinking. In particular, I have tremendously enjoyed and benefitted from María Uriarte's unique and genuine combination of smarts, intuition, and wit. She provided me with exceptional guidance and entrusted me with resources to complete the substantial fieldwork required for this project. I happily look forward to ongoing collaboration. Rob Anderson taught me mountains about niche modeling and the importance of colorful shoelaces. Robin Chazdon has X-ray goggles when it comes to understanding forest dynamics and I thank her for sharing some of that insight with me. Joel Cracraft gave me the push towards historical biogeography that all ecologists need and taught me that species don't dream (or do they?). Dustin Rubenstein heard out my ideas, offered alternative perspectives, and introduced me to the incredible Serengeti series. Although not officially on my committee, collaborators who offered guidance and insight on particular chapters include Dave Erickson, John Kress, Nate Swenson, and Jess Zimmerman.

Next, I thank the many people whose help was essential for completing the fieldwork associated with this project. In particular, Marcos Rodriguez tirelessly (and pleasantly) assisted with all aspects of fieldwork. Fabiola Areces generously shared her expert botanical knowledge and time. Aaron Hogan, Chris Nytch, T.J. Agin, Víctor José Vega López, and Seth Rifkin were companions in the field and friends on the porch. Frank Axelrod, Alejandro Cubiña, Freddie Perez, José Sustache, Papo Vives, Alberto Areces, Pedro Acevedo, Danilo Chinea, Jeanine Vélez, and Marcos Caraballo contributed data, knowledge, blood, sweat, and/or tears.

The Department of Ecology, Evolution, and Environmental Biology at Columbia University provided a supportive environment to conduct this work. Lourdes Gautier, Maria Estrada-Werst, Jetaime McFadden, and Amy Kohn made the wheels turn. Many thanks to E3B faculty for shaping ideas, reading drafts, and writing recommendations – especially Shahid Naeem, Ruth DeFries, Kevin Griffin, Duncan Menge, and Matt Palmer. The entire community of E3B students and postdocs made for a supportive and stimulating environment. I especially appreciated the many drafts read, draughts poured, and/or baby-sitting shifts covered by Meghna Agarwala, Bene Bachelot, Mary Blair, Megan Cattau, Ben Clark, Liza Comita, Marina Côrtez, Matt Fagan, Dan Flynn, Mary Heskel, Meha Jain, Jesse Lasky, Nicole Minovets, Liz Nichols, Camilo Sanin, Naomi Schwartz, Ben Taylor, Vivian Valencia, Brian Weeks and Charles Yackulic. Over at City College, Rob Boria, Pete Galante, Jamie Kass, and Mariano Soley-Guardia taught me more than I thought I wanted to know about niche modeling. Numerous other mentors and friends encouraged me to take advantage of this opportunity in the first place and supported me throughout. In particular, Alex Wilson, Ted Fleming, Amy Harwood, Ethan Rose, Sally and Stella Garrido-Spencer, Ryan Jeffery, Adam Porterfield, Kaela Economos, Norah van Dusen, and Mark Beck.

Of course, I could never overstate the importance love and support of my family in completing this work. My parents, Damian and Anita, triggered my fascination with nature during all those camping trips and they believed in me at all steps along my meandering path. Steadfast support and encouragement from my brother, Jeff, has me convinced that anything is possible. My wife, Silvia Bibbo, has given me wonderful companionship, assistance in the field, a son, and an excuse to eat pasta every night. Our son, Martin Muscarella, has refreshed my perspective on life and cooperated with the writing process much more than expected from any ordinary baby.

Finally, I thank the following institutions: The Departmento de Recursos Naturales y Ambientales (DRNA) of Puerto Rico granted permission to conduct this work under DRNA permit #2011-IC-046. The University of Puerto Rico and staff at El Verde Field Station provided logistic support. The work was supported financially by grants from the Earth Institute and the Institute of Latin American Studies at Columbia University, the Smithsonian Institute, as well as NSF grants DEB 1050957 and DEB 1311367.

DEDICATION

To all the little seedlings that will never reach the canopy

INTRODUCTION

One goal at the core of ecology is to understand the connections between biophysical processes and species interactions that unfold over vastly different scales. At regional scales, evolution and biogeography structure biotas through speciation, extinction, and dispersal (Ricklefs & Schluter 1993; Holyoak *et al.* 2005). Locally, interactions among organisms and their environment shape patterns of co-occurrence (Weiher & Keddy 1995; Cavender-Bares *et al.* 2004b). The scale-dependent nature of ecological pattern and process requires that we adopt a multi-scale approach to understand the distribution of diversity (Ricklefs 2008). This is particularly critical given the degree to which contemporary species distributions are influenced by rapid environmental change. Gaining a better understanding of how these processes change along environmental gradients is critical for a more realistic view of how global change will influence the community assembly processes that govern the structure of communities (Purvis & Hector 2000; McGill *et al.* 2006).

Patterns of species diversity hold relatively little information to help distinguish among multiple underlying processes. Traits of species that reflect physiological mechanisms that mediate species occurrence patterns (e.g., traits that convey tolerance to drought) provide a useful framework for linking process to pattern across scales. It is essential to interpret these links in a historical context, however, to determine the degree to which variation in life history strategies among species reflects shared evolutionary history rather than convergent adaptation to particular environments. This dissertation builds on recent approaches that combine information on species' functional traits and evolutionary histories to gain a refined view of how contemporary and historical processes jointly govern the distribution of biodiversity in tropical forests at an intermediate regional scale (Webb *et al.* 2002; Cavender-Bares *et al.* 2006; Graham & Fine 2008; Devictor *et al.* 2010; Kraft & Ackerly 2010; Uriarte *et al.* 2010a; Swenson 2011).

Variation in life history strategies, including resource acquisition, resource conservation, and stress tolerance, arises through evolutionary trade-offs associated with heterogeneous environments (Roff 1992; Sibly 1997; Johnson & Stinchcombe 2007). Shifts in the functional composition of communities along environmental gradients reflect the combined influences of historical and contemporary processes that generate and maintain diversity in response to underlying environmental variation (Díaz & Cabido 2001; McGill *et al.* 2006). For instance, environmental conditions can filter species across environmental gradients, restricting membership of local communities and confining species to areas with particular conditions. At the same time, functionally similar species may compete for shared resources differently than functionally distinct species (de Bello *et al.* 2009). Understanding phylogenetic relationships among species can help inform studies of functional diversity in two ways (Helmus *et al.* 2007). First, it allows us to determine the degree to which variation in life history strategies among species reflects shared evolutionary history rather than convergent adaptation to particular environments. Second, it may reveal "cryptic variation" in functional variation among species (Ives & Helmus 2010). For example, despite being difficult to measure, closely related species may be susceptible to a similar set of pathogens (Ricklefs 2010).

Existing studies that integrate information on functional traits and phylogenetic relatedness have largely focused on either stand- (e.g., Kraft *et al.* 2008) or global-scale analyses (e.g., Qian & Ricklefs 2007). Of central interest for most stand-scale studies is the relative influence of local biotic interactions versus abiotic filtering in mediating spatial and temporal patterns of community diversity. This scale, however, captures a relatively small window of environmental heterogeneity compared with the broad environmental gradients prevalent in nature. At the other end of the spatio-temporal spectrum, global studies necessarily emphasize biogeographical explanations for patterns of diversity, thereby missing the critical contribution of local interactions (Wisz *et al.* 2012). This dissertation focuses on an intermediate regional scale (i.e., the island of Puerto Rico), which provides an excellent opportunity to explore the role of local and historical processes across a wide range of abiotic conditions. Additionally, the regional scale of analysis allows this work to go beyond the use of phylogenetic history as a corrective measure (Harvey & Pagel 1991), to understand how ecological and evolutionary processes unfold over time to generate patterns of biodiversity (Jombart *et al.* 2010; Pavoine & Bonsall 2010).

In this dissertation, I use tropical forest tree communities in Puerto Rico to examine hypotheses about the distribution of taxonomic, functional, and phylogenetic diversity across broad environmental gradients. Generally, environmental gradients provide opportunities to understand patterns of diversity by facilitating mechanistic links between species distributions and environmental conditions (Fukami & Wardle 2005). These links can be invaluable for understanding how species may respond to changing environmental conditions (McGill *et al.* 2006). I focus predominantly on the island's precipitation gradient because current global circulation models project precipitation in the Caribbean to decline by as much as 50% (Neelin *et al.* 2006; Gamble & Curtis 2008). Long-term changes in precipitation, including increased frequency and intensity of short-term droughts (Allan & Soden 2008), are likely to play major roles in governing the distributions of species under future climate regimes.

The overarching aim of this dissertation is to understand the multi-scale processes that govern species distributions and patterns of community diversity across abiotic gradients. By combining information on species functional and phylogenetic relationships, each chapter tests process-driven hypotheses about how these components are expected to vary through space and time.

Tropical forests of Puerto Rico have been an ideal system to conduct the research presented here for two primary reasons. Most importantly, the island encompasses several substantial environmental gradients in a relatively small area. Six Holdridge life zones ranging from subtropical dry forest to subtropical rainforest occur on the island (Ewel & Whitmore 1973), and mean annual precipitation ranges from ca 800–5,000 mm yr^{-1}. The island's complex geologic history is reflected in its rugged topography (0–1,338 m a.s.l.) and high diversity of soil types (Bawiec 1998; Miller & Lugo 2009). Second, substantial existing data on the flora (e.g., Colon 1996; Liogier 1996; Santiago-Valentin & Olmstead 2004; Kress *et al.* 2010; Acevedo-Rodríguez & Strong 2011; Axelrod 2011), physical characteristics (e.g., Ewel & Whitmore 1973; Bawiec 1998; Lugo *et al.* 2001; Daly *et al.* 2003; Miller & Lugo 2009), and land use history (e.g., Rudel *et al.*

2000; Helmer *et al.* 2002; Yackulic *et al.* 2011) can be leveraged to ask questions at a truly regional scale.

In **Chapter 1**, “A well-resolved phylogeny of the trees of Puerto Rico based on DNA barcode sequence data”, I generate an island-wide molecular phylogeny for the trees of Puerto Rico. I compare methods for building large-scale community phylogenies and test hypotheses about the phylogenetic composition of Puerto Rican forests. The phylogeny presented in this chapter will facilitate stronger inferences about the role of historical processes in governing the assembly and composition of Puerto Rican forests, provide insight into Caribbean biogeography, and encourage the use of evolutionary history during conservation planning in Puerto Rico.

In **Chapter 2**, “Functional and phylogenetic dimensions of tree diversity reveal shifting assembly mechanisms across regional environmental gradients in Puerto Rico”, I focus on determining how the relative strength of different assembly processes changes along the island’s precipitation gradient. The results suggest geographic variation in the predominant assembly mechanisms acting across regional environmental gradients.

In **Chapter 3**, “Functional convergence and phylogenetic divergence during secondary succession of subtropical wet forests in Puerto Rico”, I use functional and phylogenetic information to elucidate the drivers of forest succession. Contrasting shifts in the functional and phylogenetic composition of forests during succession provide insight to the processes driving community dynamics during succession.

In **Chapter 4**, “Linking functional traits and ecological niche models: insights into assembly processes for Puerto Rican forests”, I investigate how species geographic distributions relate to patterns of community functional composition. I merge this analysis with an examination of trait correlations to understand the role of trait-mediated habitat filtering and fine-scale niche partitioning as drivers of species geographic distributions and community diversity.

CHAPTER ONE:

A well-resolved phylogeny of the trees of Puerto Rico based on DNA barcode sequence data

Robert Muscarella, María Uriarte, David L. Erickson, Nathan G. Swenson, Jess K. Zimmerman, and W. John Kress

Published as:
Muscarella, R., Uriarte, M., Erickson, D.L., Swenson, N.G., Zimmerman, J.K. & Kress, W.J. (2014). A well-resolved phylogeny of the trees of Puerto Rico based on DNA barcode sequence data. *PLoS ONE*. 9, e112843.

ABSTRACT

Background: The use of phylogenetic information in community ecology and conservation has grown in recent years. Two key issues for community phylogenetics studies, however, are (i) low terminal phylogenetic resolution and (ii) arbitrarily defined species pools. *Methodology/Principal Findings*: We used three DNA barcodes (plastid DNA regions *rbcL*, *matK*, and *trnH-psbA*) to infer a phylogeny for 527 native and naturalized trees of Puerto Rico, representing the vast majority of the entire tree flora of the island (89%). We used a maximum likelihood (ML) approach with and without a constraint tree that enforced monophyly of recognized plant orders. Based on 50% consensus trees, the ML analyses improved phylogenetic resolution relative to a comparable phylogeny generated with PHYLOMATIC (proportion of internal nodes resolved: constrained ML = 74%, unconstrained ML = 68%, PHYLOMATIC = 52%). We quantified the phylogenetic composition of 15 protected forests in Puerto Rico using the constrained ML and PHYLOMATIC phylogenies. We found some evidence that tree communities in areas of high water stress were relatively phylogenetically clustered. Reducing the scale at which the species pool was defined (from island to soil types) changed some of our results depending on which phylogeny (ML vs. PHYLOMATIC) was used. Overall, the increased terminal resolution provided by the ML phylogeny revealed additional patterns that were not observed with a less-resolved phylogeny. *Conclusions/Significance*: With the DNA barcode phylogeny presented here (based on an island-wide species pool), we show that a more fully resolved phylogeny increases power to detect

nonrandom patterns of community composition in several Puerto Rican tree communities. Especially if combined with additional information on species functional traits and geographic distributions, this phylogeny will (i) facilitate stronger inferences about the role of historical processes in governing the assembly and composition of Puerto Rican forests, (ii) provide insight into Caribbean biogeography, and (iii) aid in incorporating evolutionary history into conservation planning.

INTRODUCTION

The use of phylogenetic information in community ecology and conservation has grown dramatically in recent years (Losos 1996; Vellend *et al.* 2011; Cavender-Bares *et al.* 2012). This body of research has been largely stimulated by the idea that evolutionary relationships can provide insights into the historical processes governing assembly of local communities (Webb *et al.* 2002; Cavender-Bares *et al.* 2009; Vamosi *et al.* 2009). From a conservation perspective, phylogenies may reveal aspects of biodiversity that are not observable from traditional metrics of species diversity (Vane-Wright *et al.* 1991; Williams *et al.* 1991; Faith 1992; Crozier 1997; Devictor *et al.* 2010). By providing a historical context, phylogenies help merge our understanding of ecological, evolutionary, and biogeographic drivers of community composition (Ricklefs 1987).

One key issue for research in community phylogenetics is how to best estimate phylogenetic relationships among species in diverse communities (e.g., tropical forests). To date, the program PHYLOMATIC (Webb & Donoghue 2005) has become a primary method by which ecologists integrate phylogenetic information with analyses of community patterns (e.g., Kembel & Hubbell 2006; Willis *et al.* 2008; Kraft & Ackerly 2010). For plants, PHYLOMATIC generates community phylogenies by pruning a megatree of angiosperms given a user-defined species list. This approach offers a repeatable and accessible way to obtain phylogenies using existing data (also see Beaulieu *et al.* 2012), however, PHYLOMATIC phylogenies typically have low or no taxonomic resolution among closely related species (e.g., within plant families or genera). Low taxonomic resolution can reduce statistical power for detecting nonrandom patterns of community structure (Kress *et al.* 2009; Swenson 2009) and can bias estimates of phylogenetic signal (Davies *et al.* 2012). Furthermore, because single genera often contain numerous species with diverse life-history characteristics (e.g., Cavender-Bares *et al.* 2004a; Sedio *et al.* 2012), resolving evolutionary relationships among congeners is critical for interpreting the link between patterns of phylogenetic community composition and the history of trait evolution. Finally, low taxonomic resolution can preclude inferences about biogeographic influences on local assemblages. The issue is particularly acute with respect to relatively recent evolutionary history (i.e., speciation events), which arguably represent a key connection between local and regional processes (Ricklefs & Jenkins 2011).

In contrast to megatree approaches such as PHYLOMATIC, phylogenies based on genetic data typically provide comparatively high taxonomic resolution. Generating molecular phylogenies, however, requires a significant investment of resources and expert knowledge. Additionally, determining how to estimate phylogenies among the very distantly related species that are typical of community-based phylogenies (as opposed to clade-based phylogenies) remains an active area of research. One potentially promising approach is to integrate existing information on evolutionary relationships in the form of a constraint tree (Kress *et al.* 2010). More research is required, however, to determine the influence of constraint trees on phylogenetic reconstruction and downstream analyses of community phylogenetic patterns.

Another characteristic of many existing studies of community phylogenetic structure lies in the lack of consistent methodology in defining species pools when testing hypotheses about mechanisms driving community assembly (e.g., competition versus environmental filtering) (Webb *et al.* 2002). Generally, these analyses are based on null models that compare an observed metric of phylogenetic composition (e.g., NRI, the net relatedness index) with a random expectation based on assemblages drawn from a regional species pool (Kraft & Ackerly 2010). In practice, studies often delimit the ‘regional pool’ as the set of species encountered in the study, regardless of the ecological significance of the study area boundaries (e.g., forest dynamics plots). Examining species assemblages within such arbitrarily defined regions can provide information on processes occurring at certain scales (Kraft & Ackerly 2010; Uriarte *et al.* 2010b). However, varying the spatial scale at which species pools are defined can provide important opportunities to evaluate the relative strength of local assembly processes (e.g., interactions that occur among neighboring trees) versus processes that occur over larger spatial and temporal scales (e.g., evolution and biogeography) and across broader environmental gradients (Hardy 2008; de Bello *et al.* 2012; Lessard *et al.* 2012a; Eiserhardt *et al.* 2013; Brunbjerg *et al.* 2014; Münkemüller *et al.* 2014). For example, numerous studies in phylogenetic community ecology have shown that as the spatial (and taxonomic) extent of the species pool increases, the phylogenetic composition of local communities tends to appear increasingly 'clustered' (i.e., co-occurring species are more closely related than expected by random chance). Other studies have shown more mixed results (Münkemüller *et al.* 2014 *and references therein*; Parmentier *et al.* 2014), which may emerge, for example, if a larger species pool includes sister taxa absent from the smaller pool. In any case, scale-dependency of community patterns likely reflects the scales at which different assembly processes influence community structure (Pickett & Bazzaz 1978; Silvertown *et al.* 2006; Ackerly & Cornwell 2007; Cavender-Bares *et al.* 2009). As such, we can gain valuable insights on community assembly by adjusting species pools to suit particular hypotheses about the scales at which different assembly processes act (Cavender-Bares *et al.* 2006; Swenson *et al.* 2007; Vamosi *et al.* 2009; Lessard *et al.* 2012a; Lessard *et al.* 2012b; Eiserhardt *et al.* 2013; Münkemüller *et al.* 2014; Swenson & Umaña 2014).

In this study, we used DNA sequence data to generate an island-wide phylogeny for nearly all of the native and naturalized tree species of Puerto Rico. Specifically, we used sequence data from three regions of plastid DNA, which are commonly used as plant DNA barcodes (*rbcL, matK, trnH-psbA*; Kress *et al.* 2010) to resolve evolutionary relationships among 527 recognized species with a maximum likelihood (ML) approach. We compare phylogenetic resolution of two ML phylogenies (built with and without the use of an ordinal-level constraint tree) and a comparable phylogeny derived from PHYLOMATIC. We then explore the implications of these different methods in a case study where we examined the phylogenetic structure of tree communities in 15 protected forests in Puerto Rico. These 15 forests span a wide variation in environmental conditions, providing an ideal template for evaluating the effects of local environmental variation on phylogenetic community structure within the island of Puerto Rico (Table 1). We addressed the following specific questions:

1. *How does the use of a constraint tree influence (i) the level of bootstrap support in a DNA barcode phylogeny of Puerto Rican trees, and (ii) the degree to which a molecular phylogeny corresponds with currently recognized taxonomic groups?* We predicted that the constraint tree would provide higher levels of bootstrap support among unconstrained nodes and increase concordance with current taxonomy relative to the unconstrained analysis.
2. *How do patterns of community phylogenetic structure in Puerto Rican forests differ when based on a DNA barcode phylogeny versus a PHYLOMATIC phylogeny?* We predicted that an increase in statistical power provided by the higher resolution of a molecular phylogeny would lead to a stronger signal of non-random phylogenetic structure.
3. *How does phylogenetic structure in Puerto Rican forests change with respect to different species pool definitions?* We predicted co-occurring species would tend to appear relatively phylogenetically clustered with respect to the full island species pool because of a strong role for environmental filtering across broad environmental gradients. We predicted that a more restricted species pool definition would reduce the level of phylogenetic clustering if niche differentiation (competitive exclusion) becomes more apparent at small spatial scales.

MATERIALS AND METHODS

All necessary permits were obtained for the described study, which complied with all relevant regulations. Specifically, the Departmento de Recursos Naturales y Ambientales (DRNA) of Puerto Rico granted permit #2011-IC-046 to collect plant specimens in the state forests of Puerto Rico. Herbaria staff at the University of Puerto Rico, Rio Piedras and the US National Herbarium provided permission to sample tissue from their collections.

Study area and species

The island of Puerto Rico encompasses six Holdridge life zones (Holdridge 1947) ranging from subtropical dry forest to subtropical rainforest in an area of 8,740 km^2 (Ewel & Whitmore 1973). Mean annual precipitation ranges drastically, from ca. 700–4,500 mm yr^{-1} (Daly *et al.* 2003). The island's complex geologic history is reflected in its rugged topography (0–1,338 m a.s.l.) and diverse parent soil materials, which include volcanic, limestone, alluvial, and ultramafic materials (Bawiec 1998). Substantial existing data on the flora (Kress *et al.* 2010; Acevedo-Rodríguez & Strong 2011; Axelrod 2011) provide a strong foundation for our work.

We created an initial list of Puerto Rican trees with the species list from the USFS Forest Inventory and Analysis (FIA) Caribbean field guide (USFS 2006). With guidance from local experts (P. Acevedo-Rodríguez, F. Areces, F. Axelrod, M. Caraballo, J. Sustache, and P. Vives, *personal communication*), we modified this list by (1) updating nomenclature to be consistent with Acevedo-Rodríguez and Strong (2011), (2) removing species occurring only under cultivation and (3) adding native and naturalized tree species known to occur in Puerto Rico but absent from the FIA list. Our final list of target species contained 594 species of seed plants representing 33 orders, 86 families, and 304 genera (Muscarella *et al*. 2014b). Of these, we were able to compile DNA sequence data for 523 (89%) species representing all 32 orders, 85 families (99%), and 287 genera (94%). The single excluded family (Cunoniaceae) is represented in Puerto Rico by a single rare species of shrub and most of the other species missing from our dataset are relatively uncommon and distributed widely throughout taxonomic groups. As a result, we do not expect the missing species to influence overall results of community phylogenetic analyses. However, it will be enlightening to include these species when sequence data become available in order to better understand the contributions of rare species to phylogenetic diversity (Mi *et al.* 2012).

Tissue collection and lab procedures

We acquired DNA sequence data from a variety of sources. Primarily, we obtained leaf tissue either from freshly collected specimens or existing herbarium sheets. For fresh specimens, we dried leaf tissue in silica gel prior to DNA extraction. Prior to depositing voucher specimens at the US National Herbarium (US), we verified species identifications by referring to the herbarium at the University of Puerto Rico, Río Piedras (UPRRP) and through consultation with local experts (F. Areces, F. Axelrod, P. Vives, *personal communication*). For 95 species, we collected leaf tissue from dry material sampled from herbarium specimens at UPRRP or US. DNA extraction, amplification and sequencing protocols followed Kress *et al*. (2010). Specifically, we used the following lab procedures for fresh and dried leaf tissue. After disrupting tissue with a Tissuelyzer (Qiagen Cat. #85210), we incubated samples overnight at 55°C in a CTAB-based extraction buffer (AutoGen, Holliston, MA). Following incubation, we removed the supernatant and placed it in clean, 2ml 96-well plate for submission to a DNA extraction robot (AutoGen 960, Holliston, MA). We hydrated DNA extractions in 100mM Tris-HCl (pH 8.0) and then transferred them to Matrix barcode tubes (MatrixTechnologies Cat. # 3735) and stored them at −80°C. Working stocks of DNA were transferred to a microtiter plate and diluted 5× with water prior to PCR. We used routine PCR, with no more than three attempts per sample to recover PCR amplicons for each sample. The PCR cycling conditions were exactly the same for *rbcL* and *trnH-psbA* (95°C 3min, [94°C 30sec, 55°C 30sec, 72°C 1min]×35cycles, 72°C 10min) following procedures outlined in Kress & Erickson (2007). The PCR cycling conditions for *matK* required lower annealing temperatures and more cycles (95°C 3min [94°C 30sec, 49°C 30sec, 72°C 1min]×40 cycles, 72°C 10min) following Fazekas *et al*. (2008) and included DMSO at a final concentration of 5%. We purified successful PCR reactions with a 56 diluted mixture of ExoSap (USB, Cat. # 78201). For sequencing, 2–4ul of the purified PCR was used in a 12ul reaction (0.8ul BigDye terminator sequencing mixture (V3.1; ABI, Cat. 4337457), 2.0ul of a 56 buffer (400u Molar Tris-HCL pH 8.0), 1ul of 1uMolar primer and distilled water to volume). Sequencing of *matK* PCR products included DMSO to a final concentration of 4% in the reaction mixture. Cycling sequencing protocols were the same for all markers, (95°C 15sec [94°C 15sec, 50°C 15sec, 60°C 4min]×30 cycles). Following cycle sequencing, products were purified on a column of sephadex and sequence reactions were read on an ABI 3730 (Applied Biosystems).

We also incorporated existing sequence data for 143 species previously sequenced from the Luquillo Forest Dynamics Plot (Kress *et al.* 2010) and for 25 species from GenBank (Benson *et al.* 2010). We excluded 67 species from analyses for which we were unable to acquire reliable sequence data either because tissue was not available or because of failure during DNA sequencing.

Sequence editing, alignment, and assembly

We used GENEIOUS (R6, version 2.4.1; Biomatters Ltd.) to trim and assemble trace files for each marker into bidirectional contigs. Separately for each marker, we aligned sequences using SATÉ (Liu *et al.* 2012). SATÉ is an iterative algorithm that divides the original sequence data set using a tree-based decomposition; we aligned these smaller sets of sequences using MAFFT (Katoh *et al.* 2005) and merged these sub-alignments into a global alignment without disrupting the individual sub-alignments using MUSCLE (Edgar 2004). SATé is particularly effective for conducting multiple sequence alignment among very distantly related taxa through the use of merging sub-alignments among related sequences, and has been widely applied for studies of very broad phylogenetic application (Schoch *et al.* 2009; Kivlin *et al.* 2011). We then concatenated the three separate marker alignments to produce an aligned three-gene matrix. Gaps were not coded and were treated as missing data in phylogenetic reconstruction.

Phylogenetic reconstruction

We generated a phylogeny using maximum likelihood (ML) methods, implemented in RAXML (Stamatakis et al. 2005) via the CIPRES Science Gateway (Miller *et al.* 2010). Based on jModelTest2 (Darriba *et al.* 2012), we modeled nucleotide substitution using a GTR+GAMMA model, with substitution rates estimated independently for each gene. We evaluated node support for the topology with the highest likelihood using 100 bootstrap runs. In addition, we trimmed PHYLOMATIC reference tree R20120829 (Stevens 2001 onwards) to use for comparative purposes. While other methods for phylogenetic reconstruction are available (*e.g.*, parsimony), we focus here on a comparison between ML methods and a very commonly used method of generating phylogenies for community ecology (i.e., PHYLOMATIC).

Rather than including densely sampled small taxonomic units, community phylogenies often contain smaller numbers of more distantly related species (*e.g.*, 32 orders represented in our dataset, represented by 18 species, on average). Resolving both shallow and deep relationships requires distinct molecular data sets that are difficult to assemble. When strong prior information is available, one approach to confront this issue is to enforce some relationships through the use of a constraint tree (e.g., Smith *et al.* 2009). In the case of our study, the Angiosperm Phylogeny Group III (APG III 2009) represents the authoritative standard for current relationships up to the family level in angiosperms. However, within the AGP III phylogeny, relationships between species are generally not resolved beyond the family level, thus providing an ideal opportunity to use DNA barcodes to resolve these finer-scale relationships. To test the ability of a constraint tree to improve phylogenetic resolution among distantly related taxa, we repeated the ML analysis detailed above using the APG III phylogeny (APG III 2009) to constrain the topology of ordinal and deeper nodes. This approach allowed the topology within each order to be resolved with DNA barcode sequence data while ordinal and deeper nodes were enforced *a priori*. We dated both the constrained and unconstrained ML phylogenies using PATHd8 (Britton *et al.* 2007) with age constraints based on fossil records provided in the Appendix of Magallón & Castillo (2009). The input files used for this analysis are provided in Muscarella *et al*. (2014b). The constraints we used included one fixed age estimate for the angiosperm crown group and 35 minimal age estimates for other clades represented in our phylogeny (Magallón & Castillo 2009). We used this approach because dated ultrametric trees are the standard for community phylogenetics studies; however, we also provide the undated, non-ultrametric trees in Muscarella *et al*. (2014b). To explore the distribution of uncertainty across the phylogeny, we calculated the proportion of recognized taxonomic groups (orders, families, and genera) that were found to be monophyletic in each analysis and the proportion of resolved nodes within each of these groups.

Case study: Phylogenetic composition of Puerto Rican forests

We measured the phylogenetic composition of 15 protected forests in Puerto Rico based on species occurrence data (presence/absence) from Little & Wadsworth (1964) and Little *et al*. (1974). As a synthesis of observations made by local experts, these volumes are the most commonly used references to describe tree composition of Puerto Rico's protected forests. The

15 forests examined here span a wide range of environmental conditions (precipitation range: ca. 800–3,800 mm yr^{-1}, elevation range: ca. 0–1,300 m a.s.l.) and occur across four main soil parent materials: unconsolidated, limestone, volcanic, and serpentine (Table 1, Fig. 1). We excluded taxa not included in our phylogeny – these accounted for only 2% of the total observations in the community dataset. With the remaining data, we quantified phylogenetic composition of each forest using the net relatedness index (NRI) and nearest taxon index (NTI) (Webb *et al.* 2002). These indices describe whether sets of co-occurring taxa are more or less closely related than random assemblages of equal species richness drawn from a pool of species. Specifically, NRI measures the average degree of relatedness among all members of the community and thus emphasizes deeper branches of the phylogenetic tree. In contrast, NTI is based on the average distance between closest relatives in each assemblage and thus emphasizes compositional patterns at the tips of the phylogeny (Webb *et al.* 2002). These metrics are calculated as: NRI = $-(r_{\text{obs}} - \text{mean}(r_{\text{rand}}))/\text{sd}(r_{\text{rand}})$, where r is either the co-occurring taxa (for NRI) or mean phylogenetic branch length separating nearest neighbors (for NTI). The observed value is r_{obs} and r_{rand} is a distribution of values based on assemblages drawn from a species pool. We calculated NRI and NTI for each forest using two different species pools: the full list of species in our dataset (the 'island pool'), and the list of species recorded from forests on the same soil parent material (the 'soil pool'). For example, for Guánica forest (limestone soil), we calculated two values of NRI: one value ($\text{NRI}_{\text{ISLAND}}$) based on null assemblages drawn from the entire species list and another value (NRI_{SOIL}) based on the list of species recorded from all forests on limestone soil (the soil pool).

We computed NRI and NTI using the ses.mpd and ses.mntd functions of the 'picante' package (Kembel *et al.* 2010) for R v 3.1.1 (R Development Core Team 2014). We ran the analyses for 999 iterations and used the 'taxa-labels' null model. We chose this null model to control for the observed species occupancy rates and species richness of each forest. Positive values of NRI and NTI indicate phylogenetic clustering whereas negative values indicate phylogenetic evenness. We performed these analyses using the constrained ML 50% consensus tree and the PHYLOMATIC phylogeny. We based these analyses on the constrained ML 50%

consensus tree because it reflects the uncertainty of our phylogenetic hypothesis given our data, while also incorporating the strong evidence resolving deep relationships provided by the APG III constraint tree.

We quantified shifts in NRI and NTI values between the two species pool definitions using paired t-tests and we quantified the similarity of these values between phylogenies with Pearson's correlation coefficient. In addition to overall patterns of community phylogenetic composition, we used the 'nodesig' algorithm in PHYLOCOM v 4.2 (Webb *et al.* 2008) to determine the particular clades that contribute significantly more or fewer species than expected to the composition of each forest.

RESULTS

DNA barcode sequences

From fresh tissue, we successfully recovered sequence data from 85%, 75%, and 94% of samples for *rbcL*, *matK, trnH-psbA*, respectively. The final three-gene alignment comprised 3,366 base pairs (549 bp for *rbcL*, 1,070 bp for *matK*, and 1,747 bp for *trnH-psbA*). The data matrix had 62.2% missing data (including gaps coded as missing data and species for which we did not recover sequence data). This amount is far more compact than previous alignments of the same three regions that used a nested partitioning of the *trnH-psbA* alignment, resulting in > 95% missing data (Kress *et al.* 2010). Considering each region separately, the amount of missing data was 23.1%, 49.2%, and 82.1% for *rbcL, matK,* and *trnH-psbA*, respectively. The full list of species included in the analysis and GenBank accessions are provided in Muscarella *et al*. (2014b).

Phylogenetic analyses

We provide the constrained and unconstrained ML trees, with bootstrap support, as well as the PHYLOMATIC phylogeny used in our analyses in the appendix of Muscarella *et al*. (2014b). Overall, we found relatively strong support for the majority of nodes in the both the constrained and unconstrained ML trees (Fig. 2). Across all nodes, 74% of nodes in the constrained ML tree received ≥50% bootstrap support and 52% received ≥80% bootstrap support. Considering only the 468 unconstrained nodes, 71% received ≥50% bootstrap support and 46% received ≥80% bootstrap support. The unconstrained ML tree had slightly lower levels of support with 68% of nodes receiving ≥ 50% support and 43% of nodes receiving ≥ 80% support. Both the constrained and unconstrained ML trees had higher resolution than the PHYLOMATIC tree, in which only 52% of internal nodes were resolved. For the constrained ML tree, monophyly was supported for 91% of families and 87% of genera (monophyly of orders was constrained). In comparison, monophyly was supported for 72% of orders, 85% of families, and 87% of genera in the unconstrained ML tree. In both cases, the non-monophyly of currently recognized families related to the placement of taxa for which we did not have sequence data for all three barcode regions. For the constrained ML tree, the average proportion of nodes within orders, families, and genera with ≥ 50% bootstrap support was 0.81 (± SD 0.20), 0.87 (± SD 0.20), and 94% (± SD 0.19), respectively. For the unconstrained ML tree, the average proportion of nodes within orders, families, and genera with ≥ 50% bootstrap support was 0.92 (± SD 0.14), 0.89 (± SD 0.18), and 92% (± SD 0.20), respectively.

Case study: Phylogenetic composition of Puerto Rican forests

Some patterns of phylogenetic community structure varied with respect to the phylogeny and species pool used in analyses (Fig. 3). For NRI, which emphasizes tree-wide patterns, Guánica dry forest was significantly clustered (i.e., taxa were more closely related than expected) based on the full island species pool for both the ML and PHYLOMATIC phylogenies (Fig. 3a). None of the other 14 forests departed from random expectations for NRI when based on the island pool. When considering the (reduced) soil species pools, the composition of the two wettest forests (Toro Negro and El Yunque, both located on volcanic soils) were significantly overdispersed (i.e., taxa were less closely related than expected), although the NRI_{SOIL} value for Toro Negro was only significant with respect to the ML phylogeny (Fig. 3b). For NTI, which emphasizes compositional patterns at the tips of the phylogeny, Cambalache forest was significantly clustered with respect to the full island species pool but only for the ML phylogeny (Fig. 3c). None of the forests had significantly nonrandom NTI values when the analyses were based on the (reduced) soil species pools, regardless of which phylogeny was used (Fig. 3d).

None of the forests shifted from significantly clustered to significantly even when comparing NRI or NTI values based on the two different species pools. However, as we predicted, the (reduced) soil species pools caused both of these metrics to become more negative (i.e., decreased the signal of phylogenetic clustering) when calculated with the ML phylogeny (paired t-test: NRI: t=2.79, df=14, p<0.01; NTI: t=4.34, df=14, p<0.001). In contrast, these species pool definitions did not significantly change NRI or NTI when calculated with the PHYLOMATIC phylogeny (paired t-test: NRI: t=0.39, df=14, p=0.35; NTI: t=0.28, df=14, p=0.39). Values of NRI calculated with each phylogeny were strongly correlated for both species pool definitions (island pool: Pearson's *r* = 0.96, p < 0.001; soil pool: Pearson's *r* = 0.92, p <0.001) but values of NTI were less strongly correlated between these two phylogenies, and were not significantly correlated when based on the soil species pool (island pool: Pearson's *r* = 0.60, p = 0.02; soil pool: Pearson's *r* = 0.48, p = 0.06).

The node-based analysis identified particular clades that were relatively over- and under represented in each forest compared with a random expectation and, overall, the ML and PHYLOMATIC phylogenies produced largely congruent results (Table 1). One of the more consistent results was that species belonging to Melastomataceae tended to be significantly underrepresented in relatively dry forests on limestone and serpentine soils (*i.e.*, Guánica, Cambalache, Maricao, and Susúa) and relatively overrepresented in three relatively wet forests on volcanic soils (Guilarte, Luquillo and Toro Negro). Also, phylogenetic clustering of Guánica forest appears to be primarily driven by an overrepresentation of Fabaceae and Capparaceae, together with an underrepresentation of magnoliids, and Ericales (Table 1).

DISCUSSION

The island-wide phylogeny for Puerto Rican trees presented here represents the community phylogenetics approach applied at a regional scale with the use of DNA sequence data. Both the constrained and unconstrained ML phylogenies provided increased phylogenetic resolution in comparison with a corresponding PHYLOMATIC tree, a predominant tree-building approach used in studies of community phylogenetics. In this study, the use of an ordinal-level constraint tree provided slightly higher phylogenetic resolution compared to the unconstrained analysis. In our case study, we uncovered patterns of nonrandom phylogenetic structure in Puerto Rican forests that depended on the phylogeny used as well as the scale at which the regional species pool was defined. Considering the rapidly increasing availability of DNA sequence data, future regional scale work in community phylogenetics will benefit from highly resolved phylogenies that include many taxa sampled across large areas and broad environmental gradients (Swenson *et al.* 2012; Swenson 2013; Swenson & Umaña 2014).

Comparison between phylogenies and taxonomic resolution

Although the ML phylogenies generated in this study were not completely resolved, the constrained 50% consensus tree did increase tip resolution by 22% in comparison with the PHYLOMATIC tree. This relatively high degree of phylogenetic resolution has a number of important implications for community phylogenetic analyses (Kress *et al.* 2009; Swenson 2009; Davies *et al.* 2012). First, poorly resolved phylogenies tend to reduce statistical power for detecting nonrandom patterns of community structure (e.g., with NRI and NTI), an issue that appears to be more severe with larger phylogenies (Swenson 2009). Swenson (2009) found that statistical power was most strongly reduced, however, when deeper nodes were unresolved (i.e. among orders and families) as opposed to more recent nodes (i.e. among species). As a result, we expect that the remaining unresolved nodes in our ML tree have a relatively small effect on analyses of phylogenetic structure for Puerto Rican tree communities because our constraint tree fixed the resolution of the deeper nodes. At the same time, the relatively deep nodes of the PHYLOMATIC phylogeny are also resolved, suggesting that a reduction in statistical power for detecting nonrandom patterns between our ML tree and the PHYLOMATIC tree may be most pronounced for metrics that focus on phylogenetic patterns among close relatives (e.g., NTI).

A second issue related to poorly resolved phylogenies is an upward bias when estimating phylogenetic signal (Davies *et al.* 2012). In other words, the tendency for close relatives to have similar functional traits tends to be overestimated when phylogenies are poorly resolved. This bias is of particular concern when examining patterns of phylogenetic community composition given the central role of phylogenetic signal of traits relevant for species co-occurrence (Mayfield & Levine 2010). In general, the relatively high degree of tip resolution afforded by molecular data can strengthen inferences that rely on linking phylogenetic and functional patterns of community composition.

A major challenge in generating large-scale community phylogenies (and systematic biology, in general) is how to recover accurate phylogenetic relationships given limited data. Researchers have long debated the relative benefits of increasing sequence length versus increasing taxon sampling to improve the accuracy of phylogenetic reconstruction (e.g., Rannala *et al.* 1998; Poe & Swofford 1999; Felsenstein 2004; Heath *et al.* 2008; Nabhan & Sarkar 2012). This issue, however, has rarely been discussed in the context of community phylogenetics even though community-based analyses typically have relatively sparse taxon sampling compared to clade-based analyses. One implication of sparse taxon sampling is that long-branch attraction can reduce the accuracy of inferred topologies (Stefanović *et al.* 2004; Heath *et al.* 2008). We confronted this potential issue by using a constraint tree to leverage strong prior information on deep phylogenetic relationships. In our case, the Angiosperm Phylogeny Group (APG III 2009) provides a synthesis of well-supported relationships among the plant orders. Overall, bootstrap support for the constrained ML tree was higher than for the unconstrained tree although we had originally expected a stronger effect of using the constraint tree. The fairly high success of recovering recognized orders in the unconstrained analysis likely derives from the large sample size included in this study and the particular genes used; they were chosen, in part, for their high performance in phylogenetic analyses (Kress & Erickson 2007; CBOL Plant Working Group 2009).

While this study used a less sparse data matrix than previous work (Kress *et al.* 2010), the alignment procedure we use still resulted in a relatively sparse data matrix, particularly for the *trnH-psbA* region. The reason for this is that the SATÉ alignment algorithm knits together small alignments and introduces gaps when making a consensus alignment (Liu *et al.* 2012). Evidence suggests that introducing gaps does not affect the overall phylogenetic results as seen with the success of phylogenetic reconstructions using super matrix methods that produce extremely sparse alignments (McMahon & Sanderson 2006) and studies that successfully align non-coding *ITS* and chloroplast intergenic spacer data for very large phylogenetic assemblages (Edwards & Smith 2010). These studies suggest that missing data is not critical, particularly if one gene is shared among all taxa. Furthermore, while the effects of missing data on phylogenetic analyses are complex (Wiens 2006), several studies suggest that even taxa with large amounts of missing data can be accurately placed in phylogenies as long as the total number of characters sampled is large (*e.g.,* (Phillipe *et al.* 2004; Wiens & Morrill 2011)). In addition, Wiens (2006) showed that, in some cases, taxa with large amounts of missing data can improve overall phylogenetic accuracy, particularly with model-based phylogenetic methods (e.g., likelihood; Wiens 1998; *but see* Poe 2003). In our case, some instances of non-monophyly of recognized taxonomic groups were caused by individual taxa for which we did not have the full complement of three gene regions. Continued investigation of the influence of missing data on large phylogenetic analyses will help clarify the conditions under which missing data may decrease phylogenetic accuracy.

Case study: Phylogenetic composition of Puerto Rican forests

Our analysis of Puerto Rican tree communities provides an initial look at broad patterns of phylogenetic structure at a regional scale. For the most part, the ML and PHYLOMATIC phylogenies provided congruent results in terms of NRI, which is a tree-wide metric of phylogenetic composition. In contrast, NTI values, which are more sensitive to variation at the tips of a phylogeny, were not surprisingly, more variable between the two phylogenies. Another difference between the two phylogenies was how the species pool influenced the results. Reducing the scale at which the regional species pool was defined (i.e., from the island to pools in each soil type) caused a decrease for both NRI and NTI when based on the ML phylogeny but no statistically significant change based on the PHYLOMATIC phylogeny.

Based on the island species pool, one of the driest forests (Guánica, which is located at low elevation and on limestone soils) exhibited tree-wide phylogenetic clustering. Across all 15 forests, values of NRI_{ISLAND} tended to decline with mean annual precipitation, suggesting that drier forests generally comprise more phylogenetically clustered subsets of the island species pool than wetter forests. When evaluated with the reduced soil species pool, however, phylogenetic clustering of Guánica became random and only one forest in the moist life zone (Cambalache; located at low elevation and on limestone soils) had significantly clustered NTI with respect to the ML phylogeny only. The two wettest forests (Toro Negro and El Yunque, which are located on higher elevation volcanic soils) exhibited significant phylogenetic evenness in the NRI metric, although the value for Toro Negro was only significant with the ML phylogeny. One interpretation of these patterns is that water limitation represents a strong environmental filter in the dry forests and constrains the composition of local communities to the lineages that are able to persist under these harsh conditions. The issue of water stress in Puerto Rico may be exacerbated by the somewhat confounded nature of underlying geology and precipitation (Miller & Lugo 2009). Specifically, limestone soils tend to occur at lower elevations and receive less precipitation than volcanic soils. The combined influence of these variables likely compounds the effects of limited water availability for plants. In contrast, niche partitioning with respect to other factors (e.g., light use, vulnerability to pathogens) may play a stronger role in the wetter forests on volcanic soils, leading to a phylogenetically more diverse set of co-occurring species. One alternative explanation for this pattern is if *in situ* lineage diversification in Puerto Rico is a more important determinant of local species composition for higher elevation forests. For example, two closely related species of *Tabebuia*, *T. rigida* and *T. schumanniana* (Bignoniaceae) are endemic to El Yunque and Carite mountains, respectively (Axelrod 2011).

We acknowledge three main limitations in our ability to interpret these patterns. First, we did not include information on species functional traits, which are relevant to their occurrence across environmental gradients. Our interpretations depend, in part, on the degree to which functional traits relevant to species occurrence along a gradient of water availability are phylogenetically conserved. Linking key functional traits with phylogenetic relatedness would help to more strongly identify the processes that underlie compositional variation among these forests (Mayfield & Levine 2010; Adler *et al.* 2013). Second, the occurrence data we used in this analysis lacks information on species abundances. Our analysis may not detect community assembly processes that are more strongly driven by species relative abundances (i.e., dominance) than the simple presence or absence. Finally, although our null model controls for species richness within each plot, statistical power for detecting nonrandom patterns is low for forests with low species richness (Gotelli & Ulrich 2012). Nonetheless, the observed patterns provide a valuable starting point for future work aimed at addressing these limitations and providing additional insight on tree community variation across broad environmental gradients in Puerto Rico.

We found that values of NRI for each forest based on the different phylogenies were highly correlated whereas NTI values for each forest calculated with the two phylogenies were not correlated. These results reinforce the idea that low resolution among terminal tips (congeneric and confamilial taxa) may be especially problematic for recovering consistent patterns with NTI. In general, previous work has suggested that NRI may have greater power to detect nonrandom patterns of community phylogenetic structure than NTI (Letcher 2010; Vellend *et al.* 2011; Letcher *et al.* 2012).

In conclusion, our study provides a highly resolved community phylogeny for tropical trees at a regional scale: the island of Puerto Rico. We hope this regional perspective facilitates additional work to better understand the processes governing composition of local tree communities. Our case study confirms the value of a highly resolved phylogeny for detecting nonrandom patterns of phylogenetic community composition. Together with the extensive amount of existing data available in Puerto Rico on environmental conditions, land use history, species distributions and functional traits, we anticipate that the regional phylogeny provided here will help strengthen our historical perspective on the forces generating and structuring the diversity of Puerto Rican forests.

ACKNOWLEDGMENTS

We extend our gratitude to the many people who helped collect and identify specimens including especially Fabiola Areces, Frank Axelrod, Alejandro Cubiña, Freddie Perez, José Sustache, Papo Vives, Alberto Areces, Pedro Acevedo, Marcos Caraballo, Víctor José Vega López, Silvia Bibbo, Christopher J. Nytch, Marcos Rodriguez, and T.J. Agin. Danilo Chinea and provided digitized occurrence data. All necessary permits were obtained for the described study, which complied with all relevant regulations. Specifically, we thank the Departmento de Recursos Naturales y Ambientales (DRNA) of Puerto Rico for permission to collect plant specimens in the forests of Puerto Rico under DRNA permit #2011-IC-046. We also thank the herbaria and staff at the University of Puerto Rico, Rio Piedras and the US National Herbarium for permission to sample tissue from their collections; the assistance of Pedro Acevedo was invaluable for this project. This manuscript benefitted from comments by Damon Little, Robert P. Anderson, Robin Chazdon, Joel Cracraft, Dustin Rubenstein, Camilo Sanin, Brian Weeks, and two anonymous reviewers.

TABLES AND FIGURES

Table 1. Environmental characteristics and generalized results of ‘nodesig’ analysis (*i.e.,* over and underrepresented lineages) for 15 protected forests in Puerto Rico. Environmental and occurrence data are from [1]Gould *et al.* (2008), [2]Ewel & Whitmore (1973), [3]Gesch (2007), [4]Daly *et al.* (2003), [5](Bawiec 1998), [6]Little & Wadsworth (1964), and [7]Little *et al.* (1974). Forest life zones are coded as: subtropical dry (df-S), subtropical moist (mf-S), subtropical wet (wf-S), lower montane wet (wf-LM), subtropical rainforest (rf-S), lower montane rainforest (rf-LM).

Forest	**Area (ha)[1]**	**Holdridge Life Zone(s)[2]**	**Elevation Range (mean) (m)[3]**	**Mean Annual Precipitation (mm yr^{-1})[4]**	**Primary Geologic Substrate[5]**	**Species Richness[6,7]**	**Overrepresented Groups**	**Underrepresented Groups**
Aguirre	432	df-S	0–4 (1)	953	Unconsolidated	33	Combretaceae, Fabaceae	
Boquerón	623	df-S	0–5 (1)	786	Unconsolidated	19	Combretaceae	
Cambalache	649	mf-S	31–188 (157)	1,593	Limestone	152	Arecaceae, Burseraceae, Anacardiaceae, Celastraceae	Melastomataceae
Carite	2,699	wf-S, wf-LM	296–839 (657)	2,018	Volcanic	146	Lauraceae, Solanaceae, *Psychotria* (Rubiaceae), *Myrcia* (Myrtaceae), *Clusia* (Clusiaceae), Meliaceae	*Exostema*, *Guettarda*, and *Stenostomum* (Rubiaceae), Fabaceae (Mimosoidae)
Ceiba	237	df-S	0–11 (4)	1,408	Unconsolidated	5	Solanaceae, Combretaceae, Rhizophoraceae	
Guajataca	955	mf-S	192–310 (249)	1,981	Limestone	197	Nyctaginaceae, Sapindales, Meliaceae	
Guánica	3,831	df-S	0–210 (81)	876	Limestone	133	*Coccoloba* (Polygonaceae), *Crescentia* (Bignoniaceae), Capparaceae, Sapindales, Fabales	Melastomataceae, Laurales, Ericales
Monte Guilarte	1,705	wf-S, wf-LM	629–1079 (909)	2,156	Volcanic	87	*Piper* (Piperaceae), *Miconia* (Melastomataceae), Meliaceae, *Inga* (Fabaceae)	
El Yunque	11,429	mf-S, wf-S, wf-LM, rf-S, rf-LM	87–1011 (570)	3,758	Volcanic	215	Solanaceae, Melastomataceae, Meliaceae, Laurales	Rutaceae, Fabaceae, Celastraceae
Maricao	4,168	mf-S, wf-S, wf-LM	130–871 (511)	2,126	Serpentine	212	Araliaceae, Aquifoliaceae, Meliaceae	Lamiales
Rio Abajo	2,284	mf-S, wf-S	209–380 (313)	2,079	Limestone	175	Meliaceae	
Piñones	732	mf-S	0–2 (1)	1,398	Unconsolidated	31	Combretaceae, Malvaceae, *Pterocarpus* (Fabaceae)	
Susúa	1,298	mf-S	107–501 (264)	1,395	Serpentine	180	Rubiaceae	Melastomataceae
Toro Negro	2,763	wf-S, wf-LM	486–1284 (988)	2,248	Volcanic	133	Aquifoliaceae, Primulaceae, Meliaceae, *Piper* (Piperaceae), Araliaceae, Laurales, Solanaceae, Melastomataceae	*Guettarda* and *Stenostomum* (Rubiaceae), Fabaceae
Vega	482	mf-S	27–110 (67)	1,668	Limestone	86	Arecaceae, Meliaceae, Celastraceae, Moraceae, Urticaceae *Drypetes* (Putrajivaceae)	

Figure 1. A map of Puerto Rico including the 15 state forests used in this study (Gould *et al.* 2008). Forest life zones are coded as: subtropical dry (df-S), subtropical moist (mf-S), subtropical wet (wf-S), lower montane wet (wf-LM), subtropical rainforest (rf-S), lower montane rainforest (rf-LM). Refer to Table 1 for forest codes.

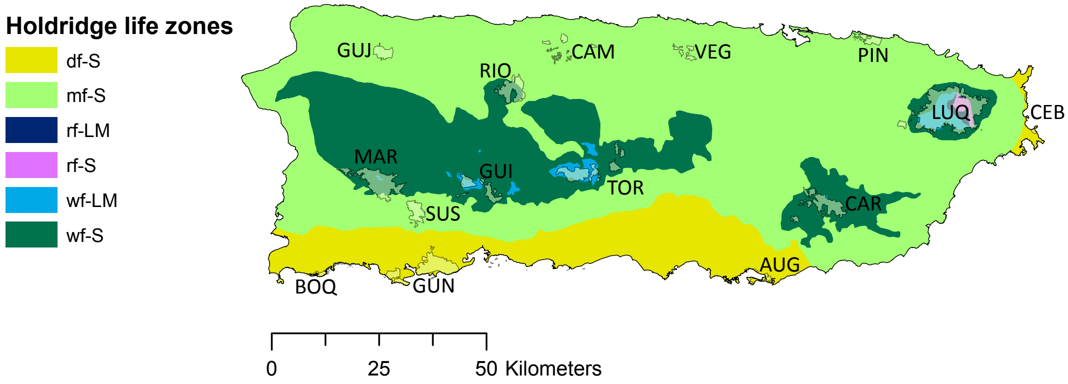


Figure 2. A maximum likelihood phylogeny constrained at the ordinal level representing 526 native and naturalized tree species of Puerto Rico (the single tree fern in the phylogeny is excluded to aid visualization). Ordinal placement according to APG III (APG III 2009) is color coded.

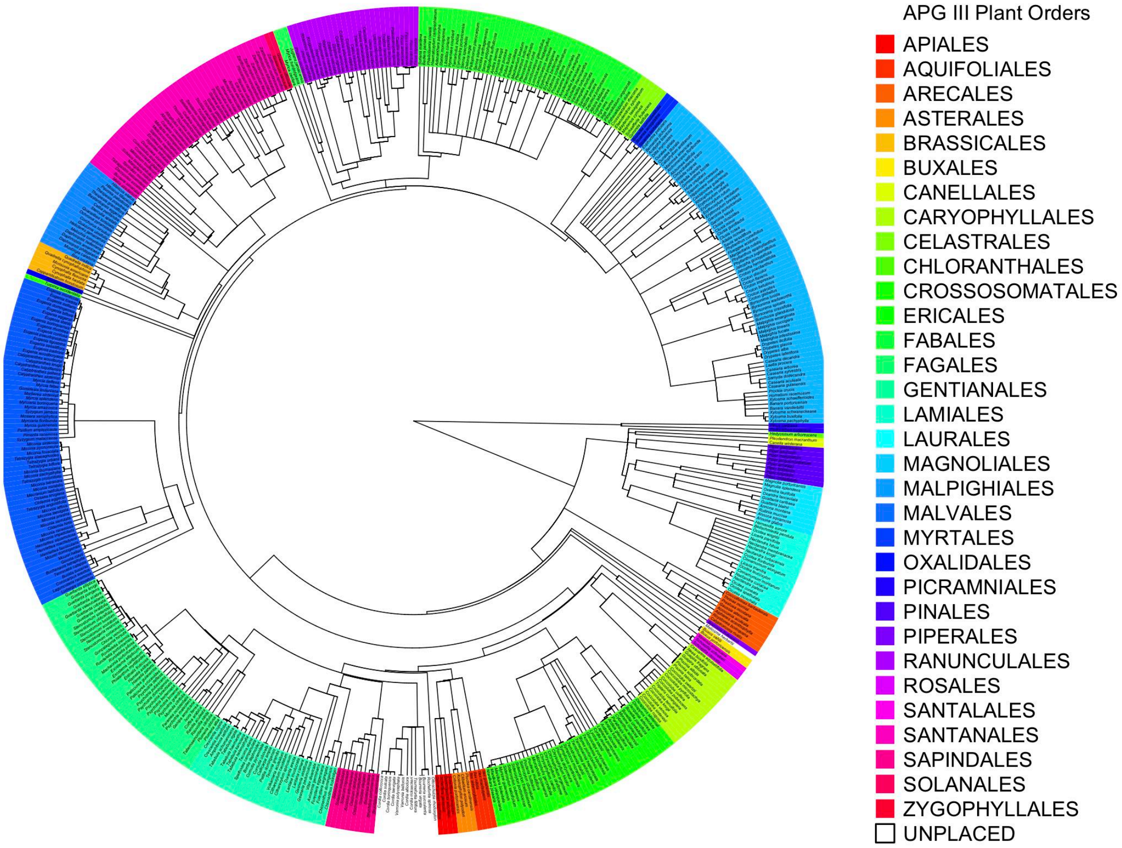
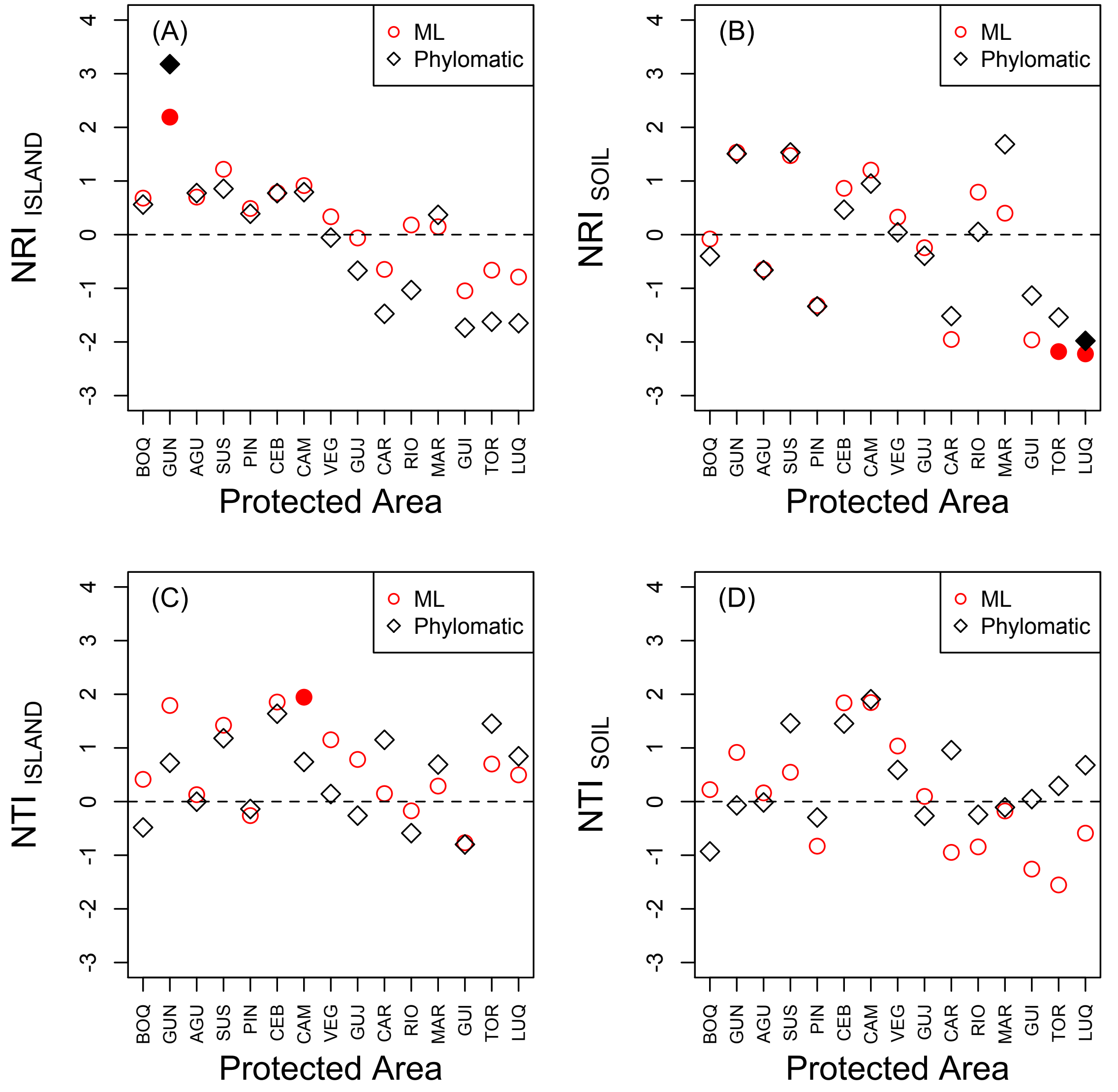


Figure 3. The net relatedness index (NRI) (A, B) and nearest taxon index (NTI; C, D) based on species occurrence records from Little & Wadsworth (1964) and Little *et al*. (1974) for 15 state forests in Puerto Rico ranked by their mean annual precipitation (Daly *et al*. 2003). Leftmost panels are based on a null model using the full island species pool; right panels (B, D) are based on species pools restricted to primary soil types. Forests are sorted from left to right in order of their mean annual precipitation. Positive values indicate phylogenetic clustering and negative values indicate phylogenetic evenness. Filled symbols indicate values that are significantly different from a null model. Refer to Table 1 for forest codes.



CHAPTER TWO:

Functional and phylogenetic dimensions of tree diversity reveal shifting assembly mechanisms across regional environmental gradients in Puerto Rico

Robert Muscarella, María Uriarte, David L. Erickson, Nathan G. Swenson, W. John Kress, and Jess K. Zimmerman

ABSTRACT

Determining how the relative strength of community assembly processes changes along abiotic gradients is a central goal in ecology. Metrics of functional and phylogenetic composition can help shed light on this issue by linking plant physiological strategies and evolutionary history to environmental heterogeneity. We censused trees in 24, 0.25-ha forest plots located along a marked gradient in precipitation and across 2 soil types (limestone and volcanic) in Puerto Rico. We characterized the composition of these plots using three functional traits with relevance to drought tolerance and light acquisition strategies (wood density [WD; g cm^{-3}], dry leaf mass per area [LMA; g cm^{-2}], and maximum height [H_{max}; m]) and a molecular phylogeny. We examined the scale at which various assembly processes act by quantifying compositional metrics using three species pools: the full island, the species recorded from the same soil type, and the species recorded from each forest. Among limestone soil plots, community-weighted mean WD and LMA decreased and H_{max} increased with increasing precipitation. Diversity of WD and H_{max} increased with precipitation, suggesting that abiotic filtering drives functional convergence in plots with limited water availability. In contrast, we found no significant shifts in community-mean trait values or functional diversity among volcanic soil plots, which occur along the wetter half of the precipitation gradient. Our study suggests geographic variation in the predominant assembly mechanisms across regional environmental gradients.

INTRODUCTION

Determining how the relative strengths of community assembly processes vary along abiotic gradients is a central goal in ecology with implications for ecosystems services (Whittaker 1967; Weiher *et al.* 2011; Hofhansl *et al.* 2014; Lasky *et al.* 2014b). One key topic is how the relative importance of competition changes with respect to abiotic conditions (Grime 1979; Tilman 1988; Goldberg & Novoplansky 1997). A large body of research in functional and phylogenetic community ecology has highlighted the hypothesis that under high resource conditions, relatively strong competitive interactions should result in high functional diversity of local communities (and high phylogenetic diversity if traits are phylogenetically conserved) as a result of limiting similarity (MacArthur & Levins 1967; Webb *et al.* 2002). In terms of coexistence theory (Chesson 2000), this hypothesis emphasizes the role of interspecific niche differences in governing local community composition. Through dominance hierarchies, however, interspecific competition can also constrain local diversity of traits that mediate fitness differences among species (Grime 2006; Navas & Violle 2009; Mayfield & Levine 2010; Kunstler *et al.* 2012; Lasky *et al.* 2014a). Thus, local patterns of functional and phylogenetic community composition may appear divergent, convergent, or random along a resource gradient depending on the conditions under which different competitive strategies are favored and the range of environmental heterogeneity examined (e.g., Bernard-Verdier *et al.* 2012; Spasojevic & Suding 2012).

Placing traits in the context of response and effect components of resource competition can help clarify hypotheses about the role of competition in structuring communities along abiotic gradients (Fig. 1; Goldberg 1990; Navas & Violle 2009; Violle *et al.* 2009). Competitive response traits enable species to persist at low resource levels and are related to Grime's (1979) concept of stress tolerance. For example, in arid environments, high wood density (WD, g cm^{-3}) conveys a survival advantage for trees because it is associated with greater resistance to drought-induced xylem cavitation (Hacke *et al.* 2001; Chave *et al.* 2009). Leaf traits associated with low water loss through transpiration (e.g., high leaf dry mass per area, LMA, g m^{-2}) also reflect tolerance to limited water availability (Wright *et al.* 2001; Poorter *et al.* 2009). Under more mesic conditions, competitive effect traits that are associated with pre-emptive or dominant resource exploitation may be favored. For instance, tall stature or rapid growth rates can convey competitive dominance for light capture that becomes a stronger determinant of community composition as water availability (and productivity) increases. Meanwhile, competitive response traits associated with shade-tolerant regeneration (*e.g.* high WD and LMA) may also represent viable strategies given the low light availability in the understory of closed canopy forests (Valladares & Niinemets 2008; Markesteijn & Poorter 2009). Overall, because the carbon and water strategies of plants are inextricably linked, understanding how traits involved with life history tradeoffs (e.g., growth / mortality) vary with respect to environmental conditions is critical for understanding how community assembly processes change along abiotic gradients. For instance, a survival advantage provided by certain traits in arid environments (e.g., high WD) is potentially diminished in wetter conditions if these are physiologically linked to lower rates of carbon acquisition.

While shifts in community-mean trait values along abiotic gradients reflect directional selection for particular life-history strategies (Shipley et al. 2011), patterns of functional diversity can also reveal variation in assembly processes along abiotic gradients (Fig. 1; Cornwell & Ackerly 2009; Bernard-Verdier *et al.* 2012; Spasojevic & Suding 2012). Navas and Violle (2009) proposed that local functional diversity should be highest in sites where the importance of resource competition for structuring communities is intermediate because traits that convey both competitive response and effect would represent viable life history strategies. According to this theory, local communities that are very strongly influenced by either abiotic stress or competitive interactions are expected to display relatively convergent functional strategies (Grime 2006; Mayfield & Levine 2010). Evaluating this hypothesis in natural systems is challenging, however, because regional environmental gradients often encompass multiple axes of environmental variation (e.g., Quesada *et al.* 2012; Coyle *et al.* 2014). As a result, directly linking particular traits to hypothesized mechanisms of resource competition across broad-scale environmental gradients may help facilitate process-based interpretation of community patterns.

Although long-term demographic data would be ideal to identify the mechanisms underlying patterns of community composition (Lasky et al. 2014a), these are unfortunately limited for diverse communities of long-lived organisms (e.g., tropical forests) across regional environmental gradients. An alternative approach is to compare the observed functional and phylogenetic composition of local communities with random expectations based on a reference pool of species (e.g., Colwell & Winkler 1984; Cavender-Bares *et al.* 2006; Swenson *et al.* 2006; Götzenberger *et al.* 2011). In practice, the reference pool is often delimited as the total list of species recorded during a study (e.g., the species list for a permanent forest plot). Patterns of community composition are, however, inherently scale-dependent and the composition of the reference pool can strongly influence this type of analysis (Gotelli & Graves 1996; Letcher 2010; Cavender-Bares *et al.* 2012; de Bello *et al.* 2012; Lessard *et al.* 2012a; Münkemüller *et al.* 2014). Numerous authors have suggested exploiting the scale-dependent nature of community patterns by conducting analyses using reference pools based on *a priori* hypotheses about the role of particular structuring processes (e.g., de Bello *et al.* 2012; Lessard *et al.* 2012a; Swenson 2013). For instance, relatively large reference pools comprise species that are absent from local communities due to a wide variety of mechanisms (*e.g.*, abiotic filtering, dispersal limitation, biogeographic history; Pärtel et al. 2011). In this context, local assemblages are likely to appear functionally and phylogenetically convergent ('clustered') (e.g., Swenson *et al.* 2006; Münkemüller *et al.* 2014). More restrictively defined reference pools have the potential to reveal assembly processes that act to structure communities at increasingly fine scales (e.g., local-scale environmental heterogeneity and biotic interactions). Ultimately, evaluating the functional and phylogenetic composition of local communities using a series of increasingly restrictive reference pools can help identify the scales at which particular filtering mechanisms drive patterns of community diversity (Lessard *et al.* 2012b; de Bello *et al.* 2013; Eiserhardt *et al.* 2013; Brunbjerg *et al.* 2014).

In this study, we quantified functional and phylogenetic aspects of tropical forest composition across marked regional gradients in precipitation and soil types on the island of Puerto Rico. We concentrated on a gradient of water availability because global climate circulation models forecast a decline of precipitation in the Caribbean region (Jury & Winter 2009; Jennings *et al.* 2014). Additionally, Puerto Rico has diverse soils that interact with precipitation to mediate water and nutrient availability for plants (Ewel & Whitmore 1973; Beinroth *et al.* 2003).

Broadly classified by underlying geologic substrate, the island's two most extensive soil types are derived from limestone and volcanic materials (Bawiec 1998), which differ in water-holding capacity and nutrient availability (Lugo & Murphy 1986; Camacho 2005). The distributions of these soil types in Puerto Rico co-vary with the island's elevation (and precipitation) gradients; limestone soils are generally found at lower (drier) elevations (range of mean annual precipitation = 698 – 2,326 mm yr^{-1}) whereas volcanic soils tend to occur at higher (wetter) elevations (718 – 4,608 mm yr^{-1}) (Bawiec 1998; Daly *et al.* 2003).

To disentangle the processes structuring tree community diversity, we focused on community-level patterns of three traits (WD, LMA, and $H_{\max}$) that are relevant to drought tolerance and light acquisition strategies. To evaluate the extent to which the functional patterns reflect shared evolutionary history as opposed to convergent response to environmental conditions, we simultaneously evaluated the phylogenetic composition of tree communities using an island-wide molecular phylogeny (Muscarella et al. 2014b). In order to better understand the scales at which different filtering processes act to determine local tree community composition, we quantified the functional and phylogenetic composition of forest plots with respect to three different reference pools: the full sample of species recorded in all sites of this study (the *island pool*), the set of species recorded from each of the two main soil types (the *soil pools*), and the set of species recorded from individual protected forests (the *forest pools*). We addressed the following specific questions:

1. *How do community mean trait values shift along a regional gradient of water availability?* Based on their relevance to physiological tradeoffs associated with drought tolerance and light-acquisition strategies, we predicted that community mean WD and LMA would decline with increasing precipitation. We expected community-mean $H_{\max}$ to increase with precipitation, as hierarchical competition for light becomes a stronger force of community structure. We expected these effects to be more pronounced among plots on the limestone soils (compared to plots on volcanic soils) because of the lower range of precipitation and water holding capacity.

2. *How does functional and phylogenetic diversity within communities vary along a regional gradient of water availability?* If tolerance to water limitation is the strongest determinant of community membership in areas of low water availability (i.e., strong abiotic filtering), we expected that dry plots on limestone soils would comprise a clustered subset of the full species pool with respect to traits related to drought-tolerance and water balance (high WD, LMA, and low H_{max}). If competitive dominance for light capture becomes a stronger determinant of local composition as water availability increases, we expected functional clustering of competitive effect traits (tall H_{max}, low WD and LMA). At the same time, we predicted that strong niche partitioning for light resources in more productive sites would lead to high functional diversity of all traits associated with light acquisition and growth strategies. Our predictions for phylogenetic patterns depend, in part, on whether these traits are conserved on the phylogeny. If biogeographic history is a strong determinant of local community composition, we expect local communities to be phylogenetically clustered regardless of functional patterns. If traits are phylogenetically conserved, we expect the phylogenetic patterns to match the functional patterns. If traits are not phylogenetically conserved, we expected random patterns of phylogenetic composition with respect to precipitation and soil type.
3. *How do patterns of community diversity change depending on how the regional species pool is defined?* We predicted that functional and phylogenetic composition of individual communities would appear less clustered (i.e., more diverse) when based on increasingly restrictive reference pools that account for various abiotic filtering mechanisms. More specifically, we predicted that local assemblages would appear relatively functionally and phylogenetically clustered in the context of the full *island* pool compared to the *soil* pools because the latter reference pool accounts for much of the abiotic filtering imposed by soil nutrient and precipitation regimes. We expected a further increase in the amount of apparent functional and phylogenetic diversity when analyses were based on individual *forest* pools because of the removal of additional abiotic filters (including dispersal limitation) as well as the increased importance of niche partitioning at smaller spatial scales.

METHODS

Study system and tree census plots

In an area of 8,740 km^2, Puerto Rico encompasses six Holdridge life zones ranging from subtropical dry forests that receive ca. 800 mm of rainfall per year to subtropical rainforests that receive upwards of 3,500 mm yr^{-1} (Ewel & Whitmore 1973; Daly *et al.* 2003). The island's complex geologic history is reflected in its rugged topography (0 – 1,338 m a.s.l.) and diverse parent soil materials, which include volcanic, limestone, alluvial, and ultramafic materials (Bawiec 1998).

In 2013, we established 24, 50m × 50m (0.25 ha) plots located in 8 protected forests in Puerto Rico (Fig. 2, Table 1). These forests occur in four life zones: subtropical dry, moist, subtropical wet, and lower montane wet forests. Each of the eight forests had three plots, and a total of twelve plots were established in soils on each of the two most extensive parent materials: limestone and volcanic. Plots were chosen randomly in mature stands lacking signs of former land-use (e.g., no plantation species and not dominated by species associated with coffee plantations) or recent natural disturbance (i.e., hurricanes). With the exception of two plots in Guánica forest, all plots were separated by at least 1 km. In each plot, we identified all woody stems (excluding lianas) ≥ 1cm diameter at 1.3 m above the ground (DBH), measured their DBH and estimated their height to the nearest 0.5 m.

Functional Traits

We collected data on three functional traits from multiple individuals of each species with deviations from the standard methods of Cornelissen *et al.* (2003) noted below. First, *wood density* (WD; g cm^{-3}) is related to a trade-off between relative growth rate and mortality (Chave *et al.* 2009; Kraft *et al.* 2010). Species with low WD tend to be relatively fast-growing compared to those with high WD because of the different levels of investment in structural materials (Wright *et al.* 2003; Chave *et al.* 2006). High WD, however, may convey resistance to drought-induced cavitation (Carlquist 1977; Hacke *et al.* 2001). For trees 10-20 cm in diameter, we measured WD using an increment borer. For species that do not typically reach this size, we included branch WD measurements and used regressions of branch and core samples from the same individuals to correct for this difference (Appendix 1; Swenson & Enquist 2008). Second, *leaf dry mass per area* (LMA; g cm^{-2}) is positively related to leaf lifespan and, more generally, to variation in plant life history strategies ranging from fast growth and high mortality by rapid photosynthetic return on carbon investment (low LMA) to slow growth and low mortality by retaining nutrients for a longer time (high LMA) (Wright *et al.* 2004; Reich *et al.* 2007). Some evidence suggests that species with low LMA may have a fitness advantage in relatively resource-rich habitats (e.g., wet and fertile soils) while species with high LMA may have a fitness advantage in more harsh conditions, including water stress conditions (Poorter *et al.* 2009; Markesteijn *et al.* 2011b; Sterck *et al.* 2014). We measured LMA on entire leaves of fully expanded foliage from the outer crown (sun-exposed, when possible) of mature trees. Finally, *maximum height* ($H_{\max}$, m) is associated with competitive ability in light-limited environments (Givnish 1995; Thomas & Bazzaz 1999) and generally increases with increasing precipitation and soil nutrients (Westoby et al. 2002). We calculated $H_{\max}$ for each species as the maximum of (i) the tallest individual observed in our study, and (ii) the maximum height reported in Little and Wadsworth (1964) and Little *et al.* (1974). Prior to analyses, we log-transformed LMA to correct for skewness. Our trait database included measurements for > 96% of the total species recorded, sampled from each forest where multiple individuals of a species were available. In this study, we used species mean trait values, which capture a majority of functional variation across plots.

Phylogenetic analyses

To evaluate functional aspects of community diversity in light of evolutionary history, we inferred a phylogenic hypothesis for 527 native and naturalized trees of Puerto Rico using three regions of plastid DNA (*rbcL*, *matK*, and *trnH-psbA*; Muscarella *et al.* 2014b). This phylogeny contains the vast majority (89%) of the island's tree flora and includes 92% of the species (accounting for 95% of the individuals) recorded in this study. We dated the phylogeny using PATHD8 (Britton et al.

2007) with fossil age constraints provided in the Appendix of Magallón & Castillo (2009). Additional details on phylogenetic analyses are provided in Muscarella *et al*. (2014b).

We calculated Pagel's λ (Pagel 1999), to quantify the phylogenetic signal of each trait – the degree to which species trait values can be explained by phylogenetic relatedness. If $\lambda = 0$ the trait has no phylogenetic signal (i.e., trait values are random with respect to phylogeny); $\lambda = 1$ if trait values correspond to a Brownian motion model of continuous trait evolution, and $\lambda > 1$ if trait values are more strongly conserved than predicted by Brownian motion. We used maximum likelihood methods to estimate lambda with 95% confidence intervals (Orme et al. 2011).

Compositional metrics and statistical analyses

To address shifts in mean trait values (*Question 1*), we quantified the community-weighted mean trait value (based on basal area) for each plot using species mean trait values. Variation in community-mean trait values along environmental gradients reflects selection for particular life-history strategies associated with environmental heterogeneity (Shipley et al. 2011). To characterize the functional and phylogenetic diversity of plots (for Questions 2 and 3), we calculated the abundance-weighted mean pairwise distance in functional trait space ($\mathrm{MPD}_{\mathrm{FUN}}$; calculated separately for each trait) and phylogenetic branch length ($\mathrm{MPD}_{\mathrm{PHY}}$) among co-occurring individuals (for saplings and trees separately) in each plot. These metrics reflect the outcome of assembly processes on the composition of local forest communities. We used a null model to convert MPD metrics to standardized effect sizes (i.e., $\mathrm{SES.MPD}_{\mathrm{FUN}}$ and $\mathrm{SES.MPD}_{\mathrm{PHY}}$). These were calculated as: $(r_{\mathrm{obs}} - \mathrm{mean}(r_{\mathrm{rand}}))/\mathrm{sd}(r_{\mathrm{rand}})$, where r is the distance in trait space (for $\mathrm{SES.MPD}_{\mathrm{FUN}}$) or phylogenetic branch length (for $\mathrm{SES.MPD}_{\mathrm{PHY}}$) among all co-occurring individuals in a plot (Webb et al. 2002). The observed value is r_{obs}. r_{rand} is a vector of values based on a null model that we used control for plot-level variation in species richness, species occupancy rates, and stem density. This consisted of generating 9999 random assemblages from the list of species in our data set by randomizing species names across the trait matrix (or phylogeny tips) and calculating a value of r_{rand} for each iteration. This procedure maintains the observed species occupancy rates and plot-level species richness while randomizing traits (or position on the phylogeny) and is appropriate when species abundances do not exhibit phylogenetic signal (Hardy 2008), as was the case in our dataset (95% CI's for λ of abundance and basal area, respectively, were: 0.00 – 0.06 and 0.00 – 0.48). Positive values of SES.MPD indicate overdispersion (or 'evenness', i.e., co-occurring individuals are less functionally or phylogenetically similar than expected by chance); negative values indicate clustering (i.e., co-occurring individuals more similar than expected by chance).

We used linear mixed models to characterize the relationship between precipitation and functional and phylogenetic composition metrics (CWM, $\mathrm{SES.MPD_{FUN}}$, and $\mathrm{SES.MPD_{PHY}}$), separately for plots on each soil type. To account for potential spatial autocorrelation among plots in the same forest, we included forest reserve as a random effect in our models. We sampled posterior distributions of model parameters with the R package arm (Gelman et al. 2006) and assigned statistical significance when 95% quantiles of parameter estimates did not overlap zero. We report marginal and conditional R^2_{GLMM} (R^2_m, the proportion of variance explained by fixed factors alone; and R^2_c, the proportion of variance explained by fixed and random factors combined) following Nakagawa & Schielzeth (2012), and using the R package MuMIn (Bartoń 2013). We also used t-tests to compare CWM, $\mathrm{SES.MPD_{FUN}}$, and $\mathrm{SES.MPD_{PHY}}$ between a subset of plots that receive similar amounts of precipitation but occur on different soil types (Guajataca, Río Abajo, Carite, and Toro Negro). We compared SES.MPD values for each plot calculated with the three different reference pools using paired t-tests. To complete our analyses, we used Phylocom (Webb et al. 2008) and R version 3.1.1 (R Development Core Team 2014) with the picante (Kembel et al. 2010) and lme4 packages (Bates et al. 2013).

RESULTS

We recorded a total of 34,301 individual trees representing 292 species (from 28 orders, 65 families, and 173 genera) across the 24 plots. Based on existing knowledge (USFS 2006; Acevedo-Rodríguez & Strong 2011; Axelrod 2011), this sample represents approximately 50% of the total tree species known to occur in Puerto Rico. Among plots on limestone soil, there were 209 species representing 25 orders, 53 families, and 141 genera. Among volcanic soil plots we recorded 138 species representing 22 orders, 48 families, and 88 genera. Fifty-five species (19% of the total) occurred in plots on both soil types and represented 17 orders, 29 families, and 47 genera. Despite the difference in total species richness between soil types, species richness rarefied by individuals did not differ between soil types and was not associated with precipitation among plots within soil types.

Question 1) Shifts in CWM values

Pairwise trait correlations were significant for all comparisons; WD and LMA were positively correlated (Pearson's $r = 0.50$, $p < 0.001$), WD and H_{max} were negatively correlated (Pearson's $r = -0.29$, $p < 0.001$), and LMA and H_{max} were negatively correlated (Pearson's $r = -0.17$, $p < 0.01$).

Among limestone soil plots, community weighted mean trait values for all three traits were significantly associated with mean annual precipitation (Fig. 3). As predicted, WD and LMA declined while H_{max} increased with increasing precipitation. Among the volcanic soil plots, in contrast, none of the community weighted mean trait values were significantly associated with mean annual precipitation. We also found no difference between community weighted mean trait values of plots that receive similar levels of precipitation but occur on different soil types ($p > 0.05$ for all traits).

Question 2) Patterns of community diversity

We found some evidence for an association between local functional diversity ($\mathrm{MPD}_{\mathrm{FUN}}$) and mean annual precipitation (Fig. 4). As we predicted, among limestone plots, local diversity of both WD and H_{max} significantly increased with precipitation when based on both the full *island* (i.e., all study plots) and *soil* (i.e., limestone plots only) reference pools (Fig. 4a,d). The relatively dry limestone soil plots comprise a significantly clustered subset of Puerto Rican trees with high WD and short stature. In contrast, local functional diversity for LMA did not vary systematically with precipitation for plots on either soil type, nor did it differ between plots of similar precipitation levels on different soil types ($p > 0.05$). When calculated with the individual forest reference pools, nearly all plots had random values of $\mathrm{MPD}_{\mathrm{FUN}}$ for all three traits and we found no evidence for significant functional overdispersion (Fig. 4h–j). Among species co-occurring in the same forest, environmental filtering does not appear to dramatically influence the distribution of local functional diversity.

Two traits (WD and LMA) showed significant phylogenetic signal that was weaker than predicted by Brownian motion (i.e., Pagel's λ was significantly different from both 0 and 1). Of these, phylogenetic signal of WD was stronger (λ [95% CI's] = 0.793 [0.627 – 0.889]) than LMA (0.557 [0.339 – 0.721]). In this dataset, $H_{\max}$ values were random with respect to the phylogeny (0.00 [0.00 – 0.556]).

Similar to the trends of WD and $H_{\max}$ diversity, phylogenetic diversity increased with precipitation among limestone soil plots when local assemblages were compared with the *island* and *soil* species pools (Fig. 5). This relationship, however, was not statistically significant when palms were excluded from the dataset or when based on the *forest* reference pools. In contrast, phylogenetic diversity for plots on volcanic soil was not associated with precipitation regardless of the reference pool used and whether palms were included.

Question 3) Effects of reference pool scaling

Overall, more restrictive reference pool definitions generally increased the apparent functional diversity of local communities. For instance, when reducing the reference pool from the *island* to the *soil* pools, functional diversity for both WD and LMA increased for all plots (Fig. 6a,b). For both of these traits, however, the shift was substantially greater among plots on volcanic soil than plots on limestone soil. For $H_{\max}$, the shift from *island* to *soil* reference pools simultaneously led to higher functional diversity for plots on limestone soil, and lower functional diversity for plots on volcanic soil (Fig. 6c).

In most cases, there were additional increases in apparent functional diversity when further restricting the reference pool to the *forest* pools (Fig. 6). Within soil types, however, several of the resulting changes varied systematically along the precipitation gradient. For instance, diversity of WD in limestone plots generally increased from the *soil* to *forest* pools. However, the driest of the limestone plots appeared substantially more diverse (higher

SES.MPD$_{FUN}$) whereas the wetter limestone plots changed relatively little (Fig. 6d). Meanwhile, the shifts in WD diversity evident in volcanic soil plots were variable and not associated with precipitation. A similar pattern was apparent for H_{max}; the driest limestone plots became apparently more diverse but the wetter limestone plots changed very little and some actually became more clustered (Fig. 6f). For LMA, functional diversity increased between the *soil* and *forests* pools (Fig. 6e).

Phylogenetic diversity also varied with the reference pool scale and differed between the two soil types. When shifting from the *island* to *soil* pools, phylogenetic diversity consistently increased for the limestone soil plots and declined for volcanic soil plots (Fig. 7). These trends did not depend on whether palms were excluded from the analysis (Fig. 7b). As with functional diversity, further restricting the reference pool to the *forest* pools changed phylogenetic diversity in ways that were partly related to precipitation. Limestone plots in the two driest forests became even less clustered whereas some of the wetter limestone plots became more clustered (Fig. 7c,d). In contrast, phylogenetic diversity for volcanic soil plots did not change consistently with this change in reference pools.

DISCUSSION

Functional and phylogenetic dimensions of community composition suggested a shift in the relative importance of different assembly mechanisms along regional abiotic gradients in forests of Puerto Rico. Specifically, the community-weighted mean trait values for LMA, WD, and H_{max} all shifted in ways consistent with our hypotheses linking functional strategies with abiotic filtering processes. Additionally, the distribution of traits and relatedness in local communities exhibited significant shifts with respect to water availability. These patterns were consistent with a strong role of abiotic filtering on competitive response traits in dry forests and an increase of niche partitioning in wet forests. Finally, scale-dependent variation in the apparent functional and phylogenetic diversity of local communities reinforced the predominant influence of different filtering mechanisms in different environmental conditions. We discuss these results in the context of how response and effect components of resource competition may drive community composition along abiotic gradients.

Question 1) Shifts in community-weighted mean traits

Directional shifts in CWM traits indicate selection for particular functional strategies along environmental gradients (Shipley et al. 2011). In our study, all three traits shifted along the precipitation gradient for limestone soil plots but not for volcanic soil plots. These shifts corresponded to our hypothesis that a strong relative importance of competitive response to drought-tolerance in dry forests gave way to stronger competitive effect for light capture in wet forests. Although the absolute range of precipitation received by plots on limestone soil was smaller than for the plots on volcanic soil (122 cm yr^{-1} vs. 187 cm yr^{-1}), climatological patterns of water availability appear to be stronger drivers of community diversity at the lower end of the precipitation gradient.

A sharp decline in the community mean WD with increasing precipitation among limestone plots, which have low water holding capacity, suggested that tolerance to drought is a major driver of community composition in these sites. Similarly, we found that community mean LMA declined with precipitation among limestone plots. Together, these responses reflect coordination between stem and leaf functional strategies to cope with water stress (Markesteijn *et al.* 2011a; Patiño *et al.* 2012). While studies across regional environmental gradients are limited, some previous studies have found similar trends in community mean WD and LMA with respect to mean annual precipitation (Chave *et al.* 2006; Swenson & Enquist 2007; Cornwell & Ackerly 2009). Although we did not examine below-ground traits in this study, previous work from Puerto Rico has shown that the root component of live-plant biomass is much higher in dry sites (particularly Guánica dry forest) than wet sites, further emphasizing the role of limited water availability in driving forest composition for these sites (Murphy & Lugo 1986).

We also observed an increase of community mean $H_{\max}$ among limestone plots that appeared to level off in the wetter half of the gradient. At the dry end of the gradient, the hydraulic limitation hypothesis (Ryan & Yoder 1997; Ryan *et al.* 2006) offers a likely explanation for constraint to maximum tree height. As water availability increases, hydraulic limitation is relaxed and the benefits of reaching tall heights in terms of maximizing carbon gain increase (Givnish 1995; Falster & Westoby 2003). However, in systems regularly exposed to wind disturbances (including Puerto Rico), tall trees have a high probability of damage during storms and most forests have a relatively low canopy height compared to forests that do not experience wind disturbances (Zimmerman *et al.* 1994; Brokaw *et al.* 2004; Mitchell 2012). This effect may be particularly acute in wet forests where hurricane return intervals are relatively short (Boose et al. 2004).

Question 2) Patterns of community diversity

Nonrandom patterns of functional and phylogenetic composition can suggest the action of particular community assembly processes (McGill et al. 2006). Contrary to the idea that niche partitioning and limiting similarity drive community structure in tropical forests, we found no evidence for significant functional overdispersion in this study, even when basing our analyses on a fairly restrictively defined reference pool (i.e., the *forest* pools). Instead, when based on the two broadest reference pools (i.e., the *island* and *soil* pools), we detected significant functional clustering of WD and $H_{\max}$ in water-limited sites and random composition of these traits in wetter sites. These results reinforce our conclusion that water availability is a primary driver of forest composition across the island of Puerto Rico, as well as among forests on limestone soil.

The results of our study provide limited support for the hypothesis proposed by Navas and Violle (2009) that functional diversity should be highest at intermediate levels of resource competition because both competitive response and effect traits would represent viable life history strategies. In some cases, functional and phylogenetic diversity appeared to peak towards the middle of the precipitation gradient but there was no significant decline in diversity with precipitation among the volcanic soil plots. Additionally, the observed shifts in community-weighted mean WD and LMA indicate selection for opposing strategies at different ends of the water availability gradient. Gaining clear support for the hypothesis of Navas and Violle (2009) is challenging, however, because of the multidimensional nature of broad scale environmental variation in natural systems. For example, we considered WD a competitive response trait under arid conditions because of its role in preventing drought-induced cavitation (Hacke *et al.* 2001; Kallarackal *et al.* 2013). In wetter sites, WD is involved in a growth/mortality tradeoff whereby some species capitalize on light acquisition via rapid growth (low WD) while others specialize on shade tolerant regeneration (high WD). Few existing studies have examined functional and phylogenetic variation along these types of multidimensional resource gradients, which are inherent features of natural systems, particularly at broad spatial scales. Future work will benefit by explicitly considering how certain traits reflect response and effect components of resource competition in the context of multidimensional resources gradients.

Phylogenetic diversity increased with precipitation among limestone plots when based on the two most inclusive species pools. The relatively low phylogenetic diversity of trees co-occurring in the driest plots may offer support to the assertion by Pennington *et al*. (2009) that association with the subtropical dry tropical forest biome represents a phylogenetically conserved niche. In particular, nitrogen-fixing leguminous trees (Fabaceae) are prominent components of subtropical dry forests, including those in Puerto Rico (Rivera 2009). Wetter sites (particularly on volcanic soils) tended to have relatively high phylogenetic diversity; a pattern that was largely driven by an abundance of palms (especially *Prestoea montana* (R. Graham) G. Nicholson) in these plots.

Question 3) Effects of reference pool scaling

Our work builds on a body of existing research that has evaluated how spatial scaling of species pools influences observed community patterns, thereby providing insight about the scales at which various filtering processes occur (e.g., Colwell & Winkler 1984; Gotelli 2000; Cavender-Bares *et al.* 2006; Swenson *et al.* 2006; Kraft *et al.* 2007; Kraft & Ackerly 2010; de Bello *et al.* 2012; Lessard *et al.* 2012a; Münkemüller *et al.* 2014). The species pools used here are based on specific factors hypothesized to strongly influence community composition across regional environmental gradients in Puerto Rico. As opposed to studies that have examined the effects of species pool scaling using arbitrary scales (e.g., various quadrat sizes within a permanent forest plot), our results demonstrate the utility of building species pools using explicit hypotheses about filtering mechanisms to gain insight to community assembly processes and generate additional hypotheses.

Interestingly, the effects of reference pool scaling differed for plots on the two soil types. Because soil types covary with precipitation in this system, it is difficult to disentangle the degree to which the *soil* pool represents the filtering effects of soil type *per se* as opposed to a more general filter of water availability. Nonetheless, plots on drier, limestone soils had a much greater change in functional diversity for both WD and LMA between the *soil* and *forest* pools whereas volcanic plots had a greater change in functional diversity between the *island* and *soil* reference pools. This result suggests a difference in the scale at which abiotic filtering structures forest communities. For forests on volcanic soils, the primary abiotic filter appears to be those that subset the full *island* pool to species that occur, generally, on volcanic soils. In other words, the processes that filter species between soil types account for the majority of community structure in volcanic soil plots. In contrast, forests on limestone soil comprise species that are filtered more strongly among forests within the soil type. Consequently, abiotic processes that filter species into different forests on limestone soil appear to occur at relatively fine spatial scales compared to forests on volcanic soil. Furthermore, among limestone soil plots, the effect of reference pool scaling on apparent functional diversity varied systematically for plots along the precipitation gradient. For example, the increase in WD and $H_{\max}$ diversity with the transition from *soil* to *forest* reference pools was strongly negatively related to precipitation. This result suggests that when compared with the *forest* pool, the driest plots (i.e., Guánica forest) appear quite locally diverse with respect to WD and $H_{\max}$. We interpret these results as evidence of small-scale habitat heterogeneity in Guánica. Previous research in Guánica forest has noted a dramatic impact of fine-scale topographic variation on moisture and organic material in the soil (Miller & Lugo 2009) as well as spatial patterns of seedling recruitment (Castilleja 1991).

Although in most cases functional diversity increased with more restrictive reference pools, volcanic soil plots appeared less diverse (more clustered) with respect to $H_{\max}$ when shifting from the full *island* to *soil* reference pools. This pattern likely emerged from the fact that the wetter limestone plots contain species with a wide range of statures even though individuals of short statured species dominate very dry plots. As such, a predominance of relatively tall statured trees in volcanic soil plots becomes even more apparent when the tall statured limestone species are excluded from the analysis.

Scaling the reference pool also affected patterns of phylogenetic diversity different for plots on the two soil types. As with the functional patterns, most limestone plots were more diverse with increasingly restrictive species pools definitions. However, similar to the results for $H_{\max}$ noted above, plots on volcanic soils were generally more phylogenetically clustered when based on the *soil* pools relative to the full *island* pool. Although the magnitude of these changes was fairly small, the difference between soil types is striking. One explanation for these patterns is that lineages represented in the volcanic soil plots are interspersed with lineages isolated on limestone soil. When limestone species are excluded, local composition of volcanic plots appears to be more clustered.

Conclusions

By examining functional and phylogenetic aspects of community diversity across regional environmental gradients, we were able to identify variation in the key processes governing Puerto Rican tree community diversity. We found particularly strong patterns of nonrandom community structure in the most water-limited sites; supporting a perspective that limited resource availability leads to strong abiotic filtering and convergence on trait values that convey strong competitive response to resource limitation. We did not find conclusive evidence for a single driver of community composition in wetter sites. Instead, community variation in these sites appears to be driven by a variety of processes that lead to random patterns of observed functional and phylogenetic composition. Our study highlights the importance of considering multi-dimensional nature of environmental gradients when investigating patterns of resource competition across large spatial scales.

ACKNOWLEDGMENTS

We thank the Departmento de Recursos Naturales y Ambientales (DRNA) of Puerto Rico as well as individual forest officers for permission to conduct this research under DRNA permit #2011-IC-046. We also thank Alexis Massol-González and Casa Pueblo for land access and accommodation. Marcos Rodriguez provided tireless assistance with all aspects of fieldwork and keen insight to the natural history of Puerto Rican trees. Fabiola Areces, Frank Axelrod, Papo Vives, and Pedro Acevedo assisted with species identification. Trait data collected by NGS was supported by a Smithsonian Center for Tropical Forest Science grant; Silvia Bibbo and Aaron Hogan helped collect additional functional trait data. This manuscript benefited from the comments of Robin Chazdon, Jesse Lasky, Benedicte Bachelot, and Benton Taylor. This work was supported by NSF DEB 1050957 to MU, NSF DEB 1311367 to MU and RM, and grants from the Earth Institute and the Institute of Latin American studies at Columbia University. The authors have no conflict of interest to declare.

TABLES AND FIGURES

Table 1. Characteristics of forest plots used in this study. Above ground biomass was estimated using allometric equations from Chave *et al*. (2005). Rarefied species richness is rarefied to 432 individuals, which was the lowest number of individuals that occurred in a single plot.

Forest	Site	Soil	Above ground biomass (Mg ha^{-1})	Basal area (m^2 ha^{-1})	Mean Annual Precipitation (cm yr^{-1})	Rarefied species richness
Cambalache	CAM1	Limestone	145.9	27.7	127	37.8
Cambalache	CAM2	Limestone	132	24.2	121	35.0
Vega	CAM3	Limestone	99.8	23.4	138	39.5
Carite	CAR1	Volcanic	164.5	26.4	169	31.2
Carite	CAR2	Volcanic	214	34.4	176	31.0
Carite	CAR3	Volcanic	84.6	22	170	38.4
Guajataca	GUA1	Limestone	186.9	27.6	170	45.5
Guajataca	GUA2	Limestone	185	24.1	165	38.4
Guajataca	GUA3	Limestone	214.1	34	158	51.5
Monte Guilarte	GUI1	Volcanic	92	16.4	182	33.9
Monte Guilarte	GUI2	Volcanic	171.6	27.8	183	34.4
Olympia	GUI3	Volcanic	73.2	15.8	191	33.1
Guanica	GUN1	Limestone	73.2	15.3	87	31.4
Guanica	GUN2	Limestone	134.4	20.2	78	37.3
Guanica	GUN3	Limestone	96.6	15.8	83	37.2
El Yunque	LUQ1	Volcanic	252.3	32	325	42.6
El Yunque	LUQ2	Volcanic	235.7	29.5	299	47.7
El Yunque	LUQ3	Volcanic	330.9	35.5	256	46.1
Río Abajo	RIO1	Limestone	95.4	18.9	171	56.3
Río Abajo	RIO2	Limestone	235.9	35.3	179	40.5
Río Abajo	RIO3	Limestone	166.2	24.9	167	49.3
Toro Negro	TOR1	Volcanic	171.6	34.8	174	28.0
Toro Negro	TOR2	Volcanic	145	28.5	179	33.7
Toro Negro	TOR3	Volcanic	167.3	24.7	182	36.1

Figure 1. Broad abiotic gradients often comprise multiple axes of environmental variation. For instance, as water availability increases productivity, leading to a reduction in understory light availability (a). The covariance of these resource gradients can lead to complex interactions between functional strategies that emphasize competitive effect and competitive response traits (b). In dry sites, traits associated with competitive response to limited water availability include those that convey resistance to drought (high WD and LMA) as well as greater hydraulic safety (short stature). Short stature may also potentially reflect a competitive effect trait if a greater proportion of biomass is allocated to underground structures for water acquisition. In wetter areas, species with low WD and LMA (and a corresponding high hydraulic connectivity) can achieve rapid growth, taking advantage of high resource conditions and thereby depleting light resources in the understory. Species with competitive response traits (i.e., traits that convey shade tolerance such as high wood density and LMA) are capable of regenerating in the understory. Predictions for shifts in community-weighted mean values and diversity for three traits are illustrated in panels (c-e).

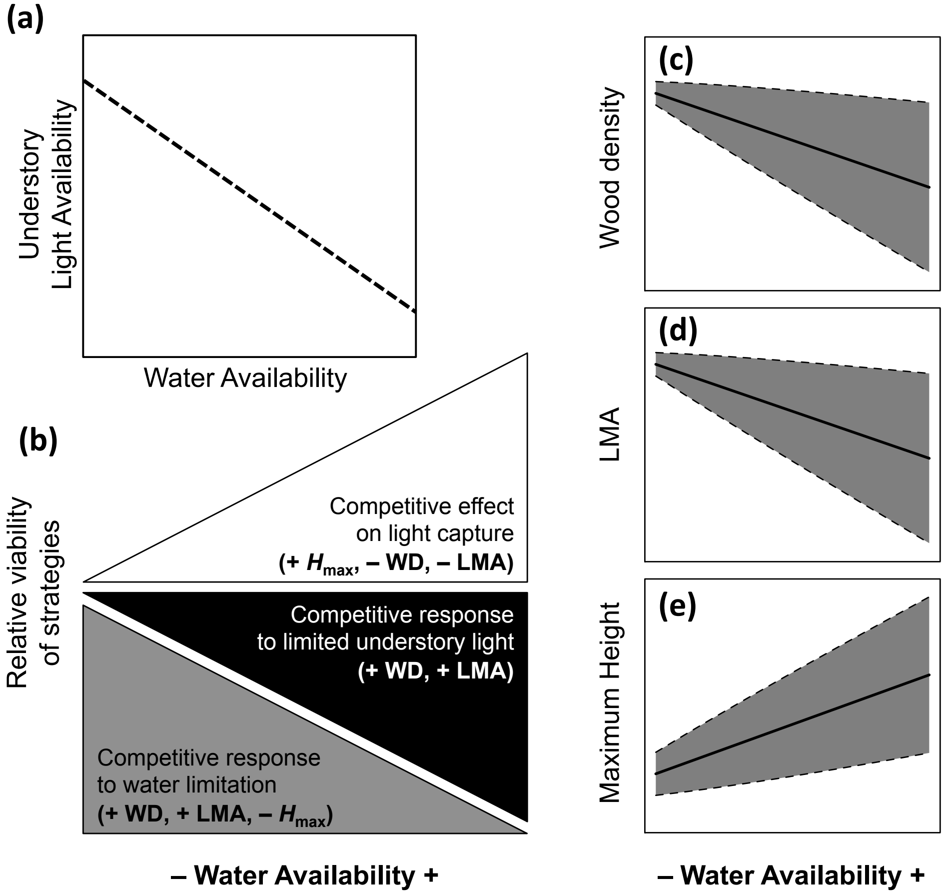


Figure 2. A map of Puerto Rico showing the locations of 24, 0.25 ha tree census plots. Circles indicate plots on limestone soils and triangles indicate plots on volcanic soils. The background shows mean annual precipitation from Daly et al. (2003) and hatched polygons represent protected forests of mainland Puerto Rico (Gould et al. 2008).

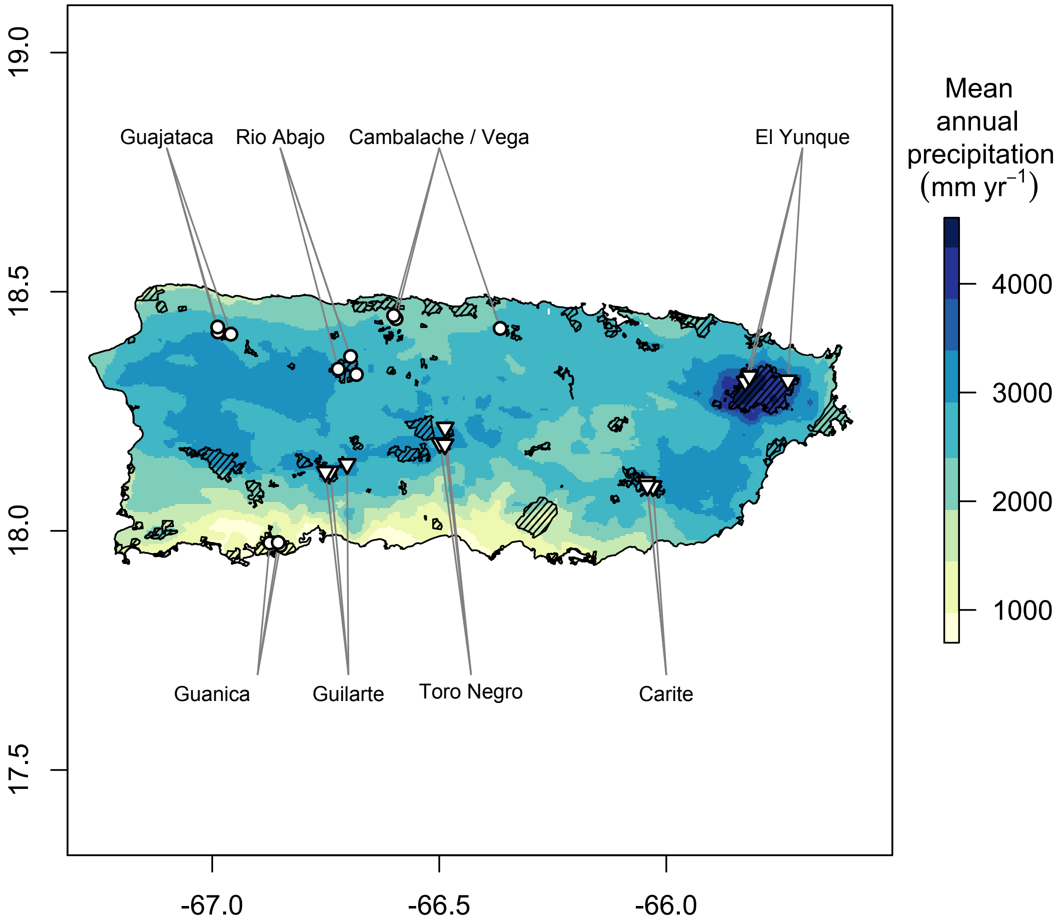


Figure 3. Community-weighted mean functional trait values for 24 0.25-ha plots in forests of Puerto Rico. Red circles and black triangles represent plots on limestone and volcanic soils, respectively. In all panels, the x-axis shows log of the mean annual precipitation for each plot. Regression lines are shown if the LMM (within soil types) had a significant slope.

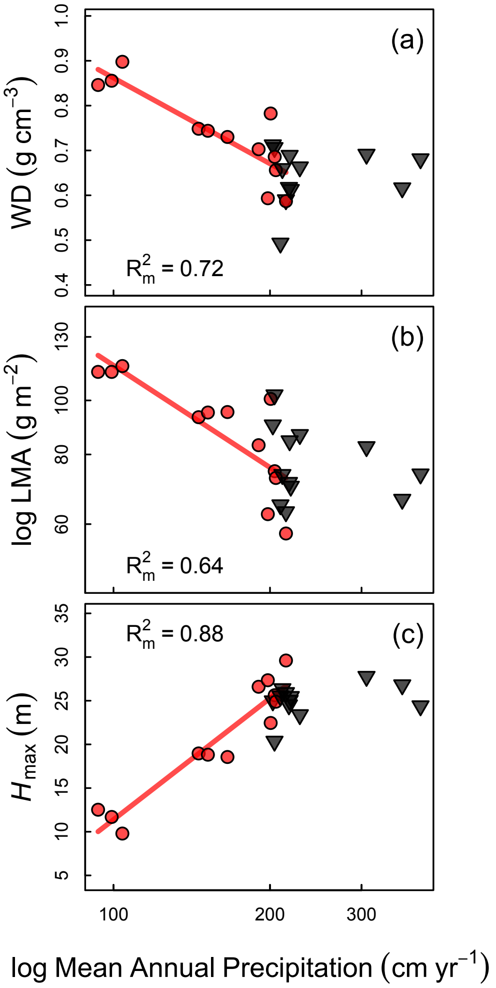


Figure 4. Standardized mean pairwise functional distance among individuals ($SES.MPD_{FUN}$) for three traits in 24 0.25 ha forest plots in Puerto Rico. Red circles and black triangles represent plots on limestone and volcanic soils, respectively. The x-axis shows mean annual precipitation for each plot on a log scale. Each column shows $SES.MPD_{FUN}$ for a different trait and each row is the $SES.MPD_{FUN}$ value calculated based on different reference pools. The top, middle, and bottom rows are based on the *island* pool, *soil* pools, and *forest* pools, respectively. In all cases, negative values indicate functionally clustered plots; positive values indicate functionally even (or 'overdispersed') plots. Points are filled if $SES.MPD_{FUN}$ was significantly different from the null expectation (α=0.05) and regression lines are shown when the LMM between $SES.MPD_{FUN}$ and precipitation (within soil types) had a significant slope.

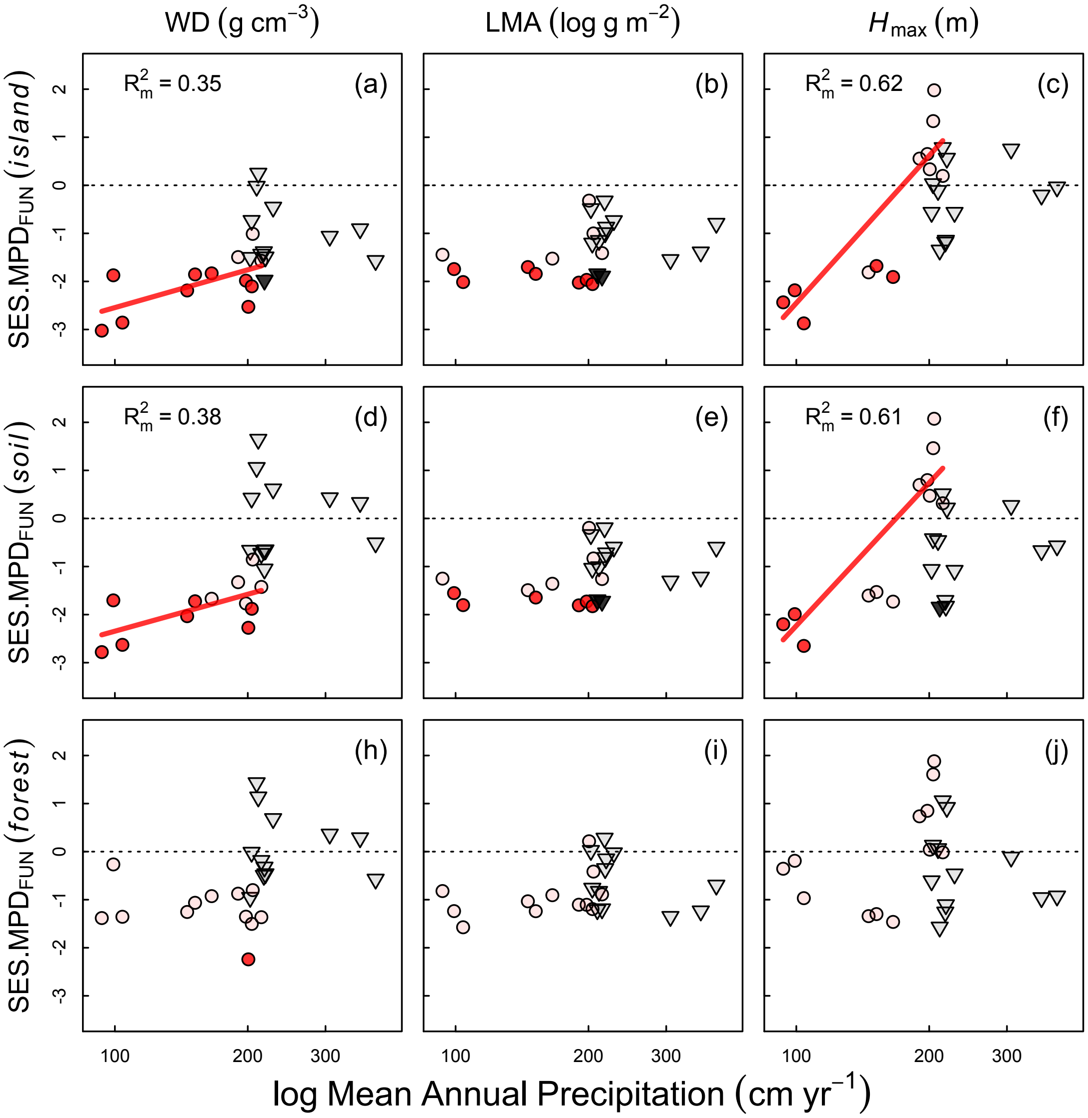


Figure 5. Standardized mean pairwise phylogenetic distance among individuals (SES.MPD$_{PHY}$) for 24, 0.25 ha plots in Puerto Rico. Red circles and black triangles represent plots on limestone and volcanic soils, respectively. The x-axis shows mean annual precipitation for each plot on a log scale. Left and right columns show SES.MPD$_{PHY}$ with palms included and excluded, respectively. Each row is the SES.MPD$_{PHY}$ value calculated based on different reference pools. The top, middle, and bottom rows are based on the *island* pool, *soil* pools, and *forest* pools, respectively. In all cases, negative values indicate phylogenetically clustered plots; positive values indicate phylogenetically even (or 'overdispersed') plots. Points are filled if SES.MPD$_{PHY}$ was significantly different from the null expectation (α=0.05) and regression lines are shown when the LMM between SES.MPD$_{PHY}$ and precipitation (within soil types) had a significant slope.

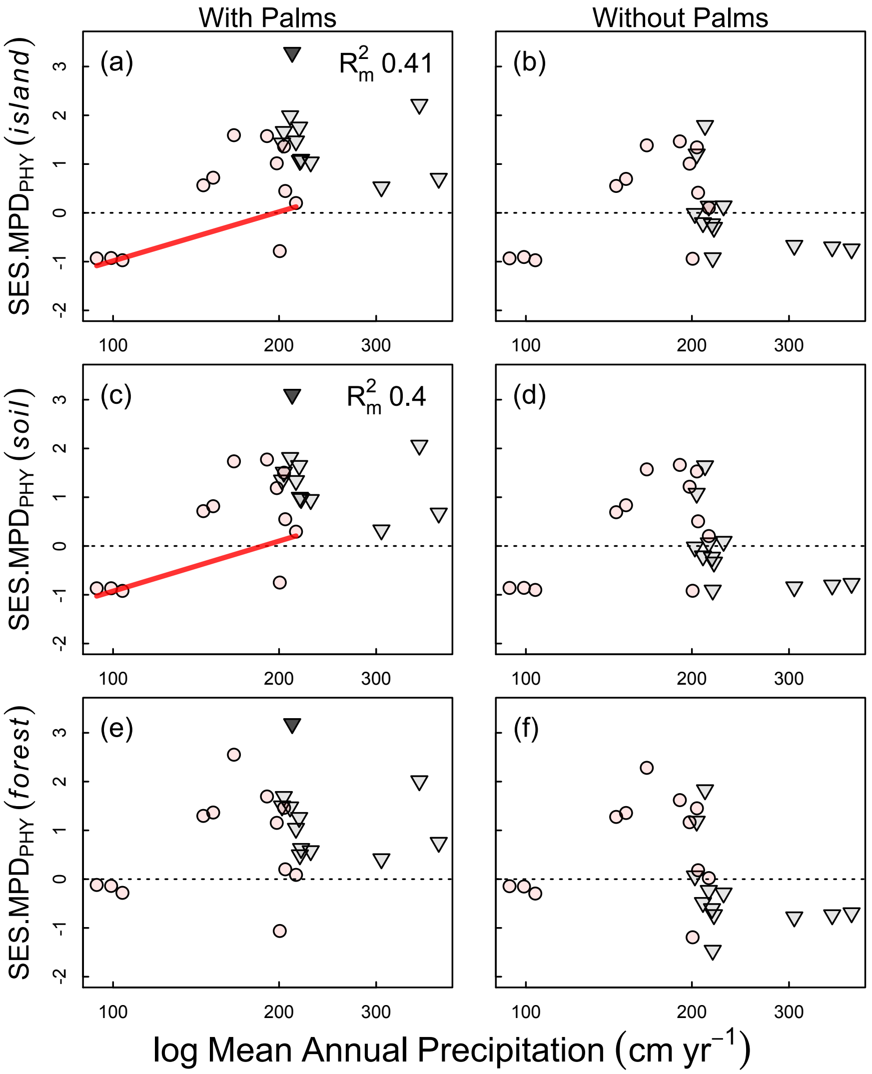


Figure 6. Difference in $SES.MPD_{FUN}$ values when calculated with different reference pools. Panels (a-c) correspond to the difference between $SES.MPD_{FUN}$ for each trait calculated with the *island* and *soil* reference pools, panels (d-f) correspond to the difference between $SES.MPD_{FUN}$ for each trait calculated with the *soil* and *forest* reference pools. Values were scaled such that positive values indicate an increase in functional diversity with the more restrictive reference pool definition whereas negative values indicate a decrease in functional diversity with the more restrictive reference pool definition.

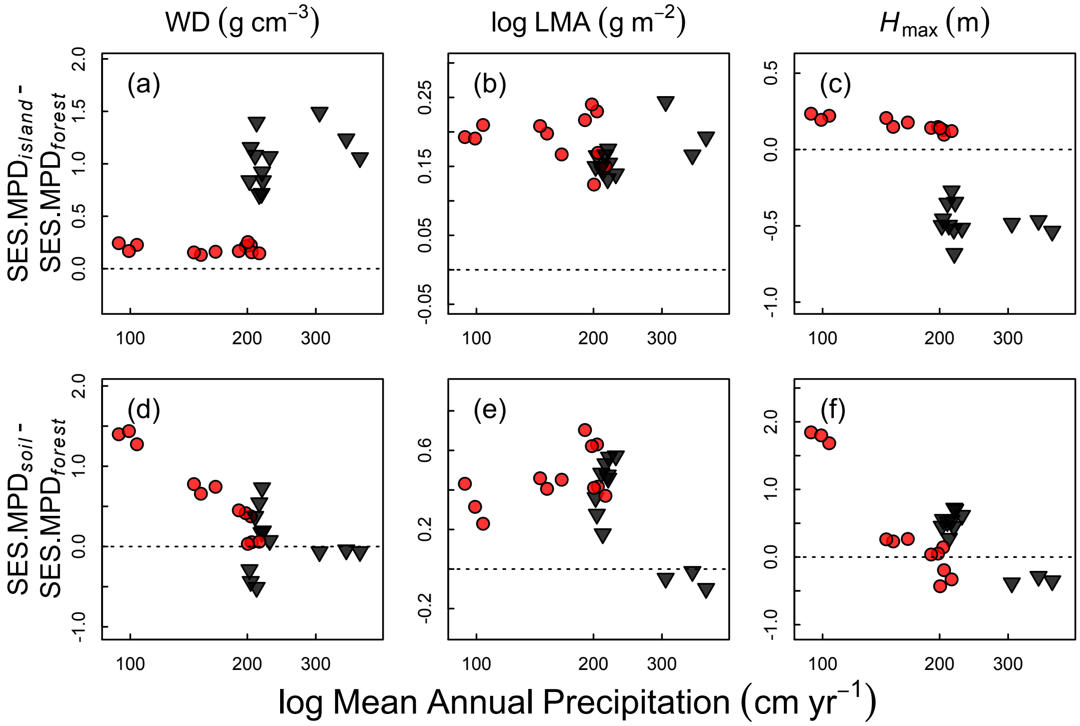
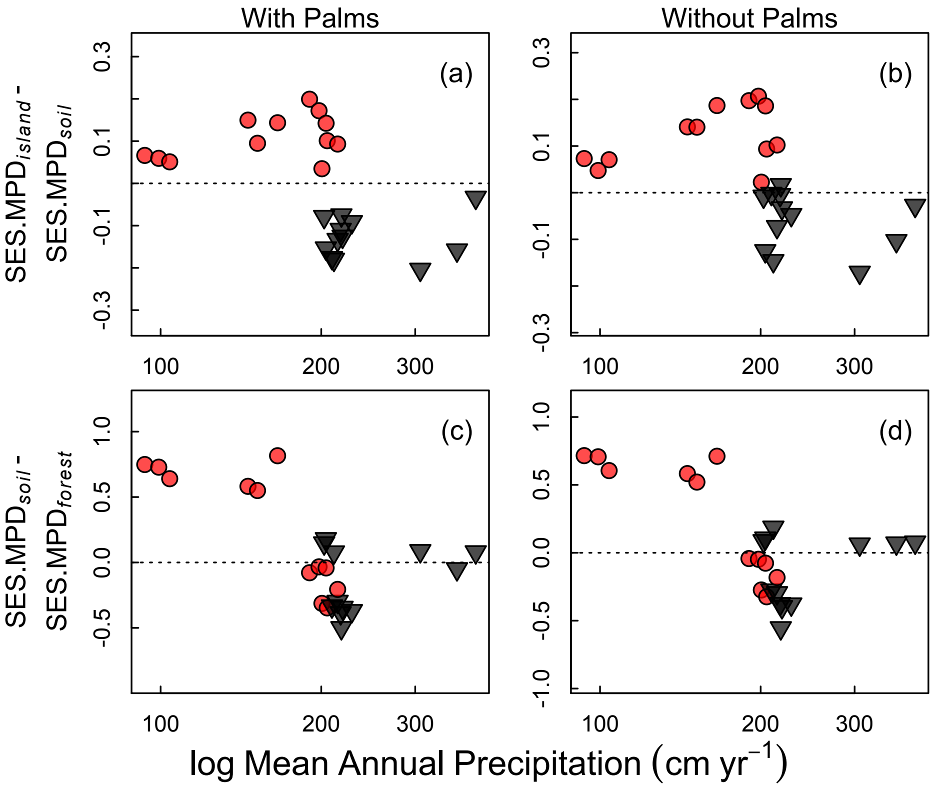


Figure 7. Difference in SES.MPD$_{PHY}$ when calculated based on different reference pools. Panels (a) and (b) correspond to the difference between SES.MPD$_{PHY}$ calculated with the *island* and *soil* reference pools; panels (c) and (d) correspond to the difference between SES.MPD$_{PHY}$ calculated with the *soil* and *forest* reference pools. Values were scaled so that positive values indicate an increase in phylogenetic diversity with the more restrictive reference pool definition whereas negative values indicate a decrease in phylogenetic diversity with the more restrictive reference pool definition.



CHAPTER THREE:
Functional convergence and phylogenetic divergence during secondary succession of subtropical wet forests in Puerto Rico

Robert Muscarella, María Uriarte, T. Mitchell Aide, David L. Erickson, Jimena Forero-Montaña, W. John Kress, Nathan G. Swenson, and Jess K. Zimmerman

ABSTRACT

Secondary forests are increasingly recognized for their potential contributions to biodiversity conservation, ecosystem services, and the vitality of local economies. Understanding of the processes that drive forest succession, however, remains elusive. We examined shifts in the functional and phylogenetic composition of sapling and adult tree communities of different ages in subtropical wet forests of Puerto Rico with the aim of providing insights into community assembly processes. We predicted that early aged stands would be dominated by species with high colonization ability (e.g., low seed mass) and acquisitive life history traits (e.g., low wood density [WD; g cm^{-3}] and leaf dry mass per area [LMA, g m^{-2}]) whereas older stands would be marked by increased abundance of species with conservative resource-use traits (e.g., high WD and LMA) and traits that confer advantages for light competition (e.g., tall maximum height, H_{max}) and shade-tolerant regeneration (e.g., large seeds). We also evaluated whether shifts in functional and phylogenetic diversity suggested an increased importance of either niche partitioning or competitive hierarchies during forest succession. Community-weighted mean LMA, H_{max} and seed mass all increased with stand age, reflecting a shift in dominance from species with acquisitive resource use strategies and smaller seeds to species with more conservative resource use and larger seeds. Some of these shifts were only evident among saplings or adults, suggesting differences in the importance of trait-mediated processes for different life stages. We also found evidence for increased importance of competitive hierarchies via a decline in functional diversity of seed mass and H_{max} across succession. Meanwhile, phylogenetic diversity tended to increase with forest age (for both saplings and adults), a pattern that was driven by high

abundance of palms in some old forest plots. These results suggest that old forest plots contain relatively distantly related species of tall stature that produce large seeds. Our study offers some support for the coordination of plant life history strategies across seed and vegetative traits. Overall, placing species along quantitative axes of niche differentiation allows us to gain insight to the processes driving community dynamics during succession.

INTRODUCTION

Today, the majority of the Earth's forests are undergoing secondary succession (FAO 2010). Secondary forests are increasingly recognized for their potential contributions to biodiversity conservation, ecosystem services, and the vitality of local economies (Chazdon 2014). Increased understanding of these forests can simultaneously clarify the long-term implications of land use change for the services these ecosystems provide and shed light on the processes that drive the assembly of ecological communities (Gardner *et al.* 2009; Lasky *et al.* 2014a).

Previous research in forest succession has generally been more successful in predicting structural aspects of change (e.g., biomass accumulation) than in characterizing the dynamics of species composition (Guariguata & Ostertag 2001; Chazdon 2003). By placing species along quantitative axes of niche differentiation, however, trait-based approaches are helping to deepen our understanding of the mechanisms that drive community assembly (Cornwell & Ackerly 2009; Lasky *et al.* 2014a). For example, declining light availability (i.e., increasing leaf area index) is a fundamental feature of humid forest succession (Bazzaz 1979; Bazzaz & Pickett 1980). Two non-mutually exclusive mechanisms that can potentially explain the community dynamics that result from this shift in light conditions are the competition-colonization tradeoff and the successional niche hypothesis (Fig. 1).

Under the competition-colonization tradeoff, species with traits associated with high colonization ability (e.g., small seeds and high fecundity) occupy available sites in early succession because of the relative dispersal limitation of competitively dominant species (Levins & Culver 1971; Tilman 1994). As succession proceeds, the importance of non-local seed input declines, environmental conditions change, and competitively dominant species exclude initial colonists, leading to a directional shift in the community-mean trait values. It is important to note that 'competitive dominance' may refer either to competitive effect traits that convey hierarchical advantages (e.g., tall adult stature: Givnish 1995; Falster & Westoby 2005) or competitive response traits that convey tolerance and the ability to regenerate under low resource conditions (e.g., large seeds; Grubb 1977; Goldberg 1990; Muller-Landau 2010).

According to the successional niche hypothesis, species that can rapidly exploit abundant resources can temporarily dominate early successional or high resource sites (e.g., gaps) regardless of a colonization advantage because species with more conservative resource-use strategies have slower growth rates (Pacala & Rees 1998). Over time, however, species able to persist and regenerate under low resource availability dominate (e.g., shade tolerant species). Under these conditions, species with acquisitive resource strategies (e.g., low wood density [WD g cm^{-3}], leaf dry mass per area [LMA, g cm^{-1}], and seed mass [g]) are predicted to flourish in early succession when light levels in the mid- and understory are high. Species with conservative traits associated with shade-tolerance (e.g., high WD, LMA, seed mass) are predicted to increase in abundance during later stages of succession (i.e., stand thinning stages).

Shifts in functional diversity during succession may also indicate the action of various community assembly processes. Successional niche theory predicts that functional diversity should be initially low when sites are dominated by early successional species with high colonization and acquisitive resource-use strategies (Kinzig & Pacala 2001; Grime 2006; Lasky *et al.* 2014b). During intermediate stages of succession, a mixture of early- and late-successional species often co-occur, resulting in a peak of functional diversity (Fig. 1). Over time, low resource conditions associated with late successional forests may constrain functional diversity for some traits that convey competitive dominance (e.g., tall $H_{\max}$) or the ability to tolerate low resource conditions (e.g., large seeds). Alternative scenarios may arise, however, if functional diversity is generally high when local environmental heterogeneity is a relatively strong driver of community composition. For example, canopy gaps that occur from mortality of large trees in the later stages of succession increase heterogeneity of the light environment, potentially favoring a high diversity of life-history strategies at small spatial scales (Bazzaz & Pickett 1980; Loehle 2000).

Conversely, patchy colonization of open sites (e.g., recently abandoned pastures) can result in high levels of light heterogeneity during early stages of succession. Highlighting the need for additional work, the few existing studies that have examined changes in plant functional diversity during tropical forest succession have reported mixed results that appear to depend, in part, on the abiotic context (*e.g.*, humid versus dry forests), the role of natural disturbances in generating community dynamics, and the life stage(s) considered (e.g., Swenson *et al.* 2012; Lasky *et al.* 2014a; Lohbeck *et al.* 2014).

In addition to shifts in functional diversity during succession, patterns of phylogenetic diversity can provide further insight to processes driving succession (e.g., Letcher 2010; Norden *et al.* 2012). If traits driving successional change exhibit phylogenetic niche conservatism (i.e., that close relatives are functionally similar), shifts in phylogenetic diversity may mirror those predicted for functional diversity outlined above. However, phylogenetic patterns can also reflect other historical processes that may or may not be evident in the functional similarity of species. Few studies, however, have simultaneously examined both functional and phylogenetic aspects of forest succession and we have very limited knowledge about how these different dimensions of community diversity interactively drive succession dynamics.

In this study, we examined shifts of functional and phylogenetic composition of sapling and adult communities along a regionally replicated chronosequence in subtropical wet forests of Puerto Rico. We integrated data on four functional traits related to colonization ability, competitive dominance, resource use and acquisition strategies, and a highly resolved molecular phylogeny to address the following questions:

1. *How does the community-level functional composition of forest plots vary with respect to forest age?* We predicted that species with high colonization ability and acquisitive life history traits (e.g., small seed size, low LMA and WD) would dominate early in succession whereas later stages of succession would be marked by increasing abundance of species with conservative resource-use strategies (e.g., high LMA and WD) and traits associated with dominance for light competition (e.g., tall $H_{\max}$) and shade-tolerant regeneration (e.g., large seed size). We expected these changes to be more pronounced among saplings than adult

trees due to their relatively rapid turnover.

2. *How do local functional and phylogenetic diversity vary with respect to forest age?* If shifts in functional diversity are driven by directional shifts in functional strategies, we expected functional diversity to peak during intermediate stages of succession. On the other hand, if an increase in light heterogeneity drives patterns of trait diversity during succession, we expected the diversity of traits related to light-use strategies (e.g., LMA) to increase with stand age. Alternatively, functional diversity of some traits may decline during with stand age if lower overall resource availability leads to convergence on traits that convey competitive dominance (e.g., tall $H_{\max}$) and the ability to tolerate low resource conditions (e.g., large seed size). If the traits underlying these functional shifts are phylogenetically conserved, we expected corresponding patterns of phylogenetic diversity during succession. Alternatively, if traits are convergent with respect to phylogeny, we expected opposite or random phylogenetic patterns. We expected these changes to be more pronounced among saplings than adults due to the relatively rapid community turnover of saplings.

METHODS

Study area and tree community surveys

We combined data from 48 secondary forest plots located in three regions of subtropical wet forests in Puerto Rico (Table 1). Most plots (n = 45) were sampled in 1995 (Aide et al. 2000) and two were newly established in 2012. We also included a subset of data from 2010 from a portion of the Luquillo Forest Dynamics Plot (LFDP) that had experienced high land-use intensity during the 1930's (i.e., cover class 1 in Thompson et al. 2002). We sub-sampled the LFDP data to match the sample design of the other plots (*see below*). Sites were selected with as similar abiotic conditions as possible in order to help meet the assumptions of the chronosequence approach (Walker et al. 2010). The age of each plot was determined with aerial photographs taken between 1936 and 1994, and/or interviews with local landowners. Together, the plots ranged in age from 4 to 91 years of natural forest growth at the time of census.

In each plot, we counted, identified, and measured the freestanding woody individuals (excluding lianas and tree ferns) that were 1 to 10 cm DBH (‘saplings’) in four parallel 1 x 50 m transects each separated by 10 m (200 m^2). We also measured larger individuals (≥ 10 cm DBH; ‘adults’) in two parallel transects measuring 10 x 50 m and separated by 10 m (1,000 m^2). For each plot, we calculated rarefied species richness based on the minimum number of stems (of both size classes) recorded from a single plot (n = 47). We accounted for potential correlations among plots in the same region using a mixed effects model (see *Statistical analyses*).

Functional traits

We collected data on four functional traits from multiple individuals of each species using methods outlined by Cornelissen *et al.* (2003) with some exceptions noted below. *Leaf dry mass per area* (LMA; g cm^{-2}) is a key component of the leaf economics spectrum and positively related to leaf longevity; species with low LMA tend to grow fast and achieve rapid return on carbon investment whereas species with high LMA have slow growth and retain nutrients for longer periods (Wright et al. 2004). We measured LMA on entire leaves (including petiole) of fully expanded foliage from the outer crown (sun-exposed, when possible) of mature trees. *Wood density* (WD; g cm^{-3}) is related to a trade-off between relative growth rate and mortality; species with low WD tend to grow relatively rapidly because of low investment in structural materials while species with high WD have higher resistance to structural damage and drought-induced mortality (Poorter *et al.* 2008; Chave *et al.* 2009). For most species, we measured WD using cores taken with an increment borer from trees 10 – 20 cm in diameter at 1 m off the ground. For species that do not reach this size, we measured WD using branch material (*see* Appendix 1; Swenson and Enquist 2008). *Maximum height* ($H_{\max}$; m) is the tallest observed height for a species and is positively associated with competitive dominance in light-limited environments and time to reproduction (Givnish 1995). *Dry seed mass* (SS; g) corresponds to a seed size–fecundity tradeoff (Smith & Fretwell 1974); empirical evidence suggests that smaller seeds have a dispersal advantage and can persist for long periods in the soil (Dalling & Brown 2009) whereas large, recalcitrant seeds tend to have higher probability of initial establishment, particularly in shady conditions (Muller-Landau 2010; Muscarella *et al.* 2013). We log-transformed species means for all traits except WD prior to analyses and we excluded palms from analyses involving WD. Across traits, we obtained data for an average of 93% of the observed species (range = 90 – 99%), representing > 97% of the individuals per plot.

Community phylogeny

We constructed a community phylogeny for the species recorded in our study by pruning a highly-resolved DNA barcode phylogeny of 527 native and naturalized trees of Puerto Rico (Muscarella *et al.* 2014b). Briefly, this phylogeny is based on three regions of chloroplast DNA (*rbc*L, *mat*K, and *trn*H-*psb*A) and contains 95% of the species (> 99% of the individuals) in our dataset. To aid in interpretation of functional and phylogenetic patterns, we used maximum likelihood to methods to estimate Pagel's λ (Pagel 1999; Orme *et al.* 2011), which quantifies the phylogenetic signal of each trait (i.e., the degree to which phylogenetic relatedness predicts functional similarity). Pagel's λ normally range from 0 to 1: $\lambda = 0$ if trait values have no phylogenetic signal (i.e., they are random with respect to phylogeny) and $\lambda = 1$ if trait values correspond to a Brownian motion model of continuous trait evolution.

Statistical analyses

To evaluate how the average functional composition of forest plots varies during succession (*Question 1*), we quantified community-weighted mean trait values separately for saplings and adults in each plot based on species' relative basal area and mean trait values. To help understand how traits may act in a coordinated manner to drive community patterns, we measured pairwise correlations between traits using Pearson's correlation coefficient with species-mean trait values. To determine how functional and phylogenetic diversity change during succession (*Question 2*), we calculated the abundance-weighted mean pairwise distance in functional trait space ($\mathrm{MPD}_{\mathrm{FUN}}$; calculated separately for each trait) and phylogenetic branch length $\mathrm{MPD}_{\mathrm{PHY}}$) among co-occurring individuals (for saplings and trees separately) in each plot. Analyses based on abundance-weighted community mean values rather than basal area provided very similar results and here we present only the results from abundance-weighted analyses. We used a null model to control for the variation in species richness, species occupancy rates, and stem density among plots. Specifically, we calculated the standard effect size of the MPD metrics (SES.MPD) as: $(r_{\text{obs}} - \text{mean}(r_{\text{rand}}))/\text{sd}(r_{\text{rand}})$, where r is the distance in trait space (for $\text{SES.MPD}_{\text{FUN}}$) or phylogenetic branch length (for $\text{SES.MPD}_{\text{PHY}}$) among all co-occurring individuals in a plot. The observed value is r_{obs} and r_{rand} is a vector of values based on a null model. For the null model, we generated 9999 random assemblages from the list of species in our data set by randomizing species names across the trait matrix (or phylogeny tips) and calculating a value of r_{rand} for each iteration. This procedure maintains the observed species occupancy rates and plot-level species richness while randomizing traits (or position on the phylogeny) and is appropriate when species abundances do not exhibit phylogenetic signal (Hardy 2008), as was the case in our dataset (95% CI's for λ of abundance = 0.00 – 0.13 for saplings; 0.00 – 0.75 for adults). Positive values of SES.MPD indicate overdispersion (or 'evenness', i.e., co-occurring individuals are less similar than expected by chance); negative values indicate clustering (i.e., co-occurring individuals more similar than expected by chance).

We used linear mixed models to regress community-mean trait values, $\text{SES.MPD}_{\text{FUN}}$, and $\text{SES.MPD}_{\text{PHY}}$ against stand age, including a random effect to account for potential correlations among plots in the same region (Table 1). We used AIC to select between models with and without a squared term for stand age to account for the possibility of non-linear responses. We sampled posterior distributions of model parameters using the R package 'arm' (Gelman et al. 2006) and assigned statistical significance when 95% quantiles did not overlap zero. We calculated marginal and conditional R^2_{GLMM} (R^2_m, variance explained by fixed factors alone; and R^2_c, variance explained by fixed and random factors combined) using the R package 'MuMIn' (Nakagawa & Schielzeth 2012; Bartoń 2013). All analyses were conducted in R 3.1.1 (R Development Core Team 2014) using Phylocom 4.2 (Webb *et al.* 2008), picante (Kembel *et al.* 2010), FD (Laliberté & Shipley 2011), caper (Orme *et al.* 2011), lme4 (Bates *et al.* 2013).

RESULTS

In total, we recorded 8,804 individuals of 135 species. Rarefied species richness ranged from 5.8 – 20.5, and increased with stand age. Stands in relatively early stages of succession tended to be dominated by *Alchornea latifolia* Sw., *Tabebuia heterophylla* (DC.) Britton, and two common exotic species, *Spathodea campanulata* P. Beauv. and *Syzygium jambos* (L.) Alston. Species that increased in abundance in medium-aged stands included *Guarea guidonia* (L.) Sleumer and *Ocotea leucoxylon* (Sw.) Laness. In older stands, several species often associated with late-successional forests were abundant (e.g., *Dacryodes excelsa* Vahl and *Manilkara bidentata* (A. DC.) A. Chev.), as well as a long-lived pioneer (*Cecropia schreberiana* Miq.) and the abundant palm, *Prestoea montana* (R. Graham) G. Nicholson.

We recorded a wide range of trait variation across species (range for LMA = 15.9 – 388.6 g m^{-2}; WD = 0.17 – 0.98 g cm^{-3}, H_{max} = 2.5 – 35.0 m, and dry seed mass = 2 x 10^{-5} – 35.2 g). Based on species mean trait values, we found significant correlations for three of the six trait pairs. Specifically, LMA was positively correlated with WD ($r = 0.38$, $p < 0.001$) and H_{max} ($r = 0.25$, $p < 0.01$), and H_{max} was positively correlated with seed mass ($r = 0.37$, $p < 0.001$). Correlations were not significant ($p > 0.05$) between seed mass and LMA ($r = 0.17$), WD and H_{max} ($r = -0.03$), and WD and seed mass ($r = 0.14$).

Question 1) Shifts in community-weighted mean traits during succession

For saplings (stems 1 – 10 cm DBH), directional shifts of community-mean values for LMA, H_{max}, and seed mass were consistent with our hypothesized transition from species with acquisitive resource strategies and high colonization abilities in early succession to species with more conservative resource strategies and higher competitive dominance in later stages of succession (Fig. 2; Table 2). Specifically, the community-mean for each of these traits increased with forest age, and linear models provided better fits than quadratic models (Table 2). As we predicted, these patterns were similar but less marked for adults (stems > 10 cm DBH): adult community-weighted mean LMA and seed mass both increased with forest age (Fig. 2). Contrary to our predictions, however, community-weighted mean WD was not significantly associated with forest age for either saplings or adults.

Question 2) Functional and phylogenetic diversity during succession

We detected two significant shifts in functional diversity during succession, and linear models provided better fits than quadratic models in all cases (Fig. 3; Table 2). Among saplings, functional diversity ($\text{SES.MPD}_{\text{FUN}}$) of seed size declined with forest age (Fig. 3d). Among adults, $\text{SES.MPD}_{\text{FUN}}$ of $H_{\max}$ declined with forest age, indicating a trend towards lower diversity of stature among adult trees during succession (Fig. 3i). All four traits showed some degree of phylogenetic signal (i.e., $\lambda > 0$), although none strongly enough to match a Brownian motion model of trait evolution (i.e., $\lambda < 1$). Seed mass had the strongest phylogenetic signal (λ [95% CI] = 0.84 [0.59 – 0.96]), followed by WD (0.69 [0.43 – 0.85]), $H_{\max}$ (0.57 [0.29 – 0.78]), and LMA (0.47 [0.09 – 0.74]).

Phylogenetic diversity also shifted during succession and the patterns contrasted with those reported for functional diversity. Among saplings, $\text{SES.MPD}_{\text{PHY}}$ increased with forest age (Fig. 3e). Early successional plots had particularly high abundances of several species of *Miconia* (Melastomataceae) as well as two species from the Bignoniaceae, *T. heterophylla* and *S. campanulata*. The increase of phylogenetic diversity in older plots was related to increasing abundance of palms (especially *P. montana*) and a generally more even distribution of species abundances. The pattern was not significant when palms were excluded from the analysis, indicating their primary role in driving this pattern (Fig. 3f). Similarly, phylogenetic diversity for adults also increased with forest age but not significantly when excluding palms (Fig. 3k,l).

DISCUSSION

Question 1) Shifts in community-mean traits

Overall, shifts of community-mean trait values reported here support both the competition-colonization tradeoff and successional niches as important drivers of successional dynamics in Puerto Rican wet forests. First, the observed increase in mean LMA during succession reflects a shift of communities dominated by species with acquisitive light-capture strategies that capitalize on rapid exploitation of high light conditions towards communities dominated by species with more conservative strategies that make higher initial investments in leaf construction with long-term payoffs (Wright et al. 2004). This finding is consistent with other studies of forest succession in wet tropical forests (e.g., Reich *et al.* 1995; Poorter *et al.* 2004), highlighting the link between high LMA and shade-tolerant life history strategies. (Wright et al. 2004, Reich et al. 1995, Poorter et al. 2004)

Because WD is generally negatively correlated with mortality rates (Chave et al. 2009), we expected that during succession long-lived species with high WD would increase in abundance and basal area, leading to an increase in community-level WD (Swenson *et al.* 2012; Lasky *et al.* 2014a). We did not, however, find a significant shift in community mean WD of trees with stand age. To date, the few other studies that have examined community-mean shifts of stem traits during succession have found inconsistent patterns that partly depend on the environmental context (e.g., dry vs. wet forests; Lohbeck et al. 2013). We suspect that at least three separate factors contribute to our results. First, *C. schreberiana*, a long-lived pioneer and gap-specialist with low WD had a high basal area in the older successional plots, driving down community mean WD. The relatively high abundance of *C. schreberiana* in these plots may be related to repeated disturbance by hurricanes, which result in a patch light environment and soil disturbance such as tree tip-ups (Flynn et al. 2010). Second, early successional plots varied widely in their community mean WD. High dispersal limitation or stochastic colonization of early successional plots followed by subsequent biotic and abiotic filtering processes could lead to such a pattern, particularly if species WD values are unrelated to their dispersal abilities. In fact, in our dataset, seed mass and WD were not significantly correlated. Finally, while short-term droughts likely have some influence on community dynamics in our system, traits conveying drought-tolerance (*e.g.*, WD) may be less important than traits associated with light capture in wet forests. Nonetheless, if plant strategies are coordinated across tissues and resources (*e.g.*, Reich 2014), traits associated with hydraulic and light capture strategies should be correlated and exhibit similar shifts of community-mean values (*but see* Baraloto et al. 2010). In fact, a positive correlation between WD and LMA (*r* = 0.38) was the strongest of the six pairwise trait correlations we examined. However, the non-significant trend in community-mean WD during succession indicates that a coordinated tradeoff across leaf and stem traits can only partly explain community-level trends in functional composition.

Consistent with the hypothesized link between seed size and colonization ability (Coomes & Grubb 2003; Muller-Landau 2010; Muscarella *et al.* 2013), we observed an increase in community-mean seed mass during succession. Theories of species coexistence that invoke seed size-related tradeoffs ultimately involve the energetic tradeoff between seed size and fecundity (Smith & Fretwell 1974). Species with small seeds tend to have a colonization advantage through high fecundity and low dispersal limitation (Coomes & Grubb 2003). Meanwhile, large seeded species are associated higher probability of establishment, particularly under low resource conditions (Muller-Landau 2010; Muscarella *et al.* 2013). Our results provide further support for the role of seed size in mediating a competition-colonization tradeoff, and as an important dimension of the successional niche. Additionally, seed mass was positively correlated with $H_{\max}$, and we observed an increase in mean $H_{\max}$ among saplings. Together, these results reveal a shift in linked life-history strategies during succession and lend support to a study by Chazdon *et al*. (2010) who reported that tree stature provided one of the strongest indications of successional strategies of tropical trees in Costa Rica. Contrary to our original expectations and the patterns observed among saplings, we did find a change in mean $H_{\max}$ among adult stems. As discussed with WD above, one likely explanation is the repeated effect of hurricane damage on Puerto Rican forests (Boose *et al.* 2004; Uriarte *et al.* 2009; Flynn *et al.* 2010). Interspecific variation in adult stature can best be understood in terms of the benefits gained in light competition versus the construction and maintenance costs of achieving tall heights (Givnish 1995; Falster & Westoby 2005). In Puerto Rico, an island with a high incidence of hurricanes, the competitive advantages of achieving tall heights may be outweighed by greater susceptibility to hurricane damage (Canham et al. 2010).

A number of recent studies have examined correlations among traits with the goal of determining the extent to which they represent coordinated axes of life history tradeoffs (e.g., Wright *et al.* 2007; Baraloto *et al.* 2010; Freschet *et al.* 2010; Fortunel *et al.* 2012; Patiño *et al.* 2012). This issue is of key importance for understanding how functional patterns relate to specific mechanisms of community assembly. In a recent review, Reich (2014) advanced the hypothesis that strong selection along trait trade-off axes, combined with biophysical constraints, has resulted in a single *fast-slow* tradeoff axis that is consistent across plant tissues and major resource types. In line with this hypothesis, Patiño *et al.* (2012) reported significant correlations of branch xylem density with both LMA and seed mass in a sample of >600 Amazonian trees. In our study, species trait values of LMA (leaf economics) were positively correlated with WD (stem economics), providing additional support for coordination of life history strategies across traits. The moderate strength of these correlations, however, indicates a substantial amount of independent trait variation. This finding has been emphasized by several other studies of Neotropical trees that reported independence between leaf and stem economic traits (Baraloto *et al.* 2010; Fortunel *et al.* 2012). Overall, we agree with Funk and Cornwell (2013) in that progress in understanding life history tradeoffs will likely come from considering how environmental context influences the strength and direction of observed trait correlations. Additionally, data on physiological traits (e.g., maximum assimilation rate, A_{max}) will help advance our understanding of the specific processes underlying community variation compared with composite traits (*e.g.*, WD and LMA) that reflect a combination of numerous physiological processes.

Question 2) Functional and phylogenetic diversity during succession

Studies using patterns of functional or phylogenetic diversity to infer community assembly processes have often emphasized a dichotomous interpretation of patterns: environmental filtering leads to convergence while limiting similarity leads to evenness (or overdispersion) (Webb et al. 2002). However, more recent perspectives emphasize that (1) multiple assembly processes can lead to similar patterns of community composition, and (2) opposing assembly mechanisms can obscure overall patterns of functional composition (e.g., Grime 2006; Lasky *et al.* 2014a). For example, competitive hierarchies can lead to functional (and, potentially, phylogenetic) convergence rather than divergence (Grime 2006; Mayfield & Levine 2010).

Additionally, fine-scale environmental heterogeneity can lead to niche partitioning and functional evenness at small spatial scales (Adler *et al.* 2013).

We observed two successional shifts in functional diversity: a decline in seed mass diversity among saplings and a decline in H_{max} diversity among adults. Together with the observed increase of community-mean values for seed mass, we conclude that our late successional plots comprised species that produce relatively large seeds compared to the pool of species found across the successional gradient. These results further support a shift in the relative importance of colonization ability towards shade-tolerant regeneration over succession.

In contrast to our expectations, functional diversity was not hump-shaped (as expected from a mixture of early- and late-successional species co-occurring in intermediate aged stands), nor did it increase with stand age (as expected from strong niche partitioning during late states of succession). Instead, our results suggest functional convergence, potentially driven by competitive hierarchies, in late successional stands. The few existing studies that have examined changes in functional diversity during tropical forest succession have reported mixed results in terms of directional shifts in trait diversity variation. In Mexican wet tropical forests, Lohbeck *et al*. (2014) reported an increase in leaf trait diversity, suggesting an increasing role of niche partitioning or competition for light during succession. In seasonally dry tropical forests of Mexico, Bhaskar *et al.* (2014) found mixed evidence for a change in functional dispersion during succession. Specifically, they reported an increase in LMA diversity with succession but diversity of WD and leaf N did not change. Lasky *et al*. (2014a) also found increased diversity of LMA at neighborhood scales during succession in Costa Rica, however, diversity of WD decreased with succession. In part, the differences in these results emerge from the differing abiotic context of the study system – particularly when comparing patterns of succession between wet and dry tropical forests, which face different limiting resources (e.g., Lohbeck et al. 2013).

Considering the degree of phylogenetic signal in the traits with decreasing functional diversity (i.e., seed mass and H_{max}), we expected similar declines of phylogenetic diversity during succession. In contrast, however, we found an increase in phylogenetic diversity during succession. For adults, this pattern was explained by the colonization of some late successional plots by distantly related palms. Among saplings, however, the pattern remained significant even when palms were excluded from the analysis. One potential explanation for this difference between life stages may the larger sample size of saplings compared to adult trees. In any case, our results are largely consistent with several previous studies that have reported an increase in phylogenetic diversity during succession (Letcher 2010; Ding *et al.* 2012; Letcher *et al.* 2012; Norden *et al.* 2012; Whitfeld *et al.* 2012). Differences between functional and phylogenetic patterns may result from a relatively weak phylogenetic signal for traits that drive succession. Although some evidence supports the idea that general successional status of tropical trees is phylogenetically conserved (e.g., Letcher 2010; Norden *et al.* 2012), few functional traits appear to be consistently related to successional habitat specialization (Letcher *et al.* in review). Even though all four traits included in our study displayed some degree of phylogenetic signal, none were conserved strongly enough to match a Brownian motion model of trait evolution. We are only aware of two other studies that have simultaneously examined functional and phylogenetic diversity during tropical succession (Ding *et al.* 2012; Swenson *et al.* 2012), and additional work will be required to better understand their interactive effects.

Conclusions

Research in tropical forest succession has been relatively successful in characterizing structural aspects of change (e.g., biomass accumulation) but understanding successional shifts of species composition has presented substantial challenges (Guariguata & Ostertag 2001; Gardner *et al.* 2009). Our study builds on current work that places species along quantitative axes of trait differentiation to help discriminate and synthesize existing theories of community assembly (Fukami *et al.* 2005; Lasky *et al.* 2014a; Lasky *et al.* 2014b). Using this framework, we linked shifts in functional and phylogenetic aspects of tree community variation with mechanisms of forest succession and life history tradeoffs. Ultimately, these advances will help establish a more general understanding of the processes driving tropical succession, a critical step in light of the increasing dominance of secondary forests worldwide.

ACKNOWLEDGMENTS

We thank the landowners who generously allowed us to work on their property. Robin Chazdon, Duncan Menge, Bénédicte Bachelot, Jesse Lasky, Naomi Schwartz, and Benjamin Taylor provided helpful comments on previous drafts of this manuscript. This work was supported by NSF DEB 1050957 to MU, NSF DEB 1311367 to MU and RM, and grants from the Earth Institute and Institute of Latin American studies at Columbia University.

TABLES AND FIGURES

Table 1. A list of plots included in this study. Values of mean annual precipitation marked with * indicate that precise plot coordinates were not available. In these cases, we estimated precipitation based on an average value from nearby locations with similar elevation. Observed species richness and stem density are quantified separately for saplings (stems < 10 cm DBH) and adults (stems ≥ 10 cm DBH). Individual-based rarefied species richness is based on the total sample form each plot (both size classes) and rarefied to the minimum number of stems from a single plot (n = 47).

						Species richness			Stem density		
Site	Region	Age when sampled (years)	Elevation (m)	Mean Annual Precipitation (cm year^{-1})	Year sampled	Saplings	Adults	Rarefied total	Saplings	Adults	Total
CAR10	CAR	25	300	165*	1995	11	10	8.71	159	82	241
CAR12	CAR	12	325	165	1995	6	5	5.78	86	76	162
CAR14	CAR	37	600	177*	1995	22	11	13.21	231	60	291
CAR15	CAR	37	500	177*	1995	16	21	17.53	141	64	205
CAR17	CAR	25	500	177*	1995	18	16	14.09	151	32	183
CAR19	CAR	4	300	165*	1995	7	1	6.17	57	13	70
CAR20	CAR	4	350	165*	1995	8	3	7.21	62	3	65
CAR21	CAR	24	350	165	1995	17	14	11.75	166	127	293
CAR23	CAR	76	710	196	1995	12	13	12.5	67	74	141
CAR24	CAR	75	650	201*	1995	21	18	16.92	52	97	149
CAR25	CAR	76	700	199	1995	15	20	15.3	44	124	168
CAR26	CAR	12	550	177	1995	17	17	16.71	123	49	172
CAR27	CAR	36	550	177	1995	23	18	20.52	87	71	158
CAR28	CAR	76	700	192	1995	17	17	16.52	51	97	148
CAR29	CAR	4	400	165*	1995	11	9	14.39	35	18	53
CAR3	CAR	76	670	201	1995	20	14	14.52	62	122	184
CAR31	CAR	13	500	177*	1995	19	4	10.79	141	75	216
CAR32	CAR	51	700	182	1995	25	13	16.32	131	75	206
CAR33	CAR	52	650	201*	1995	27	16	17.62	112	71	183
CAR34	CAR	52	650	201*	1995	9	13	11.16	49	93	142
CAR35	CAR	77	650	201*	1995	15	13	12.02	56	125	181
CAR37	CAR	76	680	201	1995	21	14	17.02	74	84	158
CAR38	CAR	51	670	177	1995	24	19	17.55	163	105	268
CAR5	CAR	76	625	177	1995	23	19	18.79	162	88	250
CAR6	CAR	24	625	177	1995	26	12	16.58	152	67	219
CAR7	CAR	36	625	172	1995	23	14	13.78	228	93	321

Table 1. (Continued)

						Species richness			Stem density		
Site	**Region**	**Age when sampled (years)**	**Elevation (m)**	**Mean Annual Precipitation (cm year^{-1})**	**Year sampled**	**Saplings**	**Adults**	**Site**	**Region**	**Age when sampled (years)**	**Elevation (m)**
CAR8	CAR	4	620	201	1995	11	1	5.78	320	3	323
CAR9	CAR	13	300	165*	1995	9	4	7.91	96	38	134
EV1	LUQ	52	540	359	2012	17	24	19.05	40	104	144
LFDPCC1	LUQ	91	360	359	2010	4	13	10.61	7	97	104
LUQ10	LUQ	9	60	208	1995	8	3	7.43	97	9	106
LUQ11	LUQ	14	70	250	1995	11	3	6.51	309	14	323
LUQ12	LUQ	10	30	175	1995	11	7	9.52	166	13	179
LUQ2	LUQ	59	50	208	1995	14	11	14.47	81	30	111
LUQ3	LUQ	15	90	230	1995	11	6	8.53	114	75	189
LUQ6	LUQ	25	220	251	1995	15	11	11.96	151	64	215
LUQ7	LUQ	37	200	251	1995	13	14	11.34	228	111	339
LUQ8	LUQ	37	50	208	1995	18	6	12.39	214	28	242
LUQ9	LUQ	60	40	208	1995	16	14	13.41	158	58	216
SB3	LUQ	82	175	205	2012	21	18	19.77	58	50	108
UTU17	UTU	27	160	187*	1995	13	6	11.32	103	49	152
UTU18	UTU	20	200	186*	1995	15	12	13.47	97	56	153
UTU19	UTU	35	210	186*	1995	18	8	14.9	83	31	114
UTU20	UTU	35	210	186*	1995	10	5	11	28	19	47
UTU21	UTU	25	325	183*	1995	14	10	9.98	179	45	224
UTU22	UTU	30	330	189*	1995	14	7	9.62	151	63	214
UTU23	UTU	30	340	189*	1995	15	13	13.09	103	45	148
UTU24	UTU	12	320	189*	1995	9	1	6.14	171	1	172

Table 2. Detailed results of linear mixed models including marginal and conditional R^2_{GLMM} (R^2_m, the proportion of variance explained by fixed factors alone; and R^2_c, the proportion of variance explained by fixed and random factors combined) follow Nakagawa & Schielzeth (2012). Traits are abbreviated as: LMA = leaf dry mass per area (g m^{-2}); WD = wood density (g cm^{-3}); H_{max} = maximum height (m); and Seed mass = seed dry mass (g). Models using community-weighted mean values, functional diversity, and phylogenetic diversity are indicated by CWM, MPD_{FUN}, and MPD_{PHY}, respectively. ΔAIC values correspond to the comparison between the two models that are identical except for the inclusion of a quadratic term for forest age ($Years^2$).

Response	**Trait**	**Size Class**	**Parameter**	**Estimate**	**Std. Error**	***t*-value**	**R^2_m**	**R^2_c**	**ΔAIC**
CWM	LMA	Saplings	(Intercept)	3.975	0.045	89.053	0.26	0.34	0
			Years	0.003	0.001	4.25			
CWM	LMA	Saplings	(Intercept)	3.985	0.064	62.512	0.26	0.33	20.6
			Years	0.003	0.003	0.813			
			$Years^2$	0	0	0.223			
CWM	LMA	Adults	(Intercept)	0.602	0.019	31.932	0.00	0.04	0
			Years	0	0	0.155			
CWM	LMA	Adults	(Intercept)	0.617	0.028	21.665	0.01	0.04	21.7
			Years	-0.001	0.002	-0.584			
			$Years^2$	0	0	0.648			
CWM	WD	Saplings	(Intercept)	2.596	0.061	42.383	0.36	0.42	0
			Years	0.006	0.001	5.354			
CWM	WD	Saplings	(Intercept)	2.438	0.077	31.838	0.43	0.43	14.5
			Years	0.016	0.004	3.788			
			$Years^2$	0	0	-2.554			
CWM	WD	Adults	(Intercept)	-4.512	0.31	-14.569	0.32	0.32	0
			Years	0.032	0.007	4.71			
CWM	WD	Adults	(Intercept)	-4.368	0.501	-8.715	0.32	0.32	16.2
			Years	0.022	0.028	0.782			
			$Years^2$	0	0	0.368			
CWM	H_{max}	Saplings	(Intercept)	3.979	0.064	61.718	0.10	0.15	0
			Years	0.003	0.001	2.375			
CWM	H_{max}	Saplings	(Intercept)	3.956	0.1	39.674	0.10	0.16	19.6
			Years	0.004	0.005	0.859			
			$Years^2$	0	0	-0.299			
CWM	H_{max}	Adults	(Intercept)	0.511	0.028	18.513	0.06	0.12	0
			Years	0.001	0.001	1.845			

Table 2. (Continued)

Response	Trait	Size Class	Parameter	Estimate	Std. Error	*t*-value	R2m	R2c	ΔAIC
CWM	H_{max}	Adults	(Intercept)	0.529	0.042	12.684	0.07	0.13	21.1
			Years	0	0.002	-0.102			
			$Years^2$	0	0	0.566			
CWM	Seed mass	Saplings	(Intercept)	3.074	0.036	85.991	0.04	0.04	0
			Years	0.001	0.001	1.4			
CWM	Seed mass	Saplings	(Intercept)	3.137	0.057	55.328	0.08	0.08	18.7
			Years	-0.003	0.003	-1.026			
			$Years^2$	0	0	1.414			
CWM	Seed mass	Adults	(Intercept)	-3.428	0.633	-5.416	0.12	0.41	0
			Years	0.023	0.008	3.035			
CWM	Seed mass	Adults	(Intercept)	-3.205	0.778	-4.117	0.12	0.41	15.8
			Years	0.008	0.031	0.254			
			$Years^2$	0	0	0.496			
MPD_{FUN}	LMA	Saplings	(Intercept)	-0.598	0.178	-3.357	0.00	0.00	0
			Years	-0.001	0.004	-0.215			
MPD_{FUN}	LMA	Saplings	(Intercept)	-1.039	0.293	-3.544	0.08	0.14	13.7
			Years	0.03	0.015	1.936			
			$Years^2$	0	0	-2.071			
MPD_{FUN}	LMA	Adults	(Intercept)	-0.266	0.188	-1.418	0.02	0.02	0
			Years	-0.004	0.004	-0.891			
MPD_{FUN}	LMA	Adults	(Intercept)	-0.318	0.304	-1.046	0.02	0.02	17.3
			Years	0	0.017	-0.004			
			$Years^2$	0	0	-0.218			
MPD_{FUN}	WD	Saplings	(Intercept)	-0.352	0.251	-1.403	0.03	0.19	0
			Years	-0.005	0.004	-1.375			
MPD_{FUN}	WD	Saplings	(Intercept)	-0.417	0.346	-1.207	0.03	0.2	17.3
			Years	-0.001	0.016	-0.06			
			$Years^2$	0	0	-0.282			
MPD_{FUN}	WD	Adults	(Intercept)	0.258	0.303	0.851	0.07	0.33	0
			Years	-0.009	0.004	-2.169			

Table 2. (Continued)

Response	Trait	Size Class	Parameter	Estimate	Std. Error	*t*-value	R2m	R2c	ΔAIC
MPD_{FUN}	WD	Adults	(Intercept)	0.677	0.375	1.804	0.11	0.36	14.1
			Years	-0.037	0.016	-2.33			
			$Years^2$	0	0	1.842			
MPD_{FUN}	H_{max}	Saplings	(Intercept)	-0.499	0.207	-2.406	0.00	0.00	0
			Years	-0.002	0.004	-0.349			
MPD_{FUN}	H_{max}	Saplings	(Intercept)	-0.261	0.359	-0.728	0.02	0.02	16.5
			Years	-0.017	0.019	-0.872			
			$Years^2$	0	0	0.813			
MPD_{FUN}	H_{max}	Adults	(Intercept)	-0.065	0.213	-0.307	0.01	0.01	0
			Years	-0.003	0.005	-0.609			
MPD_{FUN}	H_{max}	Adults	(Intercept)	0.529	0.354	1.495	0.10	0.10	13.1
			Years	-0.04	0.019	-2.152			
			$Years^2$	0	0	2.062			
MPD_{FUN}	Seed mass	Saplings	(Intercept)	-0.562	0.178	-3.156	0.15	0.15	0
			Years	-0.01	0.004	-2.756			
MPD_{FUN}	Seed mass	Saplings	(Intercept)	-0.694	0.31	-2.241	0.15	0.15	17.2
			Years	-0.002	0.016	-0.126			
			$Years^2$	0	0	-0.522			
MPD_{FUN}	Seed mass	Adults	(Intercept)	0.429	0.342	1.253	0.02	0.26	0
			Years	-0.005	0.005	-1.121			
MPD_{FUN}	Seed mass	Adults	(Intercept)	0.844	0.449	1.88	0.05	0.23	15.3
			Years	-0.031	0.021	-1.501			
			$Years^2$	0	0	1.283			
MPD_{PHY}	With palms	Saplings	(Intercept)	-0.615	0.262	-2.344	0.19	0.19	0
			Years	0.019	0.006	3.35			
MPD_{PHY}	With palms	Saplings	(Intercept)	-0.786	0.424	-1.855	0.19	0.19	16.4
			Years	0.031	0.024	1.312			
			$Years^2$	0	0	-0.517			
MPD_{PHY}	With palms	Adults	(Intercept)	-0.363	0.295	-1.231	0.04	0.04	0
			Years	0.009	0.006	1.385			

Table 2. (Continued)

Response	Trait	Size Class	Parameter	Estimate	Std. Error	*t*-value	R2m	R2c	ΔAIC
MPD_{PHY}	With palms	Adults	(Intercept)	-0.948	0.465	-2.041	0.09	0.09	13.9
			Years	0.05	0.026	1.907			
			$Years^2$	0	0	-1.612			
MPD_{PHY}	No palms	Saplings	(Intercept)	-0.645	0.355	-1.817	0.43	0.66	0
			Years	0.029	0.004	7.3			
MPD_{PHY}	No palms	Saplings	(Intercept)	-0.333	0.43	-0.775	0.45	0.66	16
			Years	0.009	0.017	0.556			
			$Years^2$	0	0	1.163			
MPD_{PHY}	No palms	Adults	(Intercept)	-0.36	0.341	-1.054	0.00	0.13	0
			Years	0.001	0.006	0.155			
MPD_{PHY}	No palms	Adults	(Intercept)	-0.508	0.529	-0.96	0	0.16	16.5
			Years	0.01	0.025	0.398			
			$Years^2$	0	0	-0.377			

Figure 1. Two potential mechanisms that underlie successional dynamics are a tradeoff in competition-colonization ability (CC; a, b) and successional niches (SN; c, d). Under the CC tradeoff, a colonization advantage (e.g., low seed mass) enables weak competitors to occupy early successional habitats where competitive dominants are absent because of dispersal limitation (shaded area of a). As the relative importance of non-local seed input declines in later stages of succession, the benefits of early colonization are diminished and weak competitors (shaded area of b) are excluded by species with traits that convey competitive dominance (e.g., conservative resource-use strategies, tall $H_{\max}$). Under the SN hypothesis (c, d), competitive ability depends on local resource levels (e.g., successional stage). Early in succession, when resource levels are high, species with acquisitive resource traits (e.g., low LMA and WD) dominate because they achieve more rapid rates of growth and reproduction (c). As resource levels become depleted, species with conservative resource-use strategies (e.g., high LMA and WD) increase in abundance and draw resources down to levels where acquisitive species cannot persist (d). Based on the functional turnover during succession predicted by the CC tradeoff and SN mechanisms, functional diversity is expected to peak at intermediate stages of succession because species associated with early- and late-succession co-occur (e).

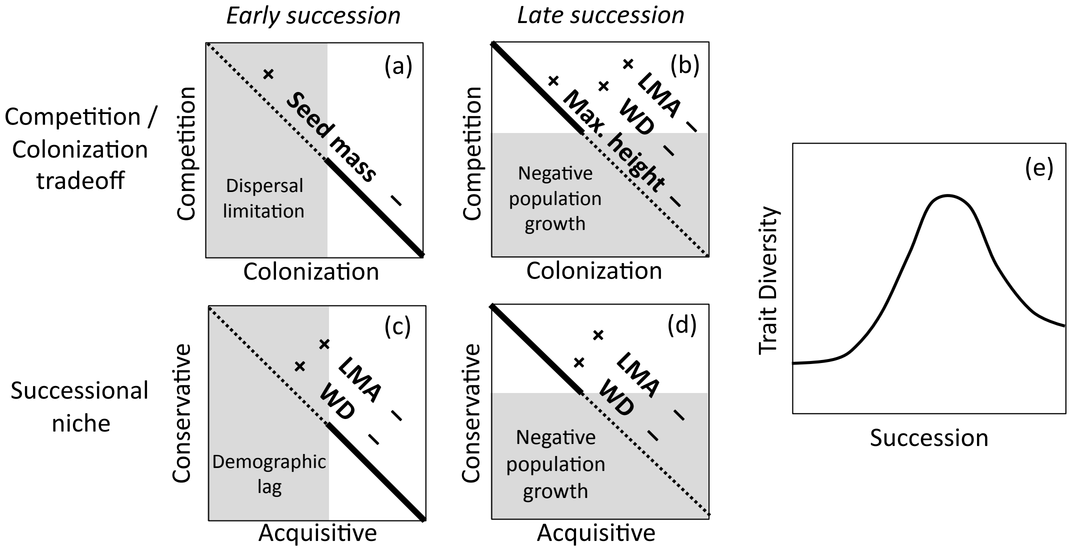


Figure 2. Community-mean trait values for saplings (1-10 cm DBH; a-d) and adults (> 10 cm DBH; e-h) in wet subtropical forests of Puerto Rico. Regression lines are shown with marginal R^2_{GLMM} (Nakagawa & Schielzeth 2012) when 95% confidence intervals of the slope did not overlap with zero. See Table 2 for full results of regressions.

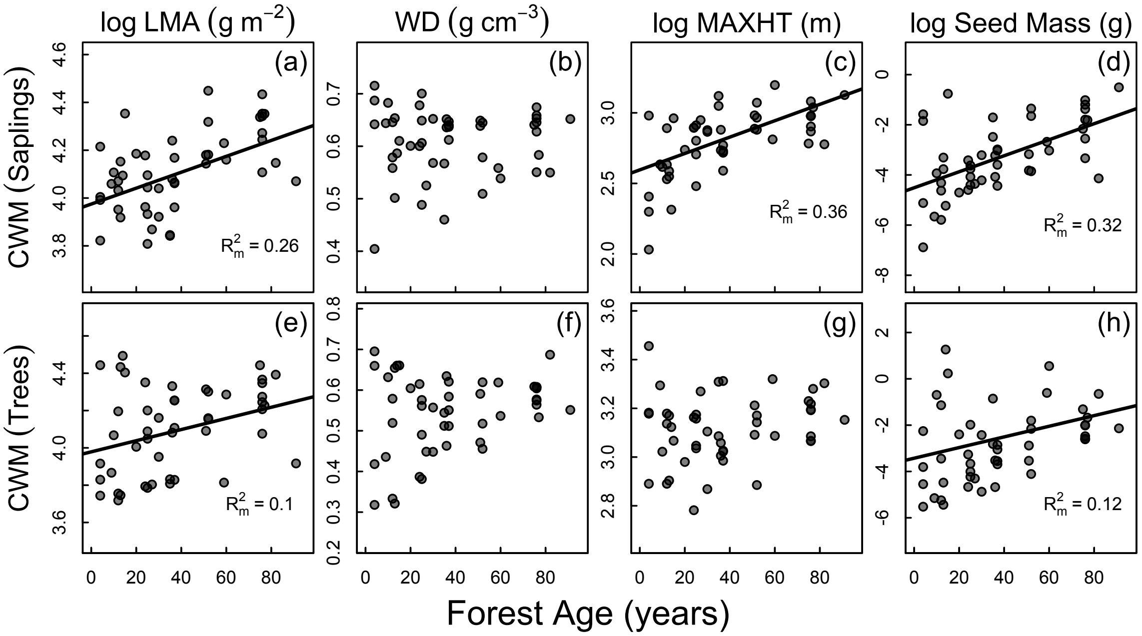
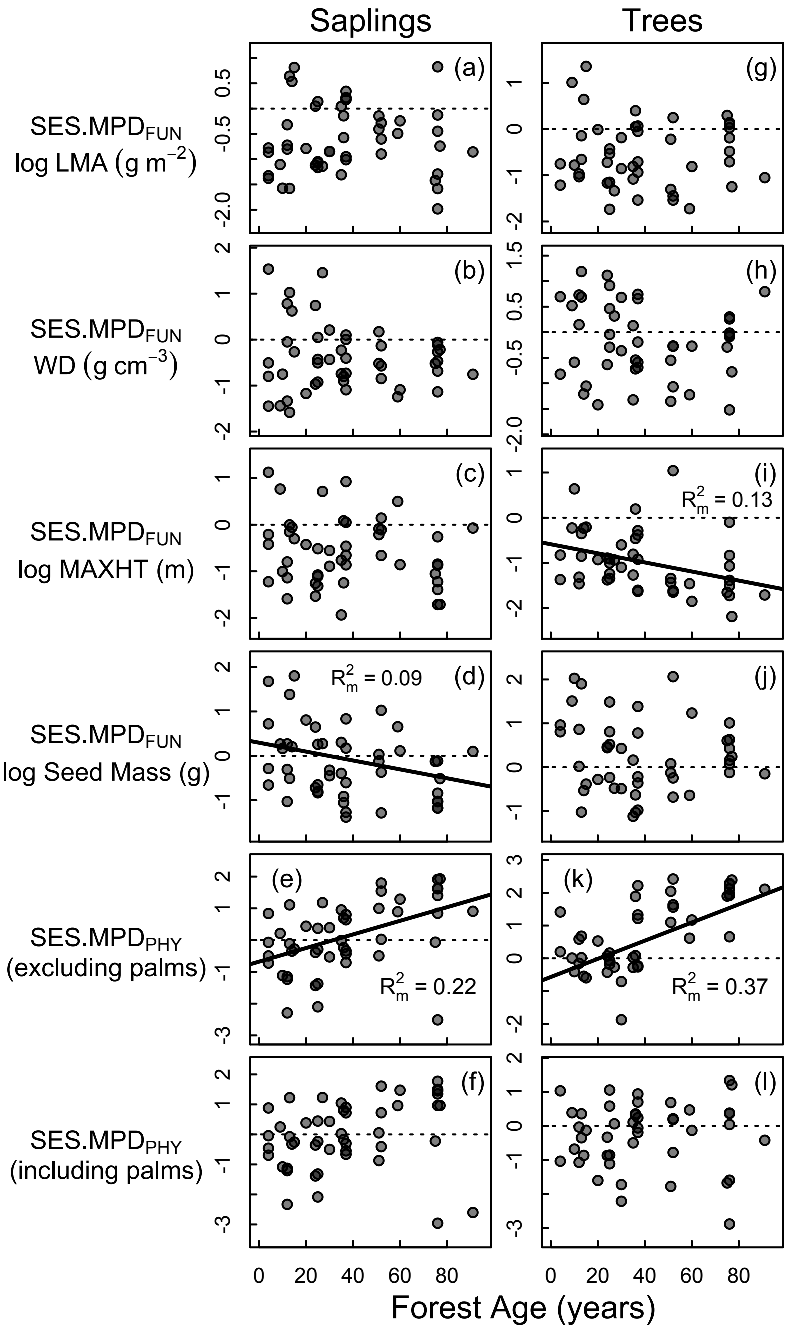


Figure 3. Mean pairwise distance in trait space (SES.MPD$_{FUN}$) and phylogenetic branch length (SES.MPD$_{PHY}$) for saplings (1-10 cm DBH; a-f) and adults (> 10 cm DBH; g-l) in wet subtropical forests of Puerto Rico. Negative values indicate functional (or phylogenetic) clustering; positive values indicate functional (or phylogenetic) evenness. Regression lines are shown with marginal R^2_{GLMM} (Nakagawa & Schielzeth 2012) when 95% confidence intervals of the slope did not overlap with zero. See Table 2 for full regression results.



CHAPTER FOUR:

Linking functional traits and ecological niche models: insights into assembly processes for Puerto Rican forests

Robert Muscarella and María Uriarte

ABSTRACT

Functional traits provide mechanistic links between abiotic gradients, species distributions, and patterns of diversity; however, the large proportion of trait variation within sites represents a major challenge for prediction in community ecology. For 173 tree species in Puerto Rico, we combined information on three functional traits, geographic distributions, and composition of forest plots across a regional precipitation gradient to test the hypothesis that habitat suitability is higher for species in locations where their trait value is closer to the local community-weighted mean (CWM). We also decomposed trait variation into within- and among-site components to evaluate their correlation structure at different scales. The relationship between habitat suitability and the difference between species traits and CWM values was negative for more species than randomly expected. Additionally, relationships between species traits and the conditions in their optimal habitats mirrored observed shifts in CWM values across the precipitation gradient. Trait-gradient analysis revealed strong pairwise trait correlations among communities, further reinforcing trait-mediated habitat filtering as a driver of species distributions and community diversity. At local scales, however, trait values were largely independent, suggesting that fine scale niche partitioning plays an important role in promoting local functional diversity. Furthermore, species with similar conditions in their optimal habitat had a wide range of trait values, raising additional questions about the mechanisms that maintain high functional diversity.

INTRODUCTION

Predicting how species and communities will respond to environmental change is both a long-standing and pressing issue in ecology. The mechanistic foundation required for predictive frameworks is, however, underdeveloped in much of community ecology. For example, correlative ecological niche models (ENMs, also called habitat suitability models or species distribution models) are a commonly used approach for understanding species responses to broad scale environmental gradients (Peterson *et al.* 2011). A major limitation with respect to the predictive capacity of ENMs, however, is the lack of physiological mechanism (Guisan & Thuiller 2005; Kearney & Porter 2009; Dormann *et al.* 2012). A functional trait perspective can help address this limitation by providing mechanistic links between species distributions, community diversity, and abiotic gradients (Schimper 1903; Mooney & Dunn 1970; McGill *et al.* 2006; Westoby & Wright 2006; Violle *et al.* 2007; Cornwell & Ackerly 2009; Sterck *et al.* 2014; Violle *et al.* 2014).

One hypothesis that underpins models of plant community assembly based on functional traits (e.g., Shipley 2010) is that if community-weighted mean trait values (i.e., plot-level trait values weighted by species abundance; CWMs) vary predictably along an abiotic gradient, then the local CWM reflects the ‘optimal’ trait value given the environmental conditions at that site. In other words, species with trait values nearest to the CWM in a particular location are expected to have relatively high fitness because they occur at relatively high abundance and contribute most strongly to the CWM (Shipley *et al.* 2011). This ‘CWM-optimality’ hypothesis is consistent with niche partitioning along resource gradients as a driver of community diversity and it emphasizes processes that constrain local functional diversity (*e.g.*, environmental filtering and competitive dominance; Grime 2006; Mayfield & Levine 2010). However, selection towards the CWM for different traits might also promote local diversity, depending on how trait axes are correlated. Although difficult to rigorously evaluate in natural systems (Adler *et al.* 2013), ENMs represent an unexploited opportunity to examine the core prediction of the ‘CWM-optimality’ hypothesis: for a given species, habitat suitability should be negatively related to the difference between its trait value and the local CWM.

In support of the ‘CWM–optimality’ hypothesis, numerous studies have demonstrated relationships between CWMs and abiotic gradients that are consistent with known physiological tradeoffs (e.g., Niinemets 2001; Wiemann & Williamson 2002; Wright & Westoby 2002; Wright *et* *al.* 2004; Chave *et al.* 2006; Niinemets & Valladares 2006; Chave *et al.* 2009; Poorter *et al.* 2009; Reich 2014). Furthermore, many of these patterns appear to hold across a range of spatial and taxonomic scales. For instance, wood density (WD) and leaf mass per area (LMA) generally decline with increasing precipitation at the community-level (Wright *et al.* 2004; Chave *et al.* 2009) and within species (Cornwell & Ackerly 2009; Violle *et al.* 2012; Niinemets 2014). Because these traits correspond to resource acquisition and utilization strategies (Poorter *et al.* 2009; Markesteijn *et al.* 2011; Reich 2014), the CWM-environment relationships can be interpreted in terms of demographic tradeoffs related to water availability (i.e., rapid growth and resource acquisition when water is abundant versus drought resistance and resource conservation in dry conditions).

Nonetheless, local communities practically always contain species spanning a wide range of values for a given trait (e.g., Westoby *et al.* 2002; Wiemann & Williamson 2002; Wright *et al.* 2004). For example, more than one-third of the total global variation in leaf mass per area (LMA) occurs within sites, as opposed to among sites that vary with respect to climatic variables (Wright *et al.* 2004). The substantial proportion of within-site variation represents a major challenge for prediction in community ecology, and for the ‘CWM-optimality’ hypothesis in particular, because it implies the combined action of mechanisms that both constrain and maintain local functional diversity (Adler *et al.* 2013; Lasky *et al.* 2014).

Fine-scale niche partitioning is one mechanism thought to contribute to the maintenance of high local diversity (Chesson 2000; Adler *et al.* 2013). A critical point here, and the subject of ongoing debate (Baraloto *et al.* 2010; Reich 2014), is the degree to which different traits represent correlated versus independent axes of life history variation. Biophysical constraints may lead to correlations between trait axes suggestive of life history tradeoffs, reducing the variety of viable strategies (e.g., Westoby *et al.* 2002; Wright *et al.* 2004; Chave *et al.* 2009; Freschet *et al.* 2010). For example, a tradeoff between hydraulic efficiency and safety appears to underlie covariation between some stem and leaf traits in tropical dry forest trees (Markesteijn *et al.* 2011). In contrast, independent variation of traits suggests higher dimensionality of resource partitioning with consequences for the maintenance of local diversity (Silvertown *et al.* 2006; Baraloto *et al.* 2010).

Ackerly and Cornwell (2007) proposed a method called trait-gradient analysis (TGA) to decompose species' trait values into alpha (within-site) and beta (among-site) components that are analogous to Whittaker's alpha and beta decomposition of species diversity (Fig. 1; Pickett & Bazzaz 1978; Ackerly & Cornwell 2007; Kooyman *et al.* 2010). For a given trait, a species' beta trait component represents its abundance-weighted mean position in communities arranged along a gradient of CWM values. A species with a high (or low) beta trait component occurs, on average, in plots with relatively high (or low) values of CWM for that trait. The alpha trait component quantifies the difference between a species' mean trait value and the mean trait value of the species with which it co-occurs. A species with a high (or low) alpha trait value has, on average, a higher (or lower) trait value compared to co-occurring species. Comparing alpha and beta components between traits can reveal variation in the strength of trait correlations at different spatial scales (Ackerly & Cornwell 2007; Kooyman *et al.* 2010; Gallagher & Leishman 2012). For example, Ackerly and Cornwell (2007) found a strong negative correlation for beta components of LMA and wood density, indicating that species that occur in plots with low CWM for LMA also tend to occur in plots with low CWM for WD. However, alpha values of LMA and WD were weakly correlated, indicating that within communities, LMA was largely independent of WD. This decoupling of traits at the local (alpha) scale implies a change in the dimensionality of niche partitioning across scales with consequences for species co-existence (Silvertown *et al.* 2006; Ackerly & Cornwell 2007; Kooyman *et al.* 2010; Gallagher & Leishman 2012).

Comparing alpha and beta trait components with estimates of habitat suitability provides another opportunity to explore the 'CWM-optimality' hypothesis in terms of the physiological mechanisms driving species geographic distributions. In particular, if CWM values covary with abiotic gradients, beta components of traits that mediate species' broad-scale distributions should be related to the abiotic conditions where the species typically occurs. In other words, the 'CWM-optimality' suggests that species should occur, on average, in plots with CWM values that correspond to the conditions of a species' optimal habitat. Correlations between alpha components and the conditions in species' optimal habitat might arise from trait-mediated fitness differences acting within communities, for example, if species that perform best in particular conditions have similar trait values (Keddy 1989).

In this study, we used TGA to decompose regional variation of three key plant functional traits (WD, LMA, and maximum height [H_{max}]) for 173 tree species across a strong precipitation gradient in Puerto Rico. We merge this analysis with ENMs to evaluate linkages between species traits and their geographic distributions. Although Ackerly & Cornwell (2007) is highly cited (270 citations, Google Scholar 10/22/2014), only a few studies have actually applied TGA and, to our knowledge, this is the first attempt to merge TGA with ENMs. We address the following specific questions:

1. *Do species' estimated habitat suitabilities decline when their trait values are more distant to the local CWM?* According to the 'CWM-optimality' hypothesis, if trait-mediated environmental filtering drives species regional distributions, we expect significant negative relationships between local habitat suitability and the difference between a species' trait value and the local CWM trait value.
2. *To what degree are different trait components correlated between traits?* We expect that trait-mediated habitat filtering acting across broad climatic gradients will result in strong pairwise correlations between beta trait components. In contrast, we expect relatively weak correlations between alpha trait components if these traits represent independent axes of life history variation among co-occurring species.
3. *How are species trait values related to precipitation in their optimal habitat?* We expect species-mean trait values to be related to the abiotic conditions of their optimal habitat in ways that mirror community-mean shifts along abiotic gradients. Specifically, we expect that species experiencing relatively high amounts of precipitation in their optimal habitat will have relatively low values of WD and LMA, and relatively high values of H_{max}. We also predict correlations between species' beta trait components and the precipitation in their optimal habitat if these trends are driven by species' geographic distributions along climatic gradients. On the other hand, we predict weak correlations between alpha trait components and

precipitation in species' optimal habitat if low pairwise trait correlation among co-occurring species increases the dimensionality of strategies enabling local species co-existence. Alternatively, correlations between alpha trait components and precipitation in species' optimal habitat could arise if trait-mediated competitive hierarchies drive species broad scale distributions.

METHODS

Study system and tree census plots

We established 12, 50 m × 50 m (0.25 ha) plots located in 4 protected forests (3 plots per forest) in Puerto Rico (Table 1). These forests occur in subtropical dry, moist, and wet forest life zones (Ewel & Whitmore 1973) and range in mean annual precipitation from ca. 800 to 2,000 mm yr^{-1} (Daly *et al.* 2003). All plots are located on limestone soils and karst topography, which constitutes about one third of Puerto Rico's land area (Lugo *et al.* 2001). Puerto Rico's karst region is noted for its high levels of biodiversity and habitat heterogeneity (Chinea 1980; Colon 1996; Lugo *et al.* 2001). Plots were located randomly in mature stands that lacked signs of former human land-use or recent natural disturbance. All plots were separated by at least 1 km except for two plots in Guánica forest, which were separated by 100 m. In each plot, we identified all woody stems (excluding lianas) ≥ 1 cm diameter at 1.3 m above the ground (DBH), measured their DBH and estimated their height to the nearest 0.5 m.

Functional traits

We collected data on 3 functional traits from multiple individuals of 173 species sampled from each forest where multiple individuals were available. *Wood density* (WD; g cm^{-3}) is related to a growth / mortality trade-off such that species with low WD have rapid growth, high hydraulic conductance, and low investment in structural materials (Chave *et al.* 2006; Chave *et al.* 2009; Kraft *et al.* 2010; Wright *et al.* 2010). Species with high WD tend to have relatively low mortality rates and are more resistant to drought-induced cavitation (Carlquist 1977; Hacke *et al.* 2001; Markesteijn *et al.* 2011). For trees 10-20 cm in diameter, we measured WD using an increment borer. For species that do not typically reach this size, we included branch WD measurements and used regressions of branch and core samples from the same individuals to correct for this difference (Appendix 1; Swenson & Enquist 2008). *Leaf dry mass per area* (LMA; g cm^{-2}) is positively related to leaf lifespan and, more generally, to variation in life history strategies ranging from fast growth and high mortality by rapid photosynthetic return on carbon investment (low LMA) to slow growth and low mortality by retaining nutrients for a longer time (high LMA) (Wright *et al.* 2004; Reich *et al.* 2007). We measured LMA on entire leaves (including petiole) of fully expanded foliage from the outer crown (sun-exposed, when possible) of mature trees. Finally, *maximum height* ($H_{\max}$; m) is associated with competitive ability in light-limited environments (Givnish 1995; Thomas & Bazzaz 1999) and generally increases with increasing precipitation and soil nutrients (Westoby *et al.* 2002). For each species, we calculated $H_{\max}$ as the larger value between (i) the tallest individual observed in our study and (ii) the maximum height reported in Little and Wadsworth (1964) or Little *et al*. (1974). We log-transformed LMA and $H_{\max}$ values to correct for skewness. In this study, we rely on species-mean trait values because the diversity and spatial extent of our study area precluded thorough measurement of intraspecific variation. This dataset represents 83% of the observed species (which account for 98% of the total basal area) recorded in the tree census plots.

Ecological niche models

We built species-specific niche models using MAXENT v 2.3 (Phillips *et al.* 2006) for all 173 species included in our trait database. We used MAXENT because it performs well with small sample sizes (Elith *et al.* 2006; Merow & Silander 2014) and a presence-background framework is appropriate for occurrence datasets based primarily on herbarium collections, such as ours. MAXENT provides gridded maps of values proportional to the expected number of occurrences per unit area for each species.

For occurrence localities, we compiled all georeferenced localities for trees in Puerto Rico from the online database GBIF (www.gbif.org). We supplemented these by manually georeferencing herbarium specimens from UPRRP, UPRM, NY, and US. Our occurrence dataset contains 17,479 observations, with 11 to 192 (mean = 54) observations for each focal species. We used the 'weighted target-group' approach to confront the spatial sampling bias inherent in this dataset (Anderson 2003). Specifically, when building an ENM for each focal species, records of all non-focal species were used as background points. Note that observations from the tree census plots were not included in the occurrence dataset. We used four climatic variables as covariates: logarithm of mean annual precipitation (MAP; mm yr^{-1}), coefficient of variation of monthly precipitation (unitless), minimum monthly temperature (°C), and mean daily temperature range (°C). These data are based on daily observations from 1963 to 1995 recorded at 108 weather stations and have a spatial resolution of 450 m^2 (Daly *et al.* 2003). We also included a categorical map of soil parent material from Bawiec (1998) as a covariate.

We conducted species-specific tuning of MAXENT settings using the R package 'ENMeval' v 0.1.1 (Muscarella *et al.* 2014a) to achieve a balance between model fit and predictive ability. Separately for each species, we calculated ΔAICc for a series of models built across a range of settings (Appendix 2) and then generated an average model prediction using AICc weights for models with ΔAICc ≤ 2 (Warren & Seifert 2011). We used these 'AICc-averaged' predictions for further analyses (*see below*). We evaluated model performance using two metrics. First, we calculated AICc-weighted average test AUC for models with ΔAICc ≤ 2. AUC values range from 0 to 1; higher values indicate better discriminatory ability (Fielding & Bell 1997). Second, we calculated the AICc-weighted average of the threshold-dependent test point omission rate based on the minimum training presence value (OR_{MTP}). Values of OR_{MTP} range from 0 to 1; progressively more overfit models have increasingly higher values of OR_{MTP} (Radosavljevic & Anderson 2014). Additional details on ENMs are provided in Appendix 2.

Statistical analyses

To evaluate the relationship between estimated habitat suitability and distance from the local CWM for each trait (*Question 1*), we calculated ΔCWM_{tip} as the absolute difference between the mean value for trait *t* of species *i* (t_i) and the community-weighted mean value of trait *t* for plot *p* (CWM_{tp}). Values of ΔCWM increase from zero as t_i becomes more different from the local CWM_t.

Separately for each species, we used OLS regression to relate ΔCWM values (separately for each trait) with habitat suitability values extracted from ENM predictions at each plot. Significantly negative slopes support the ‘CWM-optimality’ hypothesis.

We used a randomization procedure to determine (i) if the number of species with significant slopes (positive and negative) was greater than expected by chance, and (ii) if the mean value of significant slopes differed from a random expectation. For this, we shuffled habitat suitability values (within species) 999 times and, during each iteration, we (i) counted the number of significant ($p < 0.05$) OLS slopes (across species) and (ii) calculated the mean value of the significant slopes. We used the 95% quantiles of these distributions to assess the significance of the observed number of significant slopes and the observed mean value of significant slopes. This approach maintained species traits values, CWMs, and the observed distribution of habitat suitability values within species.

To determine the degree to which various trait components are correlated (*Question 2*), we performed trait-gradient analysis (TGA; Ackerly & Cornwell 2007) separately for each trait (Fig. 1). We first calculated the abundance-weighted mean trait value (CWM) for each plot *j* as:

$$\mathrm{CWM} = \frac{\sum_{\mathrm{i}=1}^{S} a_{ij} t_i}{\sum_{\mathrm{i}=1}^{S} a_{ij}}$$

where a_{ij} is the relative basal area of species *i* in plot *j* and t_{ij} is the mean trait value of species *i*. The beta trait value for species *i* is then:

$$\beta_i = \frac{\sum_{j=1}^{P} \mathrm{CWM}_j a_{ij}}{\sum_{j=1}^{P} a_{ij}}$$

where *P* is the total number of plots in the study. Beta trait values quantify a species’ abundance-weighted mean position in communities arranged along a gradient of CWM values for a particular trait. Finally, the alpha trait value for species *i* is:

$$\alpha_i = t_i - \beta_i$$

The alpha trait component quantifies the difference between a species’ mean trait value and the mean trait value of the species with which it co-occurs (Fig. 1; also see Ackerly & Cornwell 2007). We used Pearson's r to quantify pairwise correlations for CWM, species-mean, alpha, and beta trait values.

To determine how species' trait values are related to the conditions of their optimal habitat (*Question 3*), we extracted the mean annual precipitation value from grid cells with maximal values of habitat suitability for each species (restricted to limestone soils). We refer to this value as the precipitation in optimal habitat suitability (HS_{MAX}). We used OLS regression to relate species-mean traits, as well as alpha and beta trait components, to values of precipitation at HS_{MAX}. We used AIC to select between models with and without a quadratic term to allow for non-linear response. All analyses were conducted in R version 3.1.1 (R Development Core Team 2014).

RESULTS

Community-weighted mean values of all three traits varied strongly with respect to mean annual precipitation (Fig. 2). Specifically, WD and LMA declined, and H_{max} increased, with increasing precipitation (R^2= 0.75, 0.69, and 0.93, respectively). Overall, ENMs showed fairly high levels of prediction accuracy (mean AUC = 0.79) and values of OR_{MTP} were low, indicating that models were generally not overfit (mean OR_{MTP} = 0.06; Appendix 2).

Question 1) Do species' estimated habitat suitabilities decline when their trait values are more distant to the local CWM?

Histograms showing the OLS regression slopes between ΔCWM and habitat suitability are shown in Figure 3. The number (and proportion) of species with significant slopes (i.e., $p < 0.05$) was 87 (50%) for WD, 84 (49%) for LMA, and 101 (58%) for H_{max}. These values were much higher than randomly expected; the 97.5% quantile for the total number of species with significant slopes expected by random was 14 (8%) for each trait. For all traits, there were many more significantly negative slopes than randomly expected (*random 97.5% quantile* = 8; *observed values for* WD =

73, LMA = 69, and H_{max} = 81). To a much lesser degree, there were also more significantly positive slopes than randomly expected (*random 97.5% quantile* = 9; *observed values for* WD = 14, LMA = 15, and H_{max} = 20). Of the species with significant slopes, > 80% (for all three traits) were negative and the mean values of significant slopes were far below the 2.5% quantiles of the expected mean of slopes (*observed value / 2.5% quantile* for WD = -0.60 / -0.05; LMA = -0.15 / -0.02; H_{max} = -0.13 / -0.01).

Question 2) To what degree are different trait components correlated between traits?

Pairwise comparisons of CWM values for all three traits were significantly correlated. Specifically, CWM of WD was positively correlated with LMA (r = 0.97) while CWM of H_{max} was negatively correlated with both WD (r = -0.92) and LMA (r = -0.86). All three of the comparisons of species-mean trait values were significant but weaker than the plot-level trait correlations (Fig. 4). Specifically, species-mean WD was positively correlated with species-mean LMA (r = 0.54). Species-mean H_{max} was negatively correlated with both species-mean WD (r= -0.44) and LMA (r = -0.21). Beta trait components were very strongly correlated and had the same signs as the plot- and species-mean correlations (Fig. 4; r = 0.98 for WD-LMA, r = -0.95 for WD-H_{max}, and r = -0.91 for LMA-H_{max}). This indicates that species that tend to occur in plots with, for example, high CWM values of WD also tend to occur in plots with high CWM values of LMA and low CWM values of H_{max}. In contrast, alpha trait components were only weakly correlated between traits (r = 0.35 for WD-LMA, r = -0.16 for WD-H_{max}, and r = 0.16 for LMA-H_{max}). These results indicate that species with, for example, higher WD than co-occurring species (on average) are not strongly constrained in terms of their values of LMA and H_{max} relative to the species with which they typically co-occur.

Question 3) How are species trait values related to the precipitation in their optimal habitat?

As we predicted, species mean trait values were correlated with precipitation at HS_{MAX} in ways that mirrored shifts in CWM traits across the precipitation gradient. Specifically, species with relatively low values of precipitation at HS_{MAX} (i.e., species whose optimal habitat is in relatively dry sites) tended to have higher species-mean values of WD ($R^2 = 0.20$) and LMA ($R^2 = 0.23$), and lower values of $H_{\max}$ ($R^2 = 0.22$; Fig. 5) than species with higher values of precipitation at HS_{MAX}. Beta values were similarly but even more strongly correlated between traits. Wood density and LMA were negatively correlated with precipitation at HS_{MAX} ($R^2 = 0.67$ and 0.62, respectively), and beta trait values of $H_{\max}$ were positively correlated with precipitation at HS_{MAX} ($R^2 = 0.75$). These results indicate, for example, that species with optimal habitat suitability in dry areas also tend to occur in plots with relatively high values of WD and LMA, and low values of $H_{\max}$. Alpha trait values for all three traits were very weakly correlated with precipitation at HS_{MAX} ($R^2 \leq 0.04$ in all cases; Fig. 5), indicating that the level of mean annual precipitation in sites where species have their highest estimated habitat suitability is not strongly related to species' trait values relative to the species with which they co-occur.

DISCUSSION

Overall, our results (i) provide mixed evidence for the 'CWM-optimality' hypothesis, and (ii) reveal both broad scale habitat filtering and fine scale niche partitioning as important drivers of functional diversity in the limestone forests of Puerto Rico. As predicted by the 'CWM-optimality' hypothesis, habitat suitability was negatively related to ΔCWM values (i.e., the absolute difference between the species trait value and the local CWM) for many more species than randomly expected, for all three traits. However, we also found more species with positive correlations between habitat suitability and ΔCWM values than expected by random chance. Trait-gradient analysis revealed strong pairwise trait correlations among communities, reflecting trait-mediated habitat filtering. Furthermore, significant relationships between species-mean trait values and the conditions of their optimal habitat mirrored shifts in CWM values across the precipitation gradient, reinforcing the role of habitat filtering in governing broad patterns of species occurrence and community composition. At the same time, species trait values were largely independent when evaluated relative to co-occurring species, suggesting that fine scale niche partitioning plays an important role in promoting local functional diversity. We also found that species with similar conditions in their optimal habitat had a wide range of trait values, raising additional questions about the particular mechanisms that maintain high functional diversity in this system.

The 'CWM-optimality' hypothesis

A prerequisite for the 'CWM-optimality' hypothesis is that CWM trait values vary with respect to environmental gradients. Indeed, we observed clear shifts in CWM trait values along the precipitation gradient that are consistent with physiological tradeoffs mediated by water availability (Givnish 1995; Poorter *et al.* 2009; Poorter *et al.* 2010; Markesteijn *et al.* 2011). These trends imply selective mechanisms (e.g., environmental filtering, competitive dominance hierarchies) that drive local functional diversity towards the CWM (*see also* Chapter 2). Specifically, drought resistance conveyed by high WD and LMA (Hacke *et al.* 2001; Poorter *et al.* 2009; Markesteijn *et al.* 2011) appears to be particularly important for dry forest trees, especially on limestone soils with low water holding capacity (Camacho 2005). Short stature in these forests is likely the result of hydraulic limitation (Ryan & Yoder 1997; Ryan *et al.* 2006). In wetter sites, a decline in hydraulic limitation is coupled with an increased benefit in terms of carbon gain for rapid growth and tall stature (Givnish 1995; Falster & Westoby 2003).

Also consistent with the 'CWM-optimality' hypothesis, the correlation between habitat suitability and ΔCWM was negative for the overwhelming majority of species that had a significant relationship. These results provide added support to the physiological constraints (mentioned above) that appear to underlie broad scale habitat filtering in this system. Interestingly, all three traits exhibited strikingly similar patterns in this analysis despite moderate correlations between the species-mean values.

Contrary to the 'CWM-optimality' hypothesis, more species than expected also had significantly positive relationships between habitat suitability and ΔCWM, reflecting the potential success of life history strategies that deviate from the CWM. For example, species with low WD that are particularly common in the dry forest (e.g. *Bursera simarubra* and *Pisonia albida*) appear to cope with limited water availability through mechanisms such as rapid water uptake, water storage in the stem, and a deciduous habit to reduce transpirational water loss (Borchert 1994). While we were not able to examine them here, belowground traits are also undoubtedly important for describing the variety of hydrological niches (Silvertown *et al.* 2014). It is possible that strong selection towards the CWM value of belowground traits in dry areas would lend additional support to the 'CWM-optimality' hypothesis.

Additional evidence in support of the 'CWM-optimality' hypothesis comes from our observations that both species-mean trait values and beta trait components were correlated with values of precipitation at HS_{MAX} (i.e., the level of precipitation in a species' optimal habitat). Specifically, species with optimal habitat suitability in dry locations tended to have higher WD, LMA, and lower H_{max} than species with optimal habitat suitability in wet locations. On one hand, these results support the role of broad scale habitat filtering as an important mechanism in determining species geographic distributions, constraining local functional diversity, and mediating shifts of CWM values across regional abiotic gradients. On the other hand, the fairly low proportion of explained variance in these relationships (< 23% of the total trait variation) indicates that a substantial amount of functional variation among species for which similar abiotic conditions represent optimal habitat. One likely explanation for these results is that while highly aggregated climatic data (e.g., mean annual precipitation) can capture coarse trends in functional composition of forest plots, fine-scale habitat heterogeneity (and perhaps temporal niche partitioning, i.e., temporal storage effects; Chesson 2000) is critical for the maintenance of local diversity. Regions of karst topography, including our study system, are noted for particularly high levels of microhabitat heterogeneity (Lugo *et al.* 2001). Environmental conditions including soil depth, chemistry, water-holding capacity, and exposure to sun and wind vary dramatically at fine scales, thus providing the potential for fine scale niche partitioning as an important mechanism for the maintenance of local functional diversity.

Trait correlations and mechanisms that promote local diversity

The decomposition of species trait values into alpha and beta components via TGA revealed variation in the dimensionality of niche partitioning at different spatial scales. We found strong correlations for all pairwise comparisons of beta trait components that provide additional support for trait-mediated habitat filtering across the regional environmental gradients in this system. In contrast, alpha trait components were relatively weakly correlated, suggesting that the three traits considered here represent independent axes of life history variation at the local scale. These results are highly relevant for the ongoing debate about the degree to which different functional trait axes are correlated (e.g., Baraloto *et al.* 2010; Fortunel *et al.* 2012; Patiño *et al.* 2012; Reich 2014). The few existing studies that have conducted TGA have reported similar results to those found here. For instance, in woody plant communities of coastal California, Ackerly & Cornwell (2007) found strong pairwise correlations for beta components of SLA, leaf size, H_{max}, and WD. Alpha trait components, in contrast, were weakly or not significantly correlated. In tree communities of eastern Australia, Kooyman *et al*. (2010) found relatively strong correlations between beta trait components compared to alpha trait components for WD, seed size, leaf size, and H_{max}. While additional studies are required, our results contribute to the existing evidence for scale-dependency in the nature of trait correlations. If general, identifying such a pattern may help elucidate physiological mechanisms underlying the maintenance of local diversity. Ultimately, however, studies that incorporate experimental manipulations (where possible) or dynamic demographic data are necessary to identify the particular mechanisms underlying trait-mediated niche partitioning (Chesson 2000; HilleRisLambers *et al.* 2012; Adler *et al.* 2013).

Conclusions and future directions

A key challenge for predictive frameworks in community ecology is to reconcile two contrasting patterns: while abiotic gradients appear to exert strong selective pressure on species and community-level mean trait values, individual communities contain wide ranges of values for any particular trait. The pervasiveness of these two patterns implies the simultaneous action of mechanisms that constrain and promote local diversity. The infusion of a functional perspective into community ecology is providing valuable insight into both the physiological mechanisms underpinning species broad scale geographic distributions and patterns of local diversity (Adler *et al.* 2014; Sterck *et al.* 2014). However, continuing to advance our understanding of the factors governing species range dynamics will require a stronger integration of local processes with broad scale geographic distributions (Wisz *et al.* 2012; Svenning *et al.* 2014). For example, while correlative ENMs have been fairly successful in predicting patterns of species occurrence, these estimates of habitat suitability may not necessarily correspond to demographic performance (McGill 2012; Thuiller *et al.* 2014) or local abundance (VanDerWal *et al.* 2009; Fox 2012). In part, this is because local scale processes (*e.g.*, competition, predation) can preclude a species from colonizing a site or otherwise decouple demographic rates from patterns of occurrence (Sagarin & Gaines 2002; Canham *et al.* 2006; Yackulic *et al.* 2011; Clark *et al.* 2013). Further development of a functional perspective in community ecology can help to fill these gaps by providing mechanistic links between abiotic gradients and diversity patterns.

ACKNOWLEDGMENTS

We thank the Departmento de Recursos Naturales y Ambientales (DRNA) of Puerto Rico as well as individual forest officers for permission to conduct this research under DRNA permit #2011-IC-046. Marcos Rodriguez provided tireless assistance with all aspects of fieldwork and keen insight to the natural history of Puerto Rican trees. Fabiola Areces, Frank Axelrod, and Papo Vives assisted with species identification. We thank Fabiola Areces, Franklin Axelrod, Danilo Chinea, and Jeanine Vélez at the University of Puerto Rico for providing access to digitized herbarium records. Rebecca Panko and Víctor José Vega López georeferenced herbarium specimens. Silvia Bibbo and Aaron Hogan helped collect functional trait data. This work was supported by NSF DEB 1050957 to MU, NSF DEB 1311367 to MU and RM, and grants from the Earth Institute and the Institute of Latin American studies at Columbia University.

TABLES AND FIGURES

Table 1. Characteristics of the plots used in this study. Precipitation metrics are derived from Daly *et al*. (2003). Rarefied species richness is based on the minimum number of individual stems sampled in a single plot (n = 806).

Site	**Forest name**	**Mean annual precipitation (cm yr^{-1})**	**Minimum monthly precipitation (cm mo^{-1})**	**CV of monthly precipitation (unitless)**	**Basal area (m ha^{-1})**	**Stem density (stems 0.25 ha^{-1})**	**Total species richness**	**Rarefied species richness**
CAM1	Cambalache	152	83	0.3	27.9	1612	55	45.9
CAM2	Cambalache	146	81	0.29	24.3	1569	41	38.7
CAM3	Cambalache	166	86	0.26	23.4	2204	53	45
GUA1	Guajataca	204	78	0.36	27.8	1705	62	53.3
GUA2	Guajataca	198	80	0.35	24.1	945	43	42.3
GUA3	Guajataca	190	70	0.38	34.4	1417	78	63.2
GUN1	Guánica	104	23	0.59	15.3	2578	43	35.8
GUN2	Guánica	93	24	0.59	20.2	3777	49	42
GUN3	Guánica	99	24	0.58	15.9	4102	54	41.5
RIO1	Rio Abajo	205	84	0.4	19.2	1750	81	66.3
RIO2	Rio Abajo	215	70	0.44	35.8	806	45	45
RIO3	Rio Abajo	201	76	0.41	25	1530	64	57.1

Figure 1. A plot of species-mean trait values (t_i) vs. basal area-weighted community-mean trait value (CWM_j) for wood density (WD; g cm^{-3}) in 12 forest plots in Puerto Rico. Point sizes are proportional to species relative basal area and the dashed line is the 1:1 line. Vertical columns of points represent species that co-occur in a plot with a CWM value of WD shown on the x-axis. Individual species are spread along the y-axis according to their mean trait value (t_i). For each species, the weighted-mean of CWM values for plots it occupies is its β trait value. A species' α value is given by the difference between the species' mean trait value (t_i) and its β value. We highlight two species for illustrative purposes: *Thouinia striata* (triangles) and *Bursera simaruba* (diamonds). *T. striata* has a relatively high mean trait value ($t_1 \approx 0.9$). The abundance-weighted mean of CWM_j values in plots where it occurs is its beta value ($\beta_1 \approx 0.7$). The difference between these two values is the species' alpha value ($\alpha_1 \approx 0.2$), indicating that *T. striata* has higher WD, on average, than the species with which it co-occurs. *B. simaruba*, in contrast, has a relatively low mean trait value ($t_2 = 0.4$) but its beta value is higher ($\beta_2 = 0.75$) because it tends to occur in plots with higher CWM values for WD than *T. striata*. The alpha value of *B. simaruba* is negative ($\alpha_2 = -0.35$) because it has a lower species-mean value of WD than the species with which it typically co-occurs.

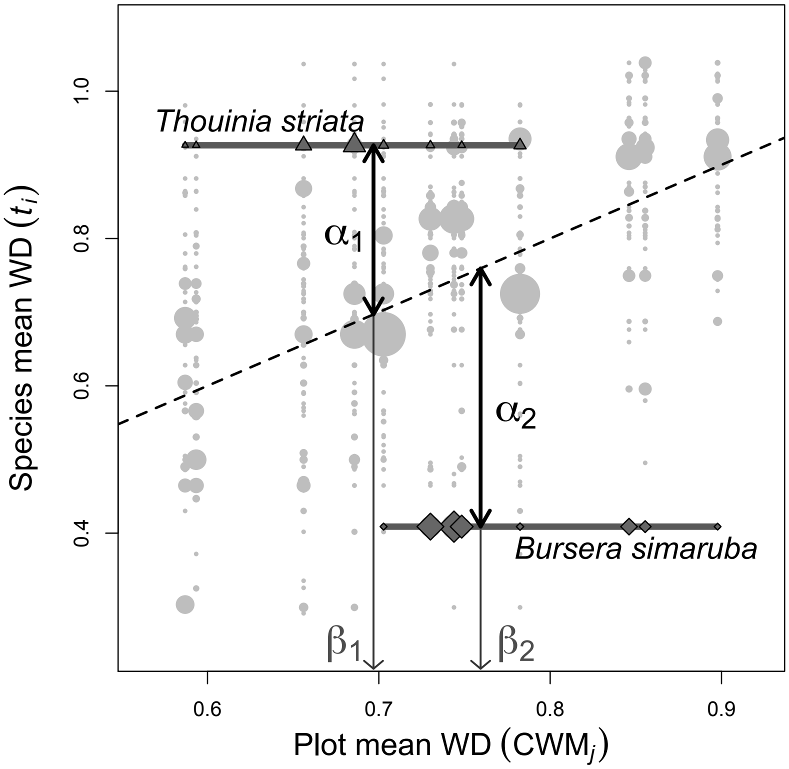


Figure 2. Community-weighted mean traits (CWMs) based on relative basal-area for wood density (a), LMA (b), and H_{max} (c) versus mean annual precipitation for 12 forest plots in limestone soils of Puerto Rico. Large circles represent CWM values, smaller points in the background represent values of species that occur in each plot. Note that LMA and H_{max} are plotted with log scales.

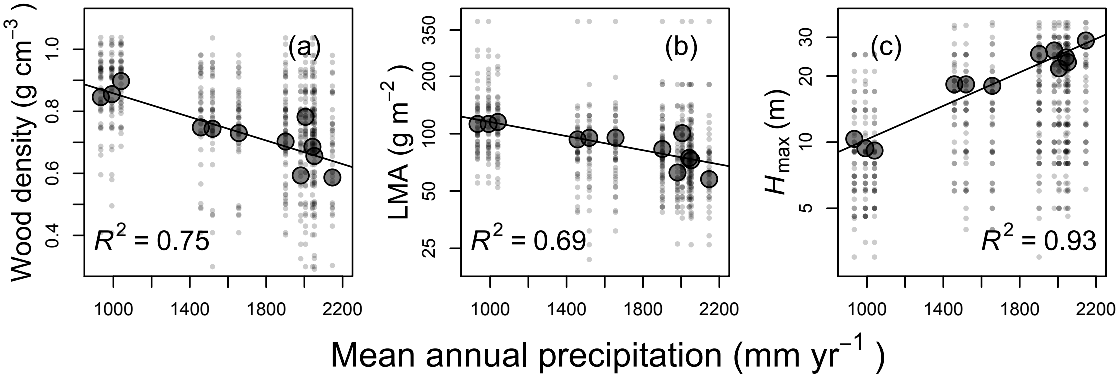


Figure 3. Histograms of OLS regression slopes for the relationship between ΔCWM_{ip} (the absolute difference between the trait value of species *i* and the CWM value in plot *p*) and habitat suitability as estimated from the ENMs. All slopes are shown with white bars and statistically significant slopes ($p < 0.05$) are highlighted in grey.

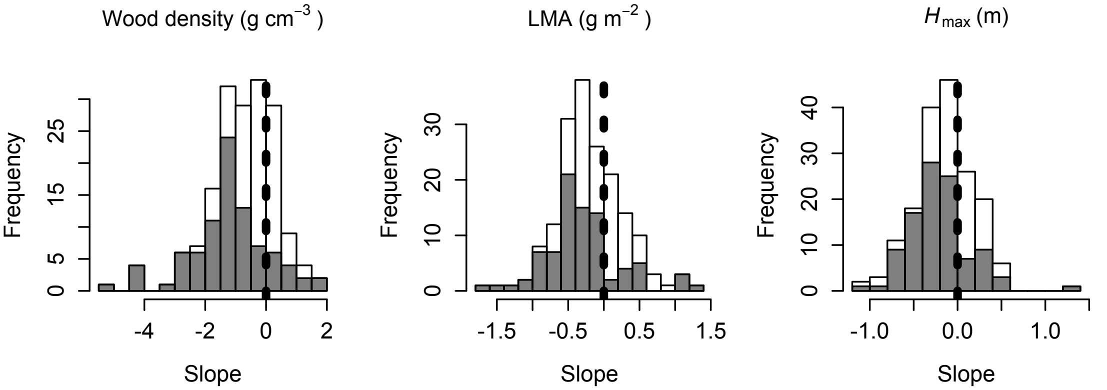


Figure 4. Plots of (a) CWM values (CWM_j), (b) species-mean values (t_i), (c) beta trait components (β_i), and (d) alpha trait components (α_i) for all pairwise combinations of wood density (WD; g cm^{-3}), leaf mass per dry area (LMA; log g m^{-2}), and maximum height (H_{max}; log m). Pearson correlation coefficients are shown with an asterisk when $p<0.05$. Positive (negative) values in (d) correspond to species that have higher (lower) trait values, on average, than the species with which they co-occur.

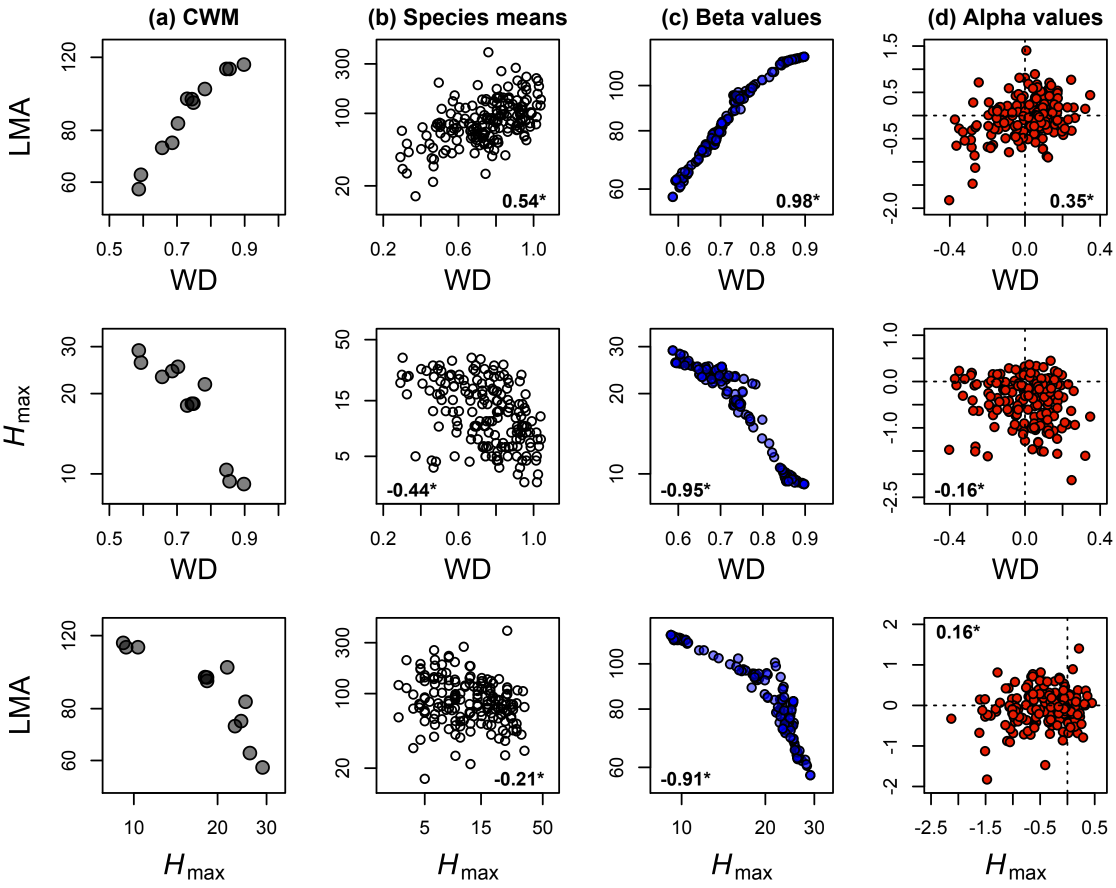
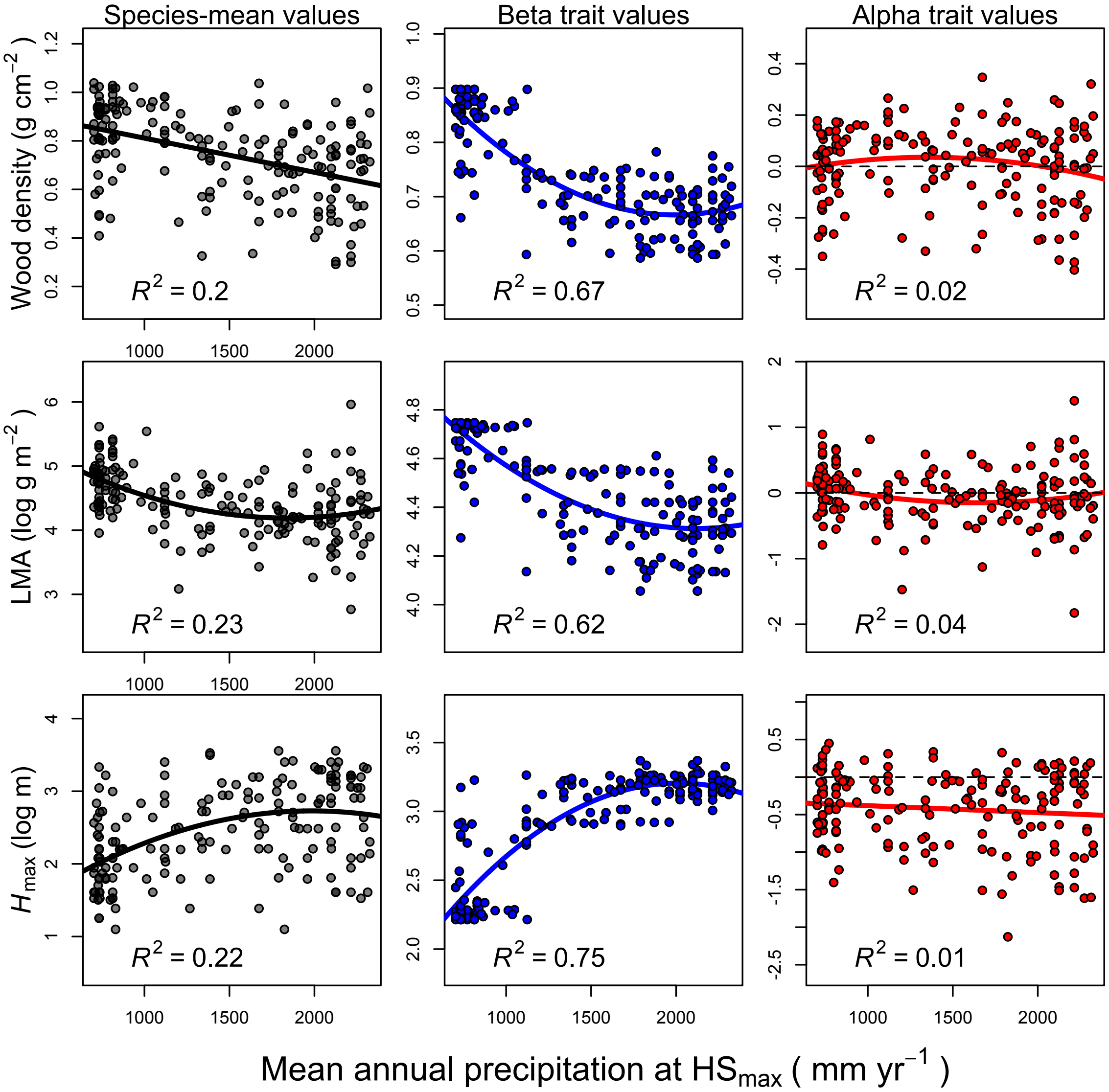


Figure 5. Relationships between precipitation in optimal habitat (HS_{MAX}, estimated from ENMs) and species-mean trait values (left column, black points), beta trait components (middle column, blue points), and alpha trait components (right column, red points).



CONCLUSION

By integrating data on species functional characteristics and evolutionary relationships, this dissertation provides new insight about how abiotic gradients mediate community assembly processes in Puerto Rican forests. An overarching theme is the relationship between abiotic gradients and processes that promote (and maintain) local diversity versus processes that reduce (and constrain) local diversity. I have argued that improving our understanding of the interactions among these processes requires a framework that embraces both regional and local perspectives on community assembly.

My hope is that this dissertation stimulates additional research towards successfully merging regional and local perspectives. In particular, the island-wide phylogeny presented in Chapter 1 opens the door to many opportunities for placing contemporary patterns of diversity, species interactions, and species geographic distributions in a historical context. Chapters 2 and 3 report compelling diversity patterns along spatial and temporal resource gradients, respectively. These chapters invite continued investigations to determine, for example, the physiological mechanisms underlying the observed compositional shifts. Furthermore, the generality of these findings for other systems remains to be seen. Chapter 4 provides robust models of species' ecological niches that can be used to test additional hypotheses regarding the links between species geographic distributions, their physiology, and their interactions with other species. In sum, the separate parts of this dissertation contribute numerous elements to the body of existing knowledge on Puerto Rican plant communities. This information can be leveraged to conduct research at a regional scale.

Together, the studies included in this dissertation suggest that understanding the degree to which species' responses to local environmental heterogeneity correspond to their responses to regional environmental gradients will be an important step towards developing predictive theory in community ecology. Ongoing and future research efforts will benefit by incorporating individual-level demographic information with both fine- and coarse-scale abiotic gradients. In this way, mechanisms that act on the scale of the individual (e.g., biotic interactions, physiological constraints) can be compared with distributions of species across broad scales. Additionally, when possible, experimental manipulations may shed additional light on the mechanisms underlying observed patterns of diversity in this system. Although these are often not feasible for diverse communities of long-lived organisms (e.g., tropical forests), they may be useful to explore regional variation in the fine-scale mechanisms governing population and community dynamics at early life stages (i.e., seeds and seedlings), which are major bottlenecks in plant demography. Ultimately, improving our understanding the convergence and diverge of species' responses to environmental gradients at different scales will give insight to the effects of environmental change on diversity patterns across a range of spatial and temporal scales.

Obviously, achieving a predictive framework for community ecology is highly ambitious, at best. Nonetheless, I hope the results of this dissertation contribute to the existing literature aimed at such a goal by re-examining existing ideas from a regional perspective. From a more immediate standpoint, I hope this dissertation provides useful information for the conservation and protection of Puerto Rico's biodiversity. Puerto Rico is truly poised to set a precedent for successful plant conservation in the tropics. To start, the amount of existing knowledge on the flora of Puerto Rico dwarfs that of most tropical locations. This is complemented by an abundance of information on land-use history and biophysical characteristics. Moreover, there are established legal mechanisms (e.g., the Endangered Species Act), a vibrant conservation community, and a research infrastructure and legacy that provide the essential ingredients for successful plant conservation. At a recent meeting of the Puerto Rico and US Virgin Island Plant Conservation Task Force, one speaker noted, "If we can't make it happen in Puerto Rico, it can't happen anywhere." I hope this dissertation contributes something towards the realization of this potential.

REFERENCES

Acevedo-Rodríguez, P. & Strong, M.T. (2011). Flora of the West Indies: catalogue of the seed plants of the West Indies.

Ackerly, D.D. & Cornwell, W. (2007). A trait-based approach to community assembly: partitioning of species trait values into within- and among-community components. *Ecology Letters*, 10, 135-145.

Adler, P.B., Fajardo, A., Kleinhesselink, A.R. & Kraft, N.J.B. (2013). Trait-based tests of coexistence mechanisms. *Ecology Letters*, 16, 1294-1306.

Adler, P.B., Salguero-Gómez, R., Compagnoni, A., Hsu, J.S., Ray-Mukherjee, J., Mbeau-Ache, C. *et al.* (2014). Functional traits explain variation in plant life history strategies. *Proceedings of the National Academy of Sciences*, 111, 740-745.

Aide, T.M., Zimmerman, J.K., Pascarella, J.B., Rivera, L. & Marcano-Vega, H. (2000). Forest regeneration in a chronosequence of tropical abandoned pastures: implications for restoration ecology. *Restoration Ecology*, 8, 328-338.

Allan, R.P. & Soden, B.J. (2008). Atmospheric Warming and the Amplification of Precipitation Extremes. *Science*, 321, 1481-1484

Anderson, R.P. (2003). Real vs. artefactual absences in species distributions: tests for *Oryzomys albigularis* (Rodentia: Muridae) in Venezuela. *Journal of Biogeography*, 30, 591-605.

APG III (2009). An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG III. *Botanical Journal of the Linnean Society*, 161, 105-121.

Axelrod, F.S. (2011). *A systematic vademecum to the vascular plants of Puerto Rico*. University of Puerto Rico, Río Piedras, PR.

Baraloto, C., Paine, C.E.T., Poorter, L., Beauchene, J., Bonal, D., Domenach, A.-M. *et al.* (2010). Decoupled leaf and stem economics in rain forest trees. *Ecology Letters*, 13, 1338-1347.

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

Bates, D., Maechler, M. & Bolker, B. (2013). lme4: Linear mixed-effects models using S4 classes. R package version 0.999999-2. http://CRAN.R-project.org/package=lme4.

Bawiec, W.J. (1998). Geology, geochemistry, geophysics, mineral occurrences, and mineral resource assessment for the commonwealth of Puerto Rico: Open-File Report 98-38, Geological Survey (U.S.).

Bazzaz, F.A. (1979). The physiological ecology of plant succession. *Annual Review of Ecology & Systematics*, 10.

Bazzaz, F.A. & Pickett, S.T.A. (1980). Physiological ecology of tropical succession: a comparative review. *Annual Review of Ecology & Systematics*, 11, 287-310.

Beaulieu, J.M., Ree, R.H., Cavender-Bares, J., Weiblen, G.D. & Donoghue, M.J. (2012). Synthesizing phylogenetic knowledge for ecological research. *Ecology*, 93, S4-S13.

Beinroth, F.H., Engel, R.J., Lugo, J.L., Santiago, C.L., Ríos, S. & Brannon, G.R. (2003). Updated taxonomic classificaiton of the soils of Puerto Rico, 2002. *Bulletin 303, University of Puerto Rico, Agricultural Experiment Station, Río Piedras, PR.*, 1-73.

Benson, D.A., Karsch-Mizrachi, I., Lipman, D.J., Ostell, J. & Sayers, E.W. (2010). GenBank. *Nucleic Acids Research*, 38, D46-D51.

Bernard-Verdier, M., Navas, M.-L., Vellend, M., Violle, C., Fayolle, A. & Garnier, E. (2012). Community assembly along a soil depth gradient: contrasting patterns of plant trait convergence and divergence in a Mediterranean rangeland. *Journal of Ecology*, 100, 1422-1433.

Bhaskar, R., Dawson, T.E. & Balvanera, P. (2014). Community assembly and functional diversity along succession post-management. *Functional Ecology*.

Boose, E.R., Serrano, M.I. & Foster, D.R. (2004). Landscape and regional impacts of hurricanes in Puerto Rico. *Ecol. Monogr.*, 74.

Borchert, R. (1994). Soil and stem water storage determine phenology and distribution of tropical dry forest trees. *Ecology*, 75, 1437-1449.

Britton, T., Anderson, C.L., Jacquet, D., Lundqvist, S. & Bremer, K. (2007). Estimating divergence times in large phylogenetic trees. *Systematic Biology*, 56, 741-752.

Brokaw, N., Fraver, S., Grear, J.S., Thompson, J., Zimmerman, J.K., Waide, R.B. *et al.* (2004). Disturbance and canopy structure in two tropical forests. In: *Tropical forest diversity and dynamism: results from a long-term tropical forest network* (eds. Losos, E & Leigh, JEG). The University of Chicago Press Chicago, IL, pp. 177-194.

Brunbjerg, A.K., Cavender-Bares, J., Eiserhardt, W.L., Ejrnæs, R., Aarssen, L.W., Buckley, H.L. *et al.* (2014). Multi-scale phylogenetic structure in coastal dune plant communities across the globe. *Journal of Plant Ecology*, 7, 101-114.

Camacho, J.L.L. (2005). The soil climate regimes of Puerto Rico - Reassessment and implications. In: *Department of Agronomy and Soils*. University of Puerto Rico Mayagüez, Puerto Rico, p. 148.

Canham, C.D., Papaik, M.J., Uriarte, M., McWilliams, W.H., Jenkins, J.C. & Twery, M.J. (2006). Neighborhood analyses of canopy tree competition along environmental gradients in New England forests. *Ecological Applications*, 16, 540-554.

Canham, C.D., Thompson, J., Zimmerman, J.K. & Uriarte, M. (2010). Variation in susceptibility to hurricane damage as a function of storm intensity in Puerto Rican tree species. *Biotropica*, 42, 87-94.

Carlquist, S. (1977). Ecological factors in wood evolution: a floristic approach. *American Journal of Botany*, 64, 887-896.

Castilleja, G. (1991). Seed germination and early establishment in a sub-tropical dry forest. In: *the graduate school*. Yale University New Haven, CT.

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

Cavender-Bares, J., Ackerly, D.D. & H., K.K. (2012). Integrating ecology and phylogenetics: the footprint of history in modern-day communities. *Ecology*, 93, S1–S3.

Cavender-Bares, J., Keen, A. & Miles, B. (2006). Phylogenetic structure of Floridian plant communities depends on taxonomic and spatial scale. *Ecology*, 87, S109-S122-S109-S122.

Cavender-Bares, J., Kitajima, K. & Bazzaz, F.A. (2004b). Multiple trait associations in relation to habitat differentiation among 17 Floridian oak species. *Ecol. Monogr.*, 74, 635-662.

Cavender-Bares, J., Kozak, K.H., Fine, P.V.A. & Kembel, S.W. (2009). The merging of community ecology and phylogenetic biology. *Ecology Letters*, 12, 693-715.

CBOL Plant Working Group (2009). A DNA barcode for land plants. *Proceedings of the National Academy of Sciences*, 106, 12794-12797.

Chave, J., Andalo, C., Brown, S., Cairns, M.A., Chambers, J.Q., Eamus, D. *et al.* (2005). Tree allometry and improved estimation of carbon stocks and balance in tropical forests. *Oecologia*, 145, 87-99.

Chave, J., Coomes, D., Jansen, S., Lewis, S.L., Swenson, N.G. & Zanne, A.E. (2009). Towards a worldwide wood economics spectrum. *Ecology Letters*, 12, 351-366.

Chave, J.m., Muller-Landau, H.C., Baker, T.R., Easdale, T.A., Steege, H.t. & Webb, C.O. (2006). Regional and phylogenetic variation of wood density across 2456 neotropical tree species. *Ecological Applications*, 16, 2356-2367.

Chazdon, R.L. (2003). Tropical forest recovery: legacies of human impact and natural disturbances. *Perspectives in Plant Ecology, Evolution and Systematics*, 6, 51-71.

Chazdon, R.L. (2014). *Second growth: the promise of tropical forest regeneration in an age of deforestation*. University of Chicago Press, Chicago, IL.

Chazdon, R.L., Finegan, B., Capers, R.S., Salgado-Negret, B., Casanoves, F., Boukili, V. *et al.* (2010). Composition and dynamics of functional groups of trees during tropical forest succession in northeastern Costa Rica. *Biotropica*, 42, 31-40.

Chesson, P. (2000). Mechanisms of maintenance of species diversity. *Annual Review of Ecology and Systematics*, 31, 343-366.

Chinea, D. (1980). The forest vegetation of the limestone hills of northern Puerto Rico. Cornell University Ithaca, NY.

Clark, J.S., Gelfand, A.E., Woodall, C.W. & Zhu, K. (2013). More than the sum of the parts: Forest climate response from joint species distribution models. *Ecological Applications*, 24, 990-999.

Colon, J.C.F. (1996). Phytogeographical trends, centers of high species richness and endemism, and the question of extinctions in the native flora of Puerto Rico. *Annals of the New York Academy of Sciences*, 776, 89-102.

Colwell, R.K. & Winkler, D.W. (1984). A null model for null models in biogeography. In: *Ecological communities: conceptual issues and the evidence* (eds. Strong, DRJ, Simberloff, D, Abele, LG & Thistle, AB). Princeton University Press Princeton, NJ, pp. 344-359.

Coomes, D.A. & Grubb, P.J. (2003). Colonization, tolerance, competition and seed-size variation within functional groups. *Trends in Ecology & Evolution*, 18, 283-291.

Cornelissen, J.H.C., Lavorel, S., Garnier, E., Díaz, S., Buchmann, N., Gurvich, D.E. *et al.* (2003). A handbook of protocols for standardised and easy measurement of plant functional traits worldwide. *Australian Journal of Botany*, 51, 335-380.

Cornwell, W.K. & Ackerly, D.D. (2009). Community assembly and shifts in plant trait distributions across an environmental gradient in coastal California. *Ecol. Monogr.*, 79, 109-126.

Coyle, J.R., Halliday, F.W., Lopez, B.E., Palmquist, K.A., Wilfahrt, P.A. & Hurlbert, A.H. (2014). Using trait and phylogenetic diversity to evaluate the generality of the stress-dominance hypothesis in eastern North American tree communities. *Ecography*, no-no.

Crozier, R.H. (1997). Preserving the information content of species: genetic diversity, phylogeny, and conservation worth. *Annual Review of Ecology and Systematics*, 28, 243-268.

Dalling, J.W. & Brown, T.A. (2009). Long-Term Persistence of Pioneer Species in Tropical Rain Forest Soil Seed Banks. *American Naturalist*, 173, 531-535.

Daly, C., Helmer, E.H. & Quiñones, M. (2003). Mapping the climate of Puerto Rico, Vieques and Culebra. *International Journal of Climatology*, 23, 1359-1381.

Darriba, D., Taboada, G.L., Doallo, R. & Posada, D. (2012). jModelTest 2: more models, new heuristics and parallel computing. *Nature Methods*, 9, 772.

Davies, J.T., Kraft, N.J.B., Salamin, N. & Wolkovich, E.M. (2012). Incompletely resolved phylogenetic trees inflate estimates of phylogenetic conservatism. *Ecology*, 93, 242-247.

de Bello, F., Lavorel, S., Lavergne, S., Albert, C.H., Boulangeat, I., Mazel, F. *et al.* (2013). Hierarchical effects of environmental filters on the functional structure of plant communities: a case study in the French Alps. *Ecography*, 36, 393-402.

de Bello, F., Price, J.N., Münkemüller, T., Liira, J., Zobel, M., Thuiller, W. *et al.* (2012). Functional species pool framework to test for biotic effects on community assembly. *Ecology*, 93, 2263-2273.

de Bello, F., Thuiller, W., Leps, J., Choler, P., Clément, J.-C., Petr, M. *et al.* (2009). Partitioning of functional diversity reveals the scale and extent of trait convergence and divergence. *Journal of Vegetation Science*, 20, 475-486.

Devictor, V., Mouillot, D., Meynard, D., Jiguet, F., Thuiller, W. & Mouquet, N. (2010). Spatial mismatch and congruence between taxonomic, phylogenetic and functional diversity: the need for integrative conservation strategies in a changing world. *Ecology Letters*, 13, 1030-1040.

Díaz, S. & Cabido, M. (2001). Vive la différence: plant functional diversity matters to ecosystem processes. *Trends in Ecology & Evolution*, 16, 646-655.

Ding, Y., Zang, R., Letcher, S.G., Liu, S. & He, F. (2012). Disturbance regime changes the trait distribution, phylogenetic structure and community assembly of tropical rain forests. *Oikos*, 121, 1263-1270.

Dormann, C.F., Schymanski, S.J., Cabral, J., Chuine, I., H., G.C., Hartig, F. *et al.* (2012). Correlation and process in species distribution models: bridging a dichotomy. *Journal of Biogeography*, 39, 2119–2131.

Edgar, R. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research*, 32, 1792–1797.

Edwards, E. & Smith, S.A. (2010). Phylogenetic analyses reveal the shady history of C4 grasses. *Proceedings of the National Academy of Sciences*, 107, 2532-2538.

Eiserhardt, W.L., Svenning, J.C., Borchsenius, F., Kristiansen, T. & Balslev, H. (2013). Separating environmental and geographical determinants of phylogenetic community structure in Amazonian palms (Arecaceae). *Botanical Journal of the Linnean Society*, 171, 244-259.

Elith, J., Graham, C.H., Anderson, R.P., Dudík, M., Ferrier, S., Guisan, A. *et al.* (2006). Novel methods improve prediction of species' distributions from occurrence data. *Ecography*, 29, 129-151.

Elith, J., Phillips, S.J., Hastie, T., Dudík, M., Chee, Y.E. & Yates, C.J. (2011) A statistical explanation of MaxEnt for ecologists. *Diversity and Distributions,* 17, 43-57.

Ewel, J.J. & Whitmore, J.L. (1973). The ecological life zones of Puerto Rico and the U.S. Virgin Islands. USDA Forest Service, Institute of Tropical Forestry.

Faith, D.P. (1992). Conservation evaluation and phylogenetic diversity. *Biological Conservation*, 61, 1-10.

Falster, D.S. & Westoby, M. (2003). Plant height and evolutionary games. *Trends in Ecology and Evolution*, 18, 337-343.

Falster, D.S. & Westoby, M. (2005). Alternative height strategies among 45 dicot rain forest species from tropical Queensland, Australia. *Journal of Ecology*, 93, 521-535.

FAO (2010). Global forest resources assessment 2010. Food and agriculture organization of the United Nations Rome, Italy.

Fazekas, A.J., Burgess, K.S., Kesanakurti, P.R., Graham, S.W., Newmaster, S.G., Husband, B.C. *et al.* (2008). Multiple multilocus DNA barcodes from the plastid genome discriminate plant species equally well. *PlosONE*, 3, e2802. (doi:2810.1371/journal.pone.0002802).

Felsenstein, J. (2004). *Inferring phylogenies*. Sinauer Associates, Sunderland, Massachusetts.

Fielding, A.H. & Bell, J.F. (1997). A review of methods for the assessment of prediction errors in conservation presence-absence models. *Environmental Conservation*, 24, 38-49.

Flynn, D.F.B., Uriarte, M., Crk, T., Pascarella, J.B., Zimmerman, J.K., Aide, T.M. *et al.* (2010). Hurricane disturbance alters secondary forest recovery in Puerto Rico. *Biotropica*, 42, 149-157.

Fortunel, C., Fine, P.V.A. & Baraloto, C. (2012). Leaf, stem and root tissue strategies across 758 Neotropical tree species. *Functional Ecology*, 26, 1153-1161.

Fox, J.W. (2012). When should we expect microbial phenotypic traits to predict microbial abundances? *Frontiers in Microbiology*, 3.

Freschet, G.T., Cornelissen, J.H.C., van Logtestijn, R.S.P. & Aerts, R. (2010). Evidence of the 'plant economics spectrum' in a subarctic flora. *Journal of Ecology*, 98, 362-373.

Fukami, T., Bezemer, T.M., Mortimer, S.R. & van der Putten, W.H. (2005). Species divergence and trait convergence in experimental plant community assembly. *Ecology Letters*, 8, 1283-1290.

Fukami, T. & Wardle, D.A. (2005). Long-term ecological dynamics: reciprocal insights from natural and anthropogenic gradients. *Proceedings of the Royal Society B-Biological Sciences*, 272, 2105-2115.

Funk, J.L. & Cornwell, W.K. (2013). Leaf traits within communities: Context may affect the mapping of traits to function. *Ecology*, 94, 1893-1897.

Gallagher, R.V. & Leishman, M.R. (2012). Contrasting patterns of trait-based community assembly in lianas and trees from temperate Australia. *Oikos*, 121, 2026-2035.

Gamble, D.W. & Curtis, S. (2008). Caribbean precipitation: review, model and prospect. *Progress in Physical Geography*, 32, 265-276.

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

Gelman, A., Su, Y.-S., Hill, J., Grazia Pittau, M., Kerman, J. & Zheng, T. (2006). Package 'arm', version: 1.0-1. Data Analysis Using Regression and Multilevel/Hierarchical Models.

Gesch, D.B. (2007). The National Elevation Dataset. In: *Digital Elevation Model Technologies and Applications: The DEM Users Manual* (ed. Maune, D). American Society for Photogrammetry and Remote Sensing Bethesda, Maryland, pp. 99-118.

Givnish, T.J. (1995). Plant stems: biomechanical adaptation for energy capture and influence on species distributions. In: *Plant stems: physiology and functional morphology* (ed. Gartner, BL). Academic Press San Diego, pp. 3-49.

Goldberg, D. (1990). Components of resource competition in plant communities. In: *Perspectiveson Plant Competition* (eds. Grace, JB & Tilman, D). Academic Press NY, pp. 27-49.

Goldberg, D.E. & Novoplansky, A. (1997). On the relative importance of competition in unproductive environments. *Journal of Ecology*, 85, 409-418.

Gotelli, N.J. (2000). Null model analysis of species co-occurrence patterns. *Ecology*, 81, 2606-2621.

Gotelli, N.J. & Graves, G.R. (1996). *Null Models in Ecology*. Smithsonian Institution Press, Washington, DC.

Gotelli, N.J. & Ulrich, W. (2012). Statistical challenges in null model analysis. *Oikos*, 121, 171-180.

Götzenberger, L., de Bello, F., Bråthen, K.A., Davison, J., Dubuis, A., Guisan, A. *et al.* (2011). Ecological assembly rules in plant communities—approaches, patterns and prospects. *Biological Reviews*, 87, 111-127.

Gould, W.A., Alarcon, C., Fevold, B., Jimenez, M.E., Martinuzzi, S., Potts, G. *et al.* (2008). The Puerto Rico Gap Analysis Project volume 1: land cover, vertebrate species distributions, and land stewardship. Gen. Tech. Rep. IITF-39.

Graham, C.H. & Fine, P.V.A. (2008). Phylogenetic beta diversity: linking ecological and evolutionary processes across space in time. *Ecology Letters*, 11, 1265-1277.

Grime, J.P. (1979). *Plant strategies and vegetation processes*. John Wiley & Sons.

Grime, J.P. (2006). Trait convergence and trait divergence in herbaceous plant communities: mechanisms and consequences. *Journal of Vegetation Science*, 17, 255-260.

Grubb, P.J. (1977). The maintenance of species-richness in plant communities: the importance of the regeneration niche. *Biological Reviews*, 52, 107-145.

Guariguata, M.R. & Ostertag, R. (2001). Neotropical secondary forest succession: changes in structural and functional characteristics. *For. Ecol. Manage.*, 148, 185-206.

Guisan, A. & Thuiller, W. (2005). Predicting species distribution: offering more than simple habitat models. *Ecology Letters*, 8, 993-1009.

Hacke, U.G., Sperry, J.S., Pockerman, W.T., Davis, S.D. & McCulloh, K.A. (2001). Trends in wood density and structure are linked to prevention of xylem implosion by negative pressure. *Oecologia*, 126, 457-461.

Hardy, O.J. (2008). Testing the spatial phylogenetic structure of local communities: statistical performances of different null models and test statistics on a locally neutral community. *Journal of Ecology*, 96, 914-926.

Harvey, P.H. & Pagel, M.D. (1991). *The Comparative Method in Evolutionary Biology*. 1 edn. Oxford University Press, Oxford.

Heath, T.A., Hedtke, S.M. & Hillis, D.M. (2008). Taxon sampling and the accuracy of phylogenetic analyses. *Journal of Systematics and Evolution*, 46, 239-257.

Helmer, E.H., Ramos, O., López, T.d.M., Quiñones, M. & Diaz, W. (2002). Mapping the forest type and land cover of Puerto Rico, a component of the Caribbean biodiversity hotspot. *Caribbean Journal of Science*, 38, 165–183.

Helmus, M.R., Savage, K., Diebel, M.W., Maxted, J.T. & Ives, A.R. (2007). Separating the determinants of phylogenetic community structure. *Ecology Letters*, 10, 917-925.

HilleRisLambers, J., Adler, P.B., Harpole, W.S., Levine, J.M. & Mayfield, M.M. (2012). Rethinking Community Assembly through the Lens of Coexistence Theory. *Annual Review of Ecology, Evolution, and Systematics*, 43, 227-248.

Hofhansl, F., Schnecker, J., Singer, G. & Wanek, W. (2014). New insights into mechanisms driving carbon allocation in tropical forests. *New Phytologist*, n/a-n/a.

Holdridge, L.R. (1947). Determination of world plant formations from simple climatic data. *Science*, 105, 367-368.

Holyoak, M., Leibold, M.A. & Holt, R.D. (2005). *Metacommunities: spatial dynamics and ecological communities*. University of Chicago Press, Chicago, IL.

Ives, A. & Helmus, M. (2010). Phylogenetic Metrics of Community Similarity. *The American Naturalist*, 176, E128-E142.

Jennings, L.N., Douglas, J., Treasure, E. & González, G. (2014). Climate change effects in El Yunque National Forest, Puerto Rico, and the Caribbean region. Gen. Tech. Rep. SRS-GTR-193. Asheville, NC: USDA-Forest Service, Southern Research Station. 47 p.

Johnson, M.T.J. & Stinchcombe, J.R. (2007). An emerging synthesis between community ecology and evolutionary biology. *Trends in Ecology & Evolution*, 22, 250-257.

Jombart, T., Pavoine, S., Devillard, S. & Pontier, D. (2010). Putting phylogeny into the analysis of biological traits: A methodological approach. *Journal of Theoretical Biology*, 264.

Jury, M.R. & Winter, A. (2009). Warming of an elevated layer over the Caribbean. *Climatic Change*, 99, 247–259.

Kallarackal, J., Otieno, D., Reineking, B.r., Jung, E.-Y., Schmidt, M.T., Granier, A. *et al.* (2013). Functional convergence in water use of trees from different geographical regions: a meta-analysis. *Trees*, 27, 787-799.

Katoh, K., Kuma, K., Toh, H. & Miyata, T. (2005). MAFFT version 5: improvement in accuracy of multiple sequence alignment. *Nucleic Acids Research*, 33.

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

Keddy, P. (1989). *Competition*. Chapman & Hall, London.

Kembel, S.W., Cowan, P.D., Helmus, M.R., Cornwell, W.K., Morlon, H., Ackerly, D.D. *et al.* (2010). Picante: R tools for integrating phylogenies and ecology. *Bioinformatics*, 26, 1463-1464.

Kembel, S.W. & Hubbell, S.P. (2006). The phylogenetic structure of a neotropical forest tree community. *Ecology*, 87, 86-99.

Kinzig, A.P. & Pacala, S.W. (2001). Successional biodiversity and ecosystem functioning. In: *The functional consequences of biodiversity: empirical progress and theoretical extensions* (eds. Kinzig, AP, Pacala, S & Tilman, GD). Princeton University Press Princeton, pp. 175-212.

Kivlin, S.N., Hawkes, C.V. & Treseder, K.K. (2011). Global diversity and distribution of arbuscular mycorrhizal fungi. *Soil Biology and Biochemistry*, 43, 2294-2303.

Kooyman, R., Cornwell, W. & Westoby, M. (2010). Plant functional traits in Australian subtropical rain forest: partitioning within-community from cross-landscape variation. *Journal of Ecology*, 98, 517-525.

Kraft, N.J.B. & Ackerly, D.D. (2010). Functional trait and phylogenetic tests of community assembly across spatial scales in an Amazonian forest. *Ecol. Monogr.*, 80, 401-422.

Kraft, N.J.B., Cornwell, W.K., Webb, C.O. & Ackerly, D.D. (2007). Trait evolution, community assembly, and the phylogenetic structure of ecological communities. *The American Naturalist*, 170, 271-283.

Kraft, N.J.B., Metz, M.R., Condit, R.S. & Chave, J. (2010). The relationship between wood density and mortality in a global tropical forest data set. *New Phytologist*, 188, 1124-1136.

Kraft, N.J.B., Valencia, R. & Ackerly, D.D. (2008). Functional traits and niche-based tree community assembly in an Amazonian forest. *Science (New York, N.Y.)*, 322, 580-582.

Kress, W.J. & Erickson, D.L. (2007). A two-locus global DNA barcode for land plants: the coding *rbcL* gene complements the non-coding *trnH-psbA* spacer region. *PLoS ONE*, 2, e508. doi:510.1371/journal.pone.0000508.

Kress, W.J., Erickson, D.L., Jones, F.A., Swenson, N.G., Perez, R., Sanjur, O. *et al.* (2009). Plant DNA barcodes and a community phylogeny of a tropical forest dynamics plot in Panama. *Proceedings of the National Academy of Sciences*, 106, 18621-18626.

Kress, W.J., Erickson, D.L., Swenson, N.G., Thompson, J., Uriarte, M. & Zimmerman, J.K. (2010). Advances in the use of DNA barcodes to build a community phylogeny for tropical trees in a Puerto Rican forest dynamics plot. *PLoS ONE*, 5, e15409.

Kunstler, G., Lavergne, S., Courbaud, B., Thuiller, W., Vieilledent, G., Zimmerman, N.E. *et al.* (2012). Competitive interactions between forest trees are driven by species' trait hierarchy, not phylogenetic or functional similarity: implications for forest community assembly. *Ecology Letters*, 15, 831-840.

Laliberté, E. & Shipley, B. (2011). FD: measuring functional diversity from multiple traits, and other tools for functional ecology. R package version 1.0-11.

Lasky, J.R., Uriarte, M., Boukili, V. & Chazdon, R.L. (2014a). Multiple effects of functional trait variation on successional forest dynamics. *Proceedings of the National Academy of Sciences*, 111, 5616-5621.

Lasky, J.R., Uriarte, M., Boukili, V.K., Erickson, D.L., John Kress, W. & Chazdon, R.L. (2014b). The relationship between tree biodiversity and biomass dynamics changes with tropical forest succession. *Ecology Letters*, 17, 1158-1167.

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

Lessard, J.-P., Borregaard, M.K., Fordyce, J.A., Rahbek, C., Weiser, M.D., Dunn, R.R. *et al.* (2012b). Strong influence of regional species pools on continent-wide structuring of local communities. *Proceedings of Royal Society B*, 279, 266-274.

Letcher, S.G. (2010). Phylogenetic structure of angiosperm communities during tropical forest succession. *Proceedings of the Royal Society B: Biological Sciences*, 277, 97-104.

Letcher, S.G., Chazdon, R.L., Andrade, A.C.S., Bongers, F., van Breugel, M., Finegan, B. *et al.* (2012). Phylogenetic community structure during succession: evidence from three neotropical forest sites. *Perspectives in Plant Ecology, Evolution and Systematics*, 14, 79-87.

Levins, R. & Culver, D. (1971). Regional coexistence of species and competition between rare species. *Proceedings of the National Academy of Sciences of the United States of America*, 68, 1246-1248.

Liogier, H.A. (1996). Botany and botanists in Puerto Rico. *Annals of the New York Academy of Sciences*, 776, 41-53.

Little, E.L. & Wadsworth, F.H. (1964). *Common trees of Puerto Rico and the Virgin Islands*. USDA Forest Service, Washington, DC.

Little, E.L., Woodbury, R.O. & Wadsworth, F.H. (1974). *Trees of Puerto Rico and the Virgin Islands*. USDA Forest Service, Washington, DC.

Liu, K., Warnow, T.J., Holder, M.T., Nelesen, S.M., Yu, J., Stamatakis, A. *et al.* (2012). SATé-II: very fast and accurate simultaneous estimation of multiple sequence alignments and phylogenetic trees. *Systematic Biology*, 61, 90-106.

Loehle, C. (2000). Strategy space and the disturbance spectrum: a life-history model for tree species coexistence. *The American Naturalist*, 156, 14-33.

Lohbeck, M., Poorter, L., Lebrija-Trejos, E., Martínez-Ramos, M., Meave, J.A., Paz, H. *et al.* (2013). Successional changes in functional composition contrast for dry and wet tropical forest. *Ecology*, 94, 1211-1216.

Lohbeck, M., Poorter, L., Martínez-Ramos, M., Rodriguez-Velázquez, J., van Breugel, M. & Bongers, F. (2014). Changing drivers of species dominance during tropical forest succession. *Functional Ecology*, 28, 1052-1058.

Losos, J.B. (1996). Phylogenetic perspectives on community ecology. *Ecology*, 77, 1344-1344.

Lugo, A.E., Castro, L.M., Vale, A., López, T.d.M., Prieto, E.H., Martinó, A.G. *et al.* (2001). Puerto Rican karst – a vital resource. General Technical Report WO-65. August 2001. Available online: http://www.fs.fed.us/global/iitf/karst.pdf.

Lugo, A.E. & Murphy, P.G. (1986). Nutrient dynamics of a Puerto Rican subtropical dry forest. *Journal of Tropical Ecology*, 2, 55-72.

MacArthur, R.H. & Levins, R. (1967). The Limiting Similarity, Convergence, and Divergence of Coexisting Species. *The American Naturalist*, 101, 377-385.

Magallón, S. & Castillo, A. (2009). Angiosperm diversification through time. *American Journal of Botany*, 96, 349-365.

Markesteijn, L. & Poorter, L. (2009). Seedling root morphology and biomass allocation of 62 tropical tree species in relation to drought- and shade-tolerance. *Journal of Ecology*, 97, 311-325.

Markesteijn, L., Poorter, L., Bongers, F., Paz, H. & Sack, L. (2011a). Hydraulics and life history of tropical dry forest tree species: coordination of species' drought and shade tolerance. *New Phytologist*, 191, 480-495.

Markesteijn, L., Poorter, L., Paz, H., Sack, L. & Bongers, F. (2011b). Ecological differentiation in xylem cavitation resistance is associated with stem and leaf structural traits. *Plant, Cell and Environment*, 34, 137-148.

Mayfield, M.M. & Levine, J.M. (2010). Opposing effects of competitive exclusion on the phylogenetic structure of communities. *Ecology Letters*, 13, 1085-1093.

McGill, B. (2012). Trees are rarely most abundant where they grow best. *Journal of Plant Ecology*, 5, 46-51.

McGill, B., Enquist, B.J., Weiher, E. & Westoby, M. (2006). Rebuilding community ecology from functional traits. *Trends in Ecology & Evolution*, 21, 178-185.

McMahon, M.M. & Sanderson, M.J. (2006). Phylogenetic supermatrix analysis of GenBank sequences from 2228 Papilionoid Legumes. *Systematic Biology*, 55, 818-836.

Merow, C., Smith, M. & Silander, J.A. (2013) A practical guide to Maxent: what it does, and why inputs and settings matter. *Ecography,* 36**,** 1-12.

Merow, C. & Silander, J.A. (2014). A comparison of Maxlike and Maxent for modelling species distributions. *Methods in Ecology and Evolution*, 5, 215-225.

Mi, X., Swenson, N.G., Valencia, R., Kress, W.J., Erickson, D.L., Pérez, A.J. *et al.* (2012). The contribution of rare species to community phylogenetic diversity across a global network of forest plots. *American Naturalist*, 180, E17-30.

Miller, G.L. & Lugo, A.E. (2009). Guide to the ecological systems of Puerto Rico. Gen. Tech. Rep. IITF-GTR-35. (eds. United States Department of Agriculture, Forest Service & International Institute of Tropical Forestry) San Juan, PR, p. 437.

Miller, M.A., Pfeiffer, W. & Schwartz, T. (2010). Creating the CIPRES Science Gateway for inference of large phylogenetic trees. In: *Proceedings of the Gateway Computing Environments Workshop (GCE)* New Orleans, LA, pp. 1-8.

Mitchell, S.J. (2012). Wind as a natural disturbance agent in forests: a synthesis. *Forestry*.

Mooney, H.A. & Dunn, E.L. (1970). Convergent evolution of Mediterranean-climate evergreen sclerophyllous shrubs. *Evolution*, 24.

Muller-Landau, H.C. (2010). The tolerance-fecundity trade-off and the maintenance of diversity in seed size. *Proceedings of the National Academy of Sciences*, 107, 4242-4247.

Münkemüller, T., Gallien, L., Lavergne, S., Renaud, J., Roquet, C., Abdulhak, S. *et al.* (2014). Scale decisions can reverse conclusions on community assembly processes. *Global Ecology and Biogeography*, 23, 620-632.

Murphy, P.G. & Lugo, A.E. (1986). Structure and biomass of a subtropical dry forest in Puerto Rico. *Biotropica*, 18, 89-96.

Muscarella, R., Galante, P.J., Soley-Guardia, M., Boria, R.A., Kass, J.M., Uriarte, M. *et al.* (2014a). ENMeval: An R package for conducting spatially independent evaluations and estimating optimal model complexity for Maxent ecological niche models. *Methods in Ecology and Evolution*, n/a-n/a.

Muscarella, R., Uriarte, M., Erickson, D.L., Swenson, N.G., Zimmerman, J.K. & Kress, W.J. (2014b). A well-resolved phylogeny of the trees of Puerto Rico based on DNA barcode sequence data. *PLoS ONE*. 9, e112843.

Muscarella, R., Uriarte, M., Forero-Montaña, J., Comita, L.S., Swenson, N.G., Thompson, J. *et al.* (2013). Life-history trade-offs during the seed-to-seedling transition in a subtropical wet forest community. *Journal of Ecology*, 101, 171-182.

Nabhan, A.R. & Sarkar, I.N. (2012). The impact of taxon sampling on phylogenetic inference: a review of two decades of controversy. *Briefings in Bioinformatics*, 13, 122-134.

Nakagawa, S. & Schielzeth, H. (2012). A general and simple method for obtaining R^2 from generalized linear mixed-effects models. *Methods in Ecology and Evolution*, 4, 133-142.

Navas, M.L. & Violle, C. (2009). Plant traits related to competition: how do they shape the functional diversity of communities? *Community Ecology*, 10, 131-137.

Neelin, J.D., Münnich, M., Su, H., Meyerson, J.E. & Holloway, C.E. (2006). Tropical drying trends in global warming models and observations. *Proceedings of the National Academy of Sciences*, 103, 6110-6115.

Niinemets, Ü. (2001). Global-scale climatic controls of leaf dry mass per area, density, and thickness in trees and shrubs. *Ecology*, 82, 453-469.

Niinemets, Ü. (2014). Is there a species spectrum within the world-wide leaf economics spectrum? Major variations in leaf functional traits in the Mediterranean sclerophyll Quercus ilex. *New Phytologist*, n/a-n/a.

Niinemets, Ü. & Valladares, F. (2006). Tolerance to shade, drought, and waterlogging of temperate northern hemisphere trees and shrubs. *Ecol. Monogr.*, 76, 521-547.

Norden, N., Letcher, S.G., Boukili, V., Swenson, N.G. & Chazdon, R. (2012). Demographic drivers of successional changes in phylogenetic structure across life-history stages in plant communities. *Ecology*, 93, S70-S82.

Orme, D., Freckleton, R.P., Thomas, G., Petzoldt, T., Fritz, S. & Isaac, N. (2011). caper: Comparative Analyses of Phylogenetics and Evolution in R.

Pacala, S.W. & Rees, M. (1998). Models Suggesting Field Experiments to Test Two Hypotheses Explaining Successional Diversity. *The American Naturalist*, 152, 729-737.

Pagel, M. (1999). Inferring the historical patterns of biological evolution. *Nature*, 401, 877–884.

Parmentier, I., Réjou-Méchain, M., Chave, J., Vleminckx, J., Thomas, D.W., Kenfack, D. *et al.* (2014). Prevalence of phylogenetic clustering at multiple scales in an African rain forest tree community. *Journal of Ecology*, n/a-n/a.

Pärtel, M., Szava-Kovats, R. & Zobel, M. (2011). Dark diversity: shedding light on absent species. *Trends in Ecology and Evolution*, 26, 124–128.

Patiño, S., Fyllas, N.M., Baker, T.R., Paiva, R., Quesada, C.A., Santos, A.J.B. *et al.* (2012). Coordination of physiological and structural traits in Amazon forest trees. *Biogeosciences*, 9, 775-801.

Pavoine, S. & Bonsall, M.B. (2010). Measuring biodiversity to explain community assembly: a unified approach. *Biological Reviews*, no-no-no-no.

Pennington, R.T., Lavin, M. & Oliveira-Filho, A. (2009). Woody plant diversity, evolution, and ecology in the tropics: Perspectives from seasonally dry tropical forests. *Annual Review of Ecology & Systematics*, 40, 437-457.

Peterson, A.T., Soberón, J., Pearson, R.G., Anderson, R.P., Martínez-Meyer, E., Nakamura, M. *et al.* (2011). *Ecological Niches and Geographic Distributions.* Monographs in Population Biology, 49. Princeton University Press, Princeton, NJ.

Phillipe, H., Snell, E.A., Bapteste, E., Lopez, P., Holland, P.W.H. & Casane, D. (2004). Phylogenomics of eukaryotes: impact of missing data on large alignments. *Molecular Biology and Evolution*, 21, 1740–1752.

Phillips, S.J., Anderson, R.P. & Schapire, R.E. (2006). Maximum entropy modeling of species geographic distributions. *Ecol. Model.*, 190, 231-259.

Pickett, S.T.A. & Bazzaz, F.A. (1978). Organization of an assemblage of early successional species on a soil moisture gradient. *Ecology*, 59, 1248-1255.

Poe, S. (2003). Evaluation of the strategy of long-branch subdivision to improve the accuracy of phylogenetic methods. *Systematic Biology*, 52, 423–428.

Poe, S. & Swofford, D.L. (1999). Taxon sampling revisited. *Nature*, 398, 299–300.

Poorter, H., Niinemets, U., Poorter, L., Wright, I.J. & Villar, R. (2009). Causes and consequences of variation in leaf mass per area (LMA): a meta-analysis. *New Phytologist*, 182, 565-588.

Poorter, L., McDonald, I., Alarcón, A., Fichtler, E., Licona, J.-C., Peña-Claros, M. *et al.* (2010). The importance of wood traits and hydraulic conductance for the performance and life history strategies of 42 rainforest tree species. *New Phytologist*, 185, 481-492.

Poorter, L., van de Plassche, M. & Boot, R.G.A. (2004). Leaf traits and herbivory rates of tropical tree species differing in successional status. *Plant Biology*, 6, 746-754.

Poorter, L., Wright, S.J., Paz, H., Ackerly, D.D., Condit, R., Ibarra-Manríquez, G. *et al.* (2008). Are functional traits good predictors of demographic rates? Evidence from five neotropical forests. *Ecology*, 89, 1908-1920.

Purvis, A. & Hector, A. (2000). Getting the measure of biodiversity. *Nature*, 405, 212-219.

Qian, H. & Ricklefs, R.E. (2007). A latitudinal gradient in large-scale beta diversity for vascular plants in North America. *Ecology Letters*, 10, 737-744.

Quesada, C.A., Phillips, O.L., Schwarz, M., Czimczik, C.I., Baker, T.R., Patiño, S. *et al.* (2012). Basin-wide variations in Amazon forest structure and function are mediated by both soils and climate. *Biogeosciences*, 9, 2203-2246.

R Development Core Team (2014). R: A language and environment for statistical computing. R Foundation for Statistical Computing Vienna, Austria.

Radosavljevic, A. & Anderson, R.P. (2014) Making better Maxent models of species distributions: complexity, overfitting and evaluation. *Journal of Biogeography,* 41, 629-643.

Rannala, B., Huelsenbeck, J.P., Yang, Z. & Nielsen, R. (1998). Taxon sampling and the accuracy of large phylogenies. *Systematic Biology*, 47, 702–710.

Reich, P.B. (2014). The world-wide ‘fast–slow’ plant economics spectrum: a traits manifesto. *Journal of Ecology*, 102, 275-301.

Reich, P.B., Ellsworth, D.S. & Uhl, C. (1995). Leaf carbon and nutrient assimilation and conservation in species of differing successional status in an oligotrophic Amazonian forest. *Functional Ecology*, 9, 65-76.

Reich, P.B., Walters, M.B. & Ellsworth, D.S. (2007). From tropics to tundra: Global convergence in plant functioning. *Proceedings of the National Academy of Sciences*, 94, 13730-13734.

Ricklefs, R. & Jenkins, D.G. (2011). Biogeography and ecology: towards the integration of two disciplines. *Philosophical Transactions of the Royal Society B-Biological Sciences*, 366, 2438-2448.

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

Ricklefs, R.E. (2008). Disintegration of the ecological community. *The American Naturalist*, 172, 741-750.

Ricklefs, R.E. (2010). Evolutionary diversification, coevolution between populations and their antagonists, and the filling of niche space. *Proceedings of the National Academy of Sciences*, 107, 1265–1272.

Ricklefs, R.E. & Schluter, D. (1993). *Species diversity in ecological communities: historical and geographical perspectives*. University of Chicago Press, Chicago, IL.

Rivera, O.A.M. (2009). Vascular flora of the Guánica dry forest, Puerto Rico. In: *Biology*. University of Puerto Rico Mayagüez.

Roff, D.A. (1992). *The evolution of life histories: theory and analysis*. Oxford University Press, Oxford.

Rudel, T.K., Perez-Lugo, M. & Zichal, H. (2000). When fields revert to forest: development and spontaneous reforestation in post-war Puerto Rico. *Professional Geographer*, 52, 386–397.

Ryan, M.G., Phillips, N. & Bond, B.J. (2006). The hydraulic limitation hypothesis revisited. *Plant Cell and Environment*, 29, 367-381.

Ryan, M.G. & Yoder, B.J. (1997). Hydraulic Limits to Tree Height and Tree Growth. *BioScience*, 47, 235-242.

Sagarin, R.D. & Gaines, S.D. (2002). The `abundant centre' distribution: to what extent is it a biogeographical rule? *Ecology Letters*, 5, 137-147.

Santiago-Valentin, E. & Olmstead, R.G. (2004). Historical biogeography of Caribbean plants: introduction to current knowledge and possibilities from a phylogenetic perspective. *Taxon*, 53, 299–319.

Schimper, A.F.W. (1903). *Plant-geography upon a physiological basis*. Clarendon Press, Oxford, UK.

Schoch, C.L., Crous, P.W., Groenewald, J.Z., Boehm, E.W.A., Burgess, T.I., de Gruyter, J. *et al.* (2009). A class-wide phylogenetic assessment of Dothideomycetes. *Studies in Mycology*, 64, 1-15-S10.

Sedio, B.E., Wright, S.J. & Dick, C.W. (2012). Trait evolution and the coexistence of a species swarm in the tropical forest understory. *Journal of Ecology*, 100, 1183-1193.

Shcheglovitova, M. & Anderson, R.P. (2013) Estimating optimal complexity for ecological niche models: A jackknife approach for species with small sample sizes. *Ecological Modelling,* 269, 9-17.

Shipley, B. (2010). *From plant traits to vegetation structure: chance and selection in the assembly of ecological communities*. Cambridge University Press, New York.

Shipley, B., Laughlin, D.C., Sonnier, G. & Otfinowski, R. (2011). A strong test of a maximum entropy model of trait-based community assembly. *Ecology*, 92, 507-517.

Sibly, R.M. (1997). Life history evolution in heterogeneous environments: a review of theory. In: *Plant life histories: Ecology, phylogeny and evolution* (eds. Silvertown, J, Franco, M & Harper, JL). Cambridge University Press Cambridge, UK.

Silvertown, J., Araya, Y. & Gowing, D. (2014). Hydrological niches in terrestrial plant communities: A review. *Journal of Ecology*, n/a-n/a.

Silvertown, J., Dodd, M., Gowing, D., Lawson, C. & McConway, K. (2006). Phylogeny and the hierarchical organization of plant diversity. *Ecology*, 87, 39-49.

Smith, C.C. & Fretwell, S.D. (1974). The optimal balance between size and number of offspring. *American Naturalist*, 108, 499–506.

Smith, S.A., Beaulieu, J. & Donoghue, M.J. (2009). Mega-phylogeny approach for comparative biology: an alternative to supertree and supermatrix approaches. *BMC Evolutionary Biology*, 9, 37.

Spasojevic, M.J. & Suding, K.N. (2012). Inferring community assembly mechanisms from functional diversity patterns: the importance of multiple assembly processes. *Journal of Ecology*, 100, 652-661.

Stefanović, S., Rice, D.W. & Palmer, J.D. (2004). Long branch attraction, taxon sampling, and the earliest angiosperms: Amborella or monocots? *BMC Evolutionary Biology*, 4, 35.

Sterck, F., Markesteijn, L., Toledo, M., Schieving, F. & Poorter, L. (2014). Sapling performance along resource gradients drives tree species distributions within and across tropical forests. *Ecology*, 95, 2514-2525.

Stevens, P.F. (2001 onwards). Angiosperm Phylogeny Website. Version 12, July 2012 [and more or less continuously updated since]. Available: www.mobot.org/MOBOT/research/APweb.

Svenning, J.-C., Gravel, D., Holt, R.D., Schurr, F.M., Thuiller, W., Münkemüller, T. *et al.* (2014). The influence of interspecific interactions on species range expansion rates. *Ecography*, no-no.

Swenson, N.G. (2009). Phylogenetic resolution and quantifying the phylogenetic diversity and dispersion of communities. *PLoS ONE*, 4, e4390-e4390.

Swenson, N.G. (2011). The role of evolutionary processes in producing biodiversity patterns, and the interrelationships between taxonomic, functional and phylogenetic biodiversity. *Am. J. Bot.*, ajb.1000289-ajb.1000289.

Swenson, N.G. (2013). The assembly of tropical tree communities – the advances and shortcomings of phylogenetic and functional trait analyses. *Ecography*, 36.

Swenson, N.G. & Enquist, B.J. (2007). Ecological and evolutionary determinants of a key plant functional trait: wood density and its community-wide variation across latitude and elevation. *American Journal of Botany*, 94, 451-459.

Swenson, N.G. & Enquist, B.J. (2008). The relationship between stem and branch wood specific gravity and the ability of each measure to predict leaf area. *American Journal of Botany*, 95, 516-519.

Swenson, N.G., Enquist, B.J., Jason, P., Thompson, J. & Zimmerman, J.K. (2006). The problem and promise of scale dependency in community phylogenetics. *Ecology*, 87, 2418-2424.

Swenson, N.G., Enquist, B.J., Thompson, J. & Zimmerman, J.K. (2007). The influence of spatial and size scale on phylogenetic relatedness in tropical forest communities. *Ecology*, 88, 1770-1780.

Swenson, N.G., Erickson, D.L., Mi, X.C., Bourg, N.A., Forero-Montaña, J., Ge, X.-J. *et al.* (2012a). Phylogenetic and functional alpha and beta diversity in temperate and tropical tree communities.

Swenson, N.G., Stegen, J.C., Davies, S.J., Erickson, D., Forero-Montaña, J., Hurlbert, A.H. *et al.* (2012b). Temporal turnover in the composition of tropical tree communities: functional determinism and phylogenetic stochasticity. *Ecology*, 93, 490-499.

Swenson, N.G. & Umaña, M.N. (2014). Phylofloristics: an example from the Lesser Antilles. *Journal of Plant Ecology*, 7, 166-175.

Thomas, S.C. & Bazzaz, F.A. (1999). Asymptotic height as a predictor of photosynthetic characteristics in Malaysian rain forest trees. *Ecology*, 80, 1607-1622.

Thompson, J., Brokaw, N., Zimmerman, J.K., Waide, R.B., Everham, E.M., Lodge, D.J. *et al.* (2002). Land use history, environment, and tree composition in a tropical forest. *Ecological Applications*, 12, 1344-1363.

Thuiller, W., Münkemüller, T., Schiffers, K.H., Georges, D., Dullinger, S., Eckhart, V.M. *et al.* (2014). Does probability of occurrence relate to population dynamics? *Ecography*, n/a-n/a.

Tilman, D. (1988). *Plant Strategies and the Dynamics and Structure of Plant Communities*. Princeton University Press, Princeton, NJ.

Tilman, D. (1994). Competition and biodiversity in spatially structured habitats. *Ecology*, 75, 2-2.

Uriarte, M., Bruna, E.M., Rubim, P., Anciães, M. & Jonckheere, I. (2010a). Effects of forest fragmentation on the seedling recruitment of a tropical herb: assessing seed vs. safe-site limitation. *Ecology*, 91, 1317-1328.

Uriarte, M., Canham, C.D., Thompson, J., Zimmerman, J.K., Murphy, L., Sabat, A.M. *et al.* (2009). Natural disturbance and human land use as determinants of tropical forest dynamics: results from a forest simulator. *Ecol. Monogr.*, 79, 423-443.

Uriarte, M., Swenson, N.G., Chazdon, R.L., Comita, L.S., Kress, W.J., Erickson, D. *et al.* (2010b). Trait similarity, shared ancestry and the structure of neighbourhood interactions in a subtropical wet forest: implications for community assembly. *Ecology Letters*, 13, 1503–1514.

USFS (2006). Forest inventory and analysis national core field guide: Caribbean version 3.0. (ed. USDA). United States Forest Service.

Valladares, F. & Niinemets, Ü. (2008). Shade Tolerance, a Key Plant Feature of Complex Nature and Consequences. *Annual Review of Ecology, Evolution, and Systematics*, 39, 237-257.

Vamosi, S.M., Heard, S.B., Vamosi, J.C. & Webb, C.O. (2009). Emerging patterns in the comparative analysis of phylogenetic community structure. *Mol. Ecol.*, 18, 572-592.

VanDerWal, J., Shoo, L.P., Johnson, C.N. & Williams, S.E. (2009). Abundance and the environmental niche: environmental suitability estimated from niche models predicts the upper limit of local abundance. *The American Naturalist*, 174, 282-291.

Vane-Wright, R.I., Humphries, C.J. & Williams, P.H. (1991). What to protect? Systematics and the agony of choice. *Biological Conservation*, 55, 235-254.

Vellend, M., Cornwell, W.K., Magnuson-Ford, K. & Mooers, A.Ø. (2011). Measuring phylogenetic biodiversity. In: *Biological diversity: frontiers in measurement and assessment* (eds. Magurran, AE & McGill, BJ). Oxford University Press Oxford, pp. 193-206.

Violle, C., Enquist, B.J., McGill, B.J., Jiang, L., Albert, C.c.H., Hulshof, C. *et al.* (2012). The return of the variance: intraspecific variability in community ecology. *Trends in Ecology & Evolution*, 27, 244-252.

Violle, C., Garnier, E., Lecoeur, J., Roumet, C., Podeur, C., Blanchard, A. *et al.* (2009). Competition, traits and resource depletion in plant communities. *Oecologia*, 160, 747-755.

Violle, C., Navas, M.L., Vile, D., Kazakou, E., Fortunel, C., Hummel, I. *et al.* (2007). Let the concept of trait be functional! *Oikos*, 116, 882-892.

Violle, C., Reich, P.B., Pacala, S.W., Enquist, B.J. & Kattge, J. (2014). The emergence and promise of functional biogeography. *Proceedings of the National Academy of Sciences*, 111, 13690-13696.

Walker, L.R., Wardle, D.A., Bardgett, R.D. & Clarkson, B.D. (2010). The use of chronosequences in studies of ecological succession and soil development. *Journal of Ecology*, 98, 725-736.

Warren, D.L. & Seifert, S.N. (2011). Ecological niche modeling in Maxent: the importance of model complexity and the performance of model selection criteria. *Ecological Applications*, 21, 335-342.

Webb, C.O., Ackerly, D.D. & Kembel, S.W. (2008). Phylocom: software for the analysis of phylogenetic community structure and trait evolution. *Bioinformatics*, 24, 2098-2100.

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

Webb, C.O. & Donoghue, M.J. (2005). Phylomatic: tree assembly for applied phylogenetics. *Molecular Ecology Notes*, 5, 181–183.

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

Weiher, E. & Keddy, P. (1995). Assembly rules, null models, and trait dispersion: new questions from old patterns. *Oikos*, 74, 159-164.

Westoby, M., Falster, D.S., Moles, A.T., Vesk, P.A. & Wright, I.J. (2002). Plant ecological strategies: some leading dimensions of variation between species. *Annual Review of Ecology & Systematics*, 33, 125-159.

Westoby, M. & Wright, I.J. (2006). Land-plant ecology on the basis of functional traits. *Trends in Ecology and Evolution*, 21, 261-269.

Whitfeld, T.J.S., Kress, W.J., Erickson, D.L. & Weiblen, G.D. (2012). Change in community phylogenetic structure during tropical forest succession: evidence from New Guinea. *Ecography*, 35, 821-830.

Whittaker, R.H. (1967). Gradient analysis of vegetation. *Biological Reviews*, 49, 207-264.

Wiemann, M.C. & Williamson, G.B. (2002). Geographic variation in wood specific gravity: effects of latitude, temperature, and precipitation. *Wood Fiber Science*, 34, 96–107.

Wiens, J.J. (1998). Does adding characters with missing data increase or decrease phylogenetic accuracy? *Systematic Biology*, 47, 625-640.

Wiens, J.J. (2006). Missing data and the design of phylogenetic analyses. *Journal of Biomedical Informatics*, 39, 34-42.

Wiens, J.J. & Morrill, M.C. (2011). Missing data in phylogenetic analysis: reconciling results from simulations and empirical data. *Systematic Biology*, 60, 719-731.

Williams, P.H., Humphries, C.J. & Vane-Wright, R.I. (1991). Measuring biodiversity: taxonomic relatedness for conservation priorities. *Australian Systematic Botany*, 4, 665–679.

Willis, C.G., Ruhfel, B., Primack, R.B., Miller-Rushing, A.J. & Davis, C.C. (2008). Phylogenetic patterns of species loss in Thoreau's woods are driven by climate change. *Proceedings of the National Academy of Sciences*, 105, 17029-17033.

Wisz, M.S., Pottier, J., Kissling, W.D., Pellissier, L., Lenoir, J., Damgaard, C.F. *et al.* (2012). The role of biotic interactions in shaping distributions and realised assemblages of species: implications for species distribution modelling. *Biological Reviews*, 88, 15-30.

Wright, I.J., Ackerly, D.D., Bongers, F., Harms, K.E., Ibarra-Manríquez, G., Martínez-Ramos, M. *et al.* (2007). Relationships among ecologically important dimensions of plant trait variation in seven neotropical forests. *Annals of Botany*, 99, 1003-1015.

Wright, I.J., Reich, P.B., Westoby, M., Ackerly, D.D., Baruch, Z., Bongers, F. *et al.* (2004). The worldwide leaf economics spectrum. *Nature*, 428, 821-827.

Wright, I.J. & Westoby, M. (2002). Leaves at low versus high rainfall: coordination of structure, lifespan and physiology. *New Phytologist*, 155, 403–416.

Wright, I.J., Westoby, M. & Reich, P.B. (2001). Strategy shifts in leaf physiology, structure and nutrient content between species of high- and low-rainfall and high- and low-nutrient habitats. *Functional Ecology*, 15, 423-434.

Wright, S.J., Kitajima, K., Kraft, N.J.B., Reich, P.B., Wright, I.J., Bunker, D.E. *et al.* (2010). Functional traits and the growth–mortality trade-off in tropical trees. *Ecology*, 91, 3664-3674.

Wright, S.J., Muller-Landau, H.C., Condit, R. & Hubbell, S.P. (2003). Gap-dependent recruitment, realized vital rates, and size distributions of tropical trees. *Ecology*, 84, 3174-3185.

Yackulic, C., Sanderson, E.W. & Uriarte, M. (2011a). Anthropogenic and intrinsic drivers of intra-specific variation in extinction risk. *Proceedings of the National Academy of Sciences, USA*, 108, 2044-2049.

Yackulic, C.B., Fagan, M., Jain, M., Jina, A., Lim, Y., Marlier, M. *et al.* (2011b). Biophysical and socioeconomic factors associated with forest transitions at multiple spatial and temporal scales. *Ecology and Society*, 16.

Zimmerman, J.K., Iii, E.M.E., Waide, R.B., Lodge, D.J., Taylor, C.M. & Brokaw, N.V.L. (1994). Responses of tree species to hurricane winds in subtropical wet forest in Puerto Rico: implications for tropical tree life histories. *Journal of Ecology*, 82, 911-922.

APPENDICES

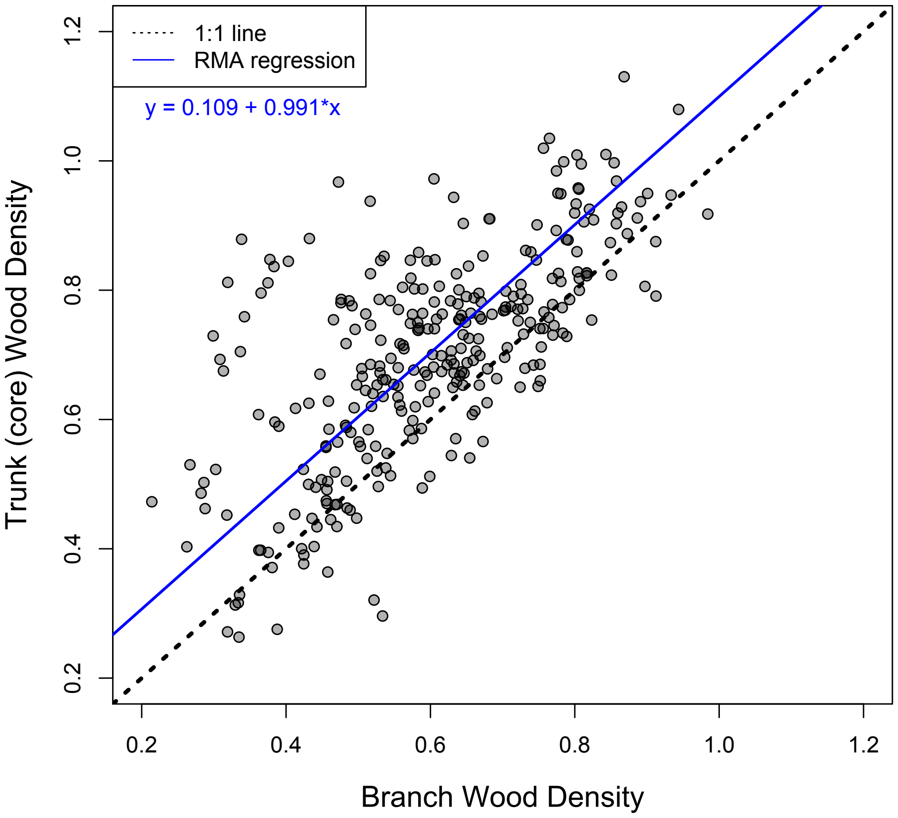
Appendix 1. Wood density sampling and conversion of branch wood samples.

Whenever possible, we measured wood density (WD; g cm^{-3}) using an increment borer in trees 10-20 cm in diameter at 1 m off the ground. A number of species in our study, however, rarely (or never) reach this size. To cope with these limitations and still generate a relatively complete and comparable database of WD for the majority of species in our study, we adapted the procedure discussed in Swenson and Enquist (2008). Specifically, we measured wood density on trunks (using an increment borer; 'core wood density') and branch sections ('branch wood density') on 302 individual trees across our study plots (Fig. S1).

For both core and branch wood samples, we measured the length and diameter of each sample while wet, and then calculated the volume based on the shape of a cylinder ($V = \pi \times radius^2 \times height$). Then, we oven-dried samples at 100° C for at least 48 hours or until weights stabilized. We measured the dry mass of the samples using an electronic balance.

We used RMA regression to determine the relationship between the core and branch WD measurements and used this relationship to convert branch wood density measurements to core wood density estimates. The estimated intercept and slope of the RMA regression were 0.109 and 0.991, respectively (Fig. S1). The R^2 value for this relationship was 0.45 ($p << 0.001$). The 95% confidence intervals for the intercept are 0.064–0.151, and for the slope are 0.921–1.064. We used these regression results to convert branch wood samples into values comparable to the core samples, which represent the predominant sample type in our dataset.

Figure S1. RMA regression for 302 core and branch wood density samples from the same individuals representing 99 species located across 8 forests in Puerto Rico.



Appendix 2. Species-specific tuning and evaluation of ecological niche models.

To achieve a balance between fit and predictive ability of ecological niche models, we conducted species-specific tuning of Maxent settings using the R package ENMeval v 0.1.1 (Muscarella *et al.* 2014a). Specifically, for each species, we built models with all combinations (N = 48) of the regularization multiplier values (0.5, 1.0, 1.5, 2.0, 2.5, 3.0, 3.5, 4.0) and feature classes (L, LQ, H, LQH, LQHP, LQHPT) (Phillips *et al*. 2006; Elith *et al*. 2011; Merow *et al*. 2013). For each unique setting combination, we used *k*-fold cross validation to evaluate model performance. This consisted of first partitioning occurrence data into *k* separate bins for evaluation. For species with fewer than 20 occurrence records (n = 19), we partitioned data using the *k*-1 jackknife method in ENMeval v 0.1.1 (Muscarella *et al.* 2014a) where the number of test bins is equal to the number of occurrence records (Shcheglovitova & Anderson 2013). For species with larger sample sizes (n = 154), we partitioned data using the 'checkerboard2' method in ENMeval v 0.1.1 (Muscarella *et al.* 2014a), with both the fine and coarse scale aggregation factors set to 5. This method spatially partitions both occurrence records and background information (i.e., the masked geographically structured approach of Radosavljevic & Anderson 2014). We evaluated model performance using (1) the threshold-independent test AUC value, and (2) the threshold-dependent test point omission rate based on the minimum training presence value (OR_{MTP}). The threshold-independent metric AUC is based on predicted suitability values for the testing localities (i.e., localities withheld during model training), averaged over *k* iterations. Higher values reflect a better ability for a model to discriminate between conditions at withheld (testing) occurrence localities and those of background localities (by ranking the former higher than the latter based on their predicted suitability values). The rank-based AUC does not indicate model fit. AUC values range from 0 to 1 and increase with the models ability to discriminate test localities from background localities (Fielding & Bell 1997). The threshold-dependent metric OR_{MTP} indicates the proportion of testing localities with suitability values (MAXENT output) lower than that associated with the lowest-ranking training locality. Omission rates greater than the expectation of zero typically indicate model overfitting. Values of OR_{MTP} range from 0 to 1; progressively more overfit models have increasingly higher values of OR_{MTP} (Radosavljevic & Anderson 2014).

For each species, we selected the set of model predictions that corresponded to ΔAICc values of ≤ 2. We used the relative AICc weights of these models to average Maxent raw output values, test AUC values, and OR_{MTP} across models. We normalized the AICc-averaged model predictions such that values were scaled from 0 to 1 for each species and represent a relative measure of habitat suitability in the study region. We did this separately for each species by subtracting the minimum suitability value from each pixel and then dividing those values by the range (maximum – minimum) of suitability values. We used these 'AICc-averaged' predictions for each species for the main analyses of this study. Additional summary statistics for our models are presented in Table S1.

Table S1. Summary of evaluation statistics for AICc-averaged ecological niche models generated for 173 species of trees in forests of Puerto Rico.

	Mean test AUC	**Mean OR_{MTP}**	**Number of occurrence records**
Mean	0.793	0.060	54.3
SD	0.083	0.056	38.9
Minimum	0.601	0.000	11
Maximum	0.962	0.325	192