THE ROLE OF EVOLUTIONARY HISTORY IN EXPLAINING THE VARIATION IN ABUNDANCE AND DISTRIBUTION OF PLANT SPECIES

by

John Robert Paul

B.S., The Evergreen State College, 1994

M.S., University of Florida, 2001

Submitted to the Graduate Faculty of

Arts and Sciences in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2008

UNIVERSITY OF PITTSBURGH

FACULTY OF ARTS AND SCIENCES

This dissertation was presented

by

John Robert Paul

It was defended on

May 8, 2008

and approved by

Dr. Anthony Bledsoe

Dr. Walter P. Carson

Dr. Susan Kalisz

Dr. Cynthia Morton

Dr. Charlotte M. Taylor

Dissertation Advisor: Dr. Stephen J. Tonsor

THE ROLE OF EVOLUTIONARY HISTORY IN EXPLAINING THE VARIATION

IN ABUNDANCE AND DISTRIBUTION OF PLANT SPECIES

John Robert Paul, PhD

University of Pittsburgh, 2008

A key observation from natural communities is that different species vary widely in their abundance and distribution. Understanding what factors are most important in explaining this variation is a fundamental goal of ecology. Here I take a comparative phylogenetic approach to address this problem. Using two clades of diverse tropical understory plants, I use information garnered from species' evolutionary relationships to test hypotheses about why some species are common while other species are rare. In a study of geographic range size variation of Neotropical *Piper* (Piperaceae) species, I used published DNA sequences to infer species' divergence times and herbarium collection records to infer their range sizes. I found that younger species have significantly smaller range sizes than older species. I examined a similar question using Mesoamerican *Psychotria* subgenus *Psychotria* (Rubiaceae) species. To infer the evolutionary relationships of species, I sequenced DNA from two loci of > 60 species in this clade. I concurrently inferred the phylogenetic relationships and absolute divergence times of species using a Bayesian relaxed-molecular clock method. I calculated two metrics of geographic range size using herbarium collection records, and predicted species' potential ranges using species distribution modeling. I found that Mesoamerican *Psychotria* subgenus *Psychotria* species have diversified primarily over the past 17 million years (Mya), and species largely fall into two clades that diverged approximately 15 Mya. In one clade, younger species have colonized a significantly smaller proportion of their potential range extent than older species.

Finally, using two genera in the clade Psychotrieae (Rubiaceae), I examined the impact of phylogenetic relatedness on the co-occurrence and variation in abundance among these species in Costa Rica, Central America. Using data collected on 240 transects nested in seven assemblages across Costa Rica and a phylogenetic hypothesis of species relationships based on DNA sequences, I found that Psychotrieae assemblages are significantly phylogenetically overdispersed, indicating that co-occurring species are less related than expected by chance. Within one heavily sampled assemblage, I found an inverse relationship between species' phylogenetic relatedness and their variation in abundance. The opposite trend was found across assemblages, where phylogenetic relatedness and variation in abundance were positively correlated.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

PREFACE

My dissertation greatly benefited from my interactions with many people whom I would like to thank. I sincerely thank all my committee members for their help and advice throughout my graduate career. I thank Tony Bledsoe for his constant enthusiasm, encouragement and excellent critiques. Tony also served as my *de facto* teaching mentor throughout my time at Pitt and my teaching skills have greatly benefited by working with him. I also had many great times birding with Tony, and I am a better birder because of it. I thank Walter Carson for giving me my first job as a field assistant in the tropics, and for his insightful and challenging comments on all aspects of my work. I also thank him for inviting me to be an active member of his lab community. The early development of my ideas greatly benefited from the marathon Carson lab meetings. I thank Susan Kalisz for her constant encouragement and excellent comments on my work. She also allowed me to conduct work in her molecular lab and use her resources that were a great benefit to my genetic work. I thank Cynthia Morton for welcoming me into her molecular lab at Carnegie Museum, and providing constant advice, encouragement, and critique on my molecular lab work. Cynthia invested a great amount of time and money into my work, and the molecular aspect of my work would have been impossible without her generous contributions. I thank Charlotte Taylor for sharing her knowledge and enthusiasm for all things *Psychotria*. She welcomed me into the Missouri Botanical Garden for two week-long research

the world around him. He has been a great inspiration and I dedicate my dissertation to him, with love.

1.0   INTRODUCTION

Understanding what factors drive variation in the distribution and abundance of organisms is a central focus of ecology. The last few decades have seen ecologists focus on understanding how biotic interactions in local communities regulate the variation in abundance and distribution of species. MacArthur's (1958) pioneering work on competition, Huchinson's (1957, 1959) work on niches, Paine's (1966) work on keystone predation, and MacArthur and Levins' (1964, 1967) work on the impact of competition on coexistence are just a few of the influential papers that inspired the next generation of researchers to determine to what extent species interactions influence emergent properties of communities, like species richness and diversity, and by what mechanisms different species come to dominate communities (e.g., resource competition, Tilman 1982). Recently, there has been a renewed interest in incorporating a historical evolutionary perspective into our understanding of ecological patterns (Wiens and Donaghue 2004), a perspective that was formerly popular early in the 20th century (Ricklefs 2004). Concurrently, there has been increased emphasis on integrating regional and local processes (e.g., Brown 1995) for a more complete picture of how communities are assembled, and why some species are found at high densities and in many places, while others are numerically rare and geographically restricted in their distributions.

A major contributing theory to this integration has been Stephen Hubbell's neutral theory in ecology (2001a), a controversial hypothesis that has simultaneously annoyed many ecologists

(e.g., Abrams 2001) while enlivening the debate regarding the importance of species traits and ecological niches in regulating patterns of diversity and commonness and rarity. Hubbell's hypothesis was built on two frameworks: island biogeography and neutral theory in population genetics. Island biogeography theory (MacArthur and Wilson 1967) has been one of the most influential and successfully predictive ideas in modern ecology. In essence, the theory states that the number of species on islands can be predicted based on the size of the island and the distance of the island from a mainland source of immigrants. Larger islands and islands closer to the mainland have more species, because they experience higher rates of immigration and lower rates of extinction. The theory has been applied more broadly than just to island situations and provides an important framework for understanding patterns of diversity in natural communities (Hubbell 2001a). One of the remarkable aspects of island biogeography is that the individual identities of species, in other words, everything about a species that makes it unique, is irrelevant to the model. Rather, rates of immigration, emigration, and extinction are the fundamental processes driving patterns of diversity.

Over roughly the same time period, neutral theory in population genetics was under development (Kimura 1968). Neutral theory essentially proposes that most mutations are neutral in their effect, and the changes in abundance of a given mutant allele are largely dependent on stochastic sampling over time (genetic drift). Hubbell (1979) introduced the idea that species might act much like alleles, drifting in their abundance through time. He recognized that the dynamics of species diversity can be predicted by largely ignoring species traits, as in island biogeography. He refined the theory over time (Hubbell and Foster 1986) and it eventually culminated in a book published in 2001 (Hubbell 2001), the same year that I started this dissertation. Hubbell's formal theory has spawned a cottage industry of publications attempting

test and claiming to refute his model (see McGill et al. 2006). In the larger picture, however, Hubbell has done a great service to ecology by reintroducing the importance of processes that occur over long time periods and large spatial scales to explain patterns of diversity (Ricklefs 2004, Lavin et al. 2004). Specifically, speciation and extinction are central to Hubbell's model, and knowing the rates of these parameters can be highly informative to predicting patterns of species richness and the shape of species-abundance curves in natural communities. While most of the controversy surrounding Hubbell's model has focused on his assertion that species are largely ecologically similar and do not show variation in their fitness due to the traits they possess, his theory provides a useful null model for how neutral dynamics would structure communities in the absence (or equivalence) of other limiting factors. Furthermore, the model makes testable predictions about how the evolutionary history of species should impact their abundances and distributions. Specifically, neutral theory echoes a prediction of a much earlier hypothesis, the Age-and-Area hypothesis, by John Willis (1922), that predicts that many species with restricted geographic ranges may simply be young species.

The past two decades have also seen another major player enter the ecological stage: phylogenetics. Inferring species' phylogenetic relatedness has become easier with new approaches (e.g., Drummond et al. 2006), computational advances (e.g., Britton et al. 2007, Zwickl 2006), and a huge influx of DNA sequence data that provides numerous characters on which to base phylogenetic inference, as well as the opportunity to develop statistical models about how DNA sequences change over time. As result, phylogenetics has gone through a renaissance, going from being largely equated with a specific division of systematics to now playing a major role in almost all aspects of evolutionary biology and more and more aspects of ecology (Webb et al. 2002). One of the most exciting applications to ecology is in

3

understanding how the phylogenetic relatedness of organisms influences the assembly of communities and ultimately the structuring of species interactions and patterns of co-occurrence and abundance (Webb et al. 2002).

In this dissertation, I investigate how phylogenetic information can inform our understanding of variation in the distribution and abundance of plant species. Specifically, I investigate the power of species age to predict range sizes in two tropical shrub genera (*Piper*, Piperceae, and *Psychotria*, Rubiaceae). In addition, I use field estimates of the co-occurrence and variation in abundance of Psychotrieae species (*Psychotria* and *Palicourea*) and their phylogenetic relatedness to investigate if assemblages of these species are phylogenetically structured differently from a random expectation, and if such structuring impacts the variation in abundance among species.

In Chapter 2, I use published DNA sequence data and herbarium collection records from Neotropical species of the diverse tropical understory shrub genus *Piper* (Piperaceae) to investigate the impact of species age on the geographic range sizes of species. I infer the phylogenetic relatedness of species using their internal-transcribed spacer (*ITS*) sequences, and infer their relative divergence times using Bayesian relaxed-clock dating. I find that species age, as determined by the divergence time of species, can be a significant predictor of range size variation, with young species having significantly smaller ranges than older species. Furthermore, I discuss the potential limitations of using molecular sequence data to infer species ages and special considerations when analyzing species' ranges sizes. This chapter was done in collaboration with Dr. Stephen Tonsor (University of Pittsburgh) and has been published as a chapter in the book *Tropical Forest Community Ecology* (Carson and Schnitzer 2008).

In Chapter 3, I conduct a more thorough investigation of the age and area hypothesis using Mesoamerican species of the diverse tropical shrub genus *Psychotria* subgenus *Psychotria* as a model system. In collaboration with Dr. Cynthia Morton (Carnegie Museum of Natural History), I generated DNA *ITS* sequence data and chloroplast *psbA-trnH* sequence data. I used both field collected material and leaf material from herbarium specimens on loan from the Missouri Botanical Garden for DNA extraction and sequencing. This allowed us to get sequence information from many rare and locally endemic species. I use both published *ITS* sequences and the *ITS* and *psbA-trnH* sequences I generated to infer the evolutionary relationships of these species and their absolute divergence times using Bayesian relaxed-clock dating. Furthermore, in collaboration with Dr. Charlotte Taylor (Missouri Botanical Garden, MBG), I used the MBG collection records to calculate two range size metrics, a measure of range occupancy and a measure of range extent, for the species in this study. Using these collections data and species distribution modeling, I infer the potential range occupancies and range extents of species to ask if younger species have filled less of their potential range occupancies and colonized less of their potential range extents. I find that *Psychotria* subgenus P*sychotria* species in Mesoamerica can be divided into two clades and that these clades have different relationships regarding species age and range size. I find that species age can be a significant predictor of range size variation in one clade, with younger species colonizing less of their potential range extents that older species. This chapter is co-authored with Dr. Morton, Dr. Taylor, and Dr. Tonsor, and will be submitted to *The American Naturalist*.

In Chapter 4, I investigate the role that phylogenetic relatedness plays in regulating the co-occurrence and abundance of Neotropical species in the diverse clade Psychotrieae (Rubiaceae), focusing on the ecologically similar species found in two genera, *Psychotria* and

Palicourea. I use data from 240 transects nested in seven geographic locations that I surveyed in 2003 and 2005 in Costa Rica to document patterns of co-occurrence and abundance. Using my *ITS* sequence data from the previous chapter, combined with all the available published Psychotrieae *ITS* sequences available on GenBank, I infer the evolutionary relationships of over 300 species using maximum likelihood and scale this tree by time using a new, fast method of generating ultrametric trees (Britton et al. 2007). I then assess if species co-occurring on the surveyed transects are more or less related than would be expected by chance. In addition, I examine the relationship between species richness, variation in abundance, and phylogenetic structure at a number of different spatial scales. I find that Psychotrieae species across all transects are phylogenetically overdispersed, indicating that they are less related than expected by chance. In addition, I find that the average phylogenetic structure estimates of assemblages from the seven geographic locations are significantly different from one another, with some assemblages phylogenetically overdispersed and some assemblages phylogenetically clustered (more related than expected by chance). Finally, I find that within one assemblage for which I had the largest sample size, the variation in abundance among species decreases with increasing relatedness, and that the average abundance of species decreases with relatedness. These results indicate that when closely-related species of Psychotrieae co-occur their abundances are more similar to one another, yet their abundances are depressed compared to their assemblage-wide average abundances.

Finally, in Chapter 5, I briefly discuss the implications and overall significance of my dissertation research, and suggest future avenues for research that integrate evolutionary history and ecology.

2.0   EXPLAINING GEOGRAPHIC RANGE SIZE BY SPECIES AGE: A TEST USING NEOTROPICAL *PIPER* SPECIES

2.1   ABSTRACT

Tropical plant species vary dramatically in their geographic range sizes. Theory predicts that narrowly endemic species may simply be young species that have not had sufficient time to expand their ranges. If two assumptions are met, namely that new species start with small range sizes and that the probability of extinction is inversely related to range size, then older species should, on average, have larger range sizes than younger species. This conjecture, originally formulated by John Willis as the Age-and-Area Hypothesis, and recently predicted by models of neutral community dynamics, has not been adequately tested in tropical plant taxa. To test this hypothesis, I focus on Neotropical species of the tropical understory shrub genus *Piper* (Piperaceae). I use published internal-transcribed spacer (*ITS*) sequences to infer species' divergence times using Bayesian relaxed-clock methods and herbarium records to estimate range sizes. I ask if there is a positive relationship between species age and range size. Using linear regression, I find that relative species age significantly explains a quarter of the variance in range size among species in this prominent tropical plant genus. This result confirms that species age can be a significant predictor of range size, and is notable in light of uncertainties in divergence time estimation using limited sequence data and incomplete sampling. I discuss the generality of

the results to other tropical plant taxa and briefly review the limited data on species-level age estimates from tropical plants. Furthermore, I discuss the potential limitations and difficulties of using divergence times as proxies for species ages, particularly when applied to analyses involving range and population sizes of new species. I suggest that the wealth of new genetic and biogeographic data on tropical plant species promise broader explorations of the impact of species age on species' range sizes in the near future.

2.2   INTRODUCTION

What accounts for rarity and endemism? Ecology, the study of distribution and abundance of species, remains without a coherent and consistent answer to this question. In tropical forest communities, the vast majority of species have few individuals and small geographic ranges (Dobzhansky 1950, Hubbell 2001a, Wallace 1878). Explaining how rare species differ from more common species, and elucidating the relative importance of various factors that regulate species' abundance and distribution is a central goal of ecology. However, the complicating influence of both deterministic and stochastic forces acting at various levels of biological organization and temporal duration make this a difficult task. In this chapter, I concentrate on the role of evolutionary history in structuring the abundance and distribution of plant species in tropical forests. Specifically, I address how the age of species can help explain patterns of rarity and endemism.

The potential importance of species age as a predictor of range size was first championed by Willis (1922). His "age-and-area hypothesis" asserted that, on average, older species will have larger ranges than younger species. He drew much of his evidence from studies of the

tropical flora of Ceylon (now Sri Lanka) where he observed that putatively ancestral species were more widely distributed than derived forms. Willis published a number of papers on the subject, and his ideas were subsequently debated and, in some cases, even ridiculed (e.g., Fernald 1924, Gleason 1924). In time, Willis' hypothesis failed to gain support (Stebbins & Major 1965) and his most lasting influence may actually have been in phylogenetics, via Yule's (1925) seminal paper that mathematically derived a model of a pure-birth speciation process, using Willis' ideas as the theoretical foundation.

Recently, the potential effects of historical processes on the distribution and abundance of organisms has received renewed attention (e.g., Ricklefs 2004, Wiens and Donoghue 2004). Much of this interest has been driven by two factors: the influx of molecular data on organisms that provide the potential to age the divergence dates of species, and the publication of Hubbell's *Unified Neutral Theory of Biodiversity and Biogeography* (2001a), which incorporates the large-scale, long-term effects of speciation and extinction on the abundance and distribution of species. Hubbell's neutral theory also specifically predicts that most rare, endemic species will be young species, while most wide-ranging species will be old (Hubbell 2001a,b); in effect, Hubbell's model makes a prediction similar to Willis' hypothesis. This prediction can be viewed as a general expectation, rather than a prediction specific to Hubbell's model. A positive relationship between species age and range size can be expected if two assumptions are met: 1) Species start with small population and range sizes, and 2) Extinction risk is inversely proportional to population and/or range size. Under these assumptions, new (young) species will have small population and range sizes and will face a high probability of extinction, while species that do persist and increase in size will face a decreasing probability of extinction. As a result, on average, young species are expected to be narrowly-endemic species, while wide-ranging species

are expected to be old. Interestingly, some of the strongest criticism of Hubbell's neutral model has focused on the expected age of common species. Specifically, if common species reach high abundance via ecological drift, the expected age of these species is unrealistically old, because of the slow pace of drift (Leigh 1999, Nee 2005, Ricklefs 2003). In contrast, if fitness deviations are accepted in the model, species can reach high abundance or go extinct much more quickly (e.g., Fuentes 2004, Yu et al. 1998). As a result, a positive age and range size relationship may be expected to persist much longer in clades that have been primarily driven by neutral processes than in clades where selection has driven species with high relative fitness' to occupy large ranges.

Of course, the relationship between species age and range size may take many forms, and Willis' age-and-area hypothesis (1922) is only one of several models of post-speciation range size transformation. For example, Gaston and colleagues (Chown 1997, Chown & Gaston 2000, Gaston 1998, Gaston 2003) have summarized a series of models of post-speciation range-size transformations (e.g., cyclical, random, stasis, etc.) that could potentially better explain the age and area relationships of some species. For example, the age and area relationship may be explained by a hump-shaped curve, where species start with small range sizes, reach their maximum range size at an intermediate age, and then decline towards extinction when they are old. Such a pattern was found for the proportion of fossil assemblages occupied by Cenozoic mollusks (Foote et al. 2007). Because there are a variety of processes that can expand or reduce species' ranges, individual clades may have their own unique age-area relationships. Thus, the utility of species age as a broad explanatory variable remains to be seen. In this chapter, I briefly review the few empirical tests of age and area and present an analysis using a clade of tropical understory shrubs (*Piper*). I discuss how the species age and range size relationship can be

viewed more broadly than the simple hypothesis presented by Willis (1922) and how this can lead to new hypotheses and understandings of the impact of historical processes on the current distribution and abundance of species.

2.3   EMPIRICAL TESTS OF AGE AND AREA

There have been few explicit tests of the age-and-area hypothesis. Two studies of marine fossil fauna have found evidence that indirectly support a positive age and area relationship. Jablonski (1987) documented a positive relationship between age (species duration) and geographic range size in the beginning of fossil mollusk's species' lifetimes, followed by long periods of stasis, but the focus of this study was on the possibility of species-level selection, rather than testing age and area *per se*. Similarly, Miller (1997) found that in Ordovician marine genera, older genera had larger ranges. Studying birds, Gaston and Blackburn (1997) found that for the entire New World avifauna, there was no relationship between mean range size of a clade and clade age, but there was a weak positive relationship between evolutionary age and total clade range size. In another study, Webb & Gaston (2000) examined six clades of birds and found various forms of the age and range-size relationship. Overall, roughly 20-50% of the variance in range size could be accounted for by species age (inferred from standard mitochondrial DNA molecular clock divergence estimates of 2% divergence per million years, Mya), but only one clade showed a positive age-and-area relationship; three showed a negative relationship and two a humped-shaped relationship. A study on *Sylvia* warblers found a weakly significant positive relationship between breeding range size and species age, but in this study the relationship could be better explained by older species generally having better dispersal

11

abilities than younger species (Böhning-Gaese et al. 2006). Finally, Jones et al. (2005) analyzed large molecular datasets of primates and carnivores and found evidence of a weakly negative age and area relationship (see this paper for a more detailed overview of Willis' age-and-area hypothesis and approaches to testing it).

Overall, a convincing positive age and area relationship predicted by Willis is not supported by these empirical data. However, a careful look at the published data reveals two trends. First, analyses that use fossil samples and measures of species duration as a proxy for age tend to find some evidence for a significant age and area relationship (e.g., Jablonski 1987, Miller 1997). In contrast, studies that examine extant species using molecular divergence dates as a proxy for age generally tend to find either no significant relationship between species age and range size, or a mixture of positive and negative relationships (see Table 7.1 in Jones et al. 2005). This discrepancy may be due, at least partially, to the different sampling methods. For example, a species' fossil record potentially allows sampling along the entire history of a species' range size trajectory over time (Fig. 2.1a). This is the ideal situation, in which the range size for a given species can be estimated at multiple ages. In contrast, molecular dating methods generally permit a single snapshot of a species' age and range size at a given point in time, and by looking at multiple species, I can infer the general trend of the age and area relationship for a group of organisms. Only having snapshots of a species age and range size relationship can introduce considerable variance into the relationship, particularly if all species follow varying range-transformation trajectories over time (even if the general shape of the relationship is similar, e.g., hump-shaped, Fig. 2.1b). However, it is likely that the majority of future age estimates for most taxa will be derived through molecular-based inference; thus, understanding

12

how these measures can potentially bias relationships such as age and area is critical to robust interpretation of results.

In addition to the potential discrepancies introduced through fossil versus molecular analysis of age and area, studies on extant species suggest that the phylogenetic level of the analysis is important. In studies of large clades containing many well-defined and potentially divergent subgroups (e.g., mammals, carnivores, or birds), general analyses of age and area find no or weak relationships (Jones et al. 2005), while studies of individual clades within these broad groups often find significant, but inconsistent relationships (e.g., the six clades of birds studied by Webb and Gaston 1998). This discrepancy suggests that the signal of an age and area relationship may be obscured when clades with distinct evolutionary histories are combined.

2.4   AN EMPIRICAL TEST USING A TROPICAL PLANT GENUS

Willis developed the age-and-area hypothesis thinking about tropical floras, and even his critics acknowledged that the hypothesis might be more important in the tropics (Gleason 1924), which were seen as stable and relatively homogenous. Despite this early attention to the tropics, to my knowledge, there have been no explicit tests of the hypothesis using tropical plants. The immense diversity of tropical plant species is only beginning to receive a genetic treatment, and my estimates of species' range sizes are imperfect, but slowly improving (e.g., Pitman et al. 2001). Most of the molecular dating of tropical plants to date has been conducted at higher phylogenetic levels; typically these studies are concerned with the general age of families and genera, and inferring when and where these groups of species diversified (e.g., Davis et al. 2005,

Muellner et al. 2006, Zerega et al. 2005).   In contrast, analyses of age and area require species-level resolution to properly address the hypothesis.

Here I examine the relationship between relative species age and range size in the diverse shrub genus *Piper* (Piperaceae) using publicly available internal-transcribed spacer (*ITS*) sequences from GenBank (www.ncbi.nlm.nih.gov; Appendix A). Most of these sequences were originally published in Jaramillo & Manos 2001, Jaramillo & Callejas 2004a, and Jaramillo & Callejas 2004b. I chose *Piper* because its species are prominent and important members of many rain forest communities throughout the world (Jaramillo & Manos 2001, Marquis 2004), there is a reasonably large amount of species-level informative genetic data available, and this taxon is an ideal model system for the study of ecology and evolution (Dyer & Palmer 2004). I focused my analysis on Neotropical species because many sequences were available for these species, the biogeography of Neotropical species has been studied (Marquis 2004, Quijano-Abril et al. 2006), and the range sizes of many species could be estimated using data from the Missouri Botanical Garden's online database, W^{3}Tropicos (http://www.tropicos.org./).

I used Bayesian inference to infer a phylogenetic tree, and then used this tree topology to estimate relative divergence dates among the species using the program BEAST (Drummond & Rambaut 2003), which uses a Bayesian relaxed clock approach to divergence time estimation (Drummond et al. 2006). For the phylogenetic inference, I aligned 113 sequences from 101 *Piper* (and *Macropiper*) species and five outgroup species using ClustalW (Thompson et al. 1994), followed by manual corrections. I used ModelTest (Posada & Crandall 1998) to evaluate the most appropriate model of molecular evolution for the analysis, which was determined by AIC model selection to be the general time reversible model with gamma distributed rates and a proportion of invariable sites (GTR + I + G). I ran the analysis in MrBayes 3.1.1 (Ronquist &

14

Huelsenbeck 2003), using model specifications for the GTR+I+G model, with a Dirichlet prior on substitution rates and state frequencies, and an unconstrained, exponential prior distribution on branch lengths. All analyses with MrBayes used two concurrent runs, each with four Markov Chain Monte Carlo (MCMC) chains (one "cold" and three "heated" chains). I examined an initial run of two million generations of MCMC simulations to assess if the chain had reached a stable distribution. Although the -log likelihood values stabilized by approximately 200,000 generations, clade probabilities failed to stabilize until nearly 1.5 million generations (assessed using the program "Are We There Yet?", Wilgenbusch et al. 2004). As a result, I ran a second analysis for five million generations, discarding the initial two million generations as burnin. This analysis effectively sampled from a stable distribution (with samples taken every 100 generations), resulting in a total of 60,000 trees after combining the two runs, from which a majority rule consensus tree was derived (Fig 2.2). This tree recovered the major clades described for *Piper* in previous work on ITS sequences (Jaramillo and Callejas 2004b).

I then used the topology of this phylogenetic tree as the input tree for the relative age analysis in BEAST. I held the topology of the tree constant for the analysis and fixed the mean substitution rate to one. BEAST uses MCMC sampling to assess branch lengths and divergence times by varying substitution parameters and the rate distribution based on a model of molecular evolution (I used the GTR+I+G). A preliminary analysis running for two million generations did not stabilize and the effective sample sizes of many parameters were low. The analysis presented here ran for ten million generations, with the first four million discarded as burnin. The resulting samples (taken every 100 generations) showed a stable –log likelihood distribution and good effective sample sizes for all parameters. I assessed the posterior probability densities of ages (divergence times of two species subtending these nodes) for 47 nodes on the

15

phylogenetic tree (Fig. 2.2). The mean divergence time values of these nodes were used to determine the relative ages of the Neotropical *Piper* species for the age and area analysis (Table 1). Since BEAST analyses have a stochastic element, I also ran the same analysis two additional times. The results were nearly identical (e.g., correlation coefficients of nodes ages between runs were > 0.99) so only the first run results are presented here.

To estimate range sizes, I counted the number of 1° x 1° latitude-longitude squares occupied by geo-referenced herbarium records in W^{3}Tropicos. This is effectively an area of occurrence measure (Gaston 1994). A few species for which I determined the age did not have records in W^{3}Tropicos; most of these were species listed as endemic to Columbia in Trelease and Yuncker (1950). Therefore, I present the analysis excluding these species; however, I also provided generous range-size estimates for these species and ran the analyses including them; the results were nearly identical and are thus not included here. The distribution of ranges sizes I calculated for the species with W^{3}Tropicos records is presented in Figure 2.3. The distribution is characterized by a few species with large ranges sizes and a long tail of species with small ranges (< 10 of 1° x 1° latitude-longitude squares).

To assess the relationship between relative species age and range size, I used linear least-squares regression using SAS 8.2 (SAS Institute 2001). I log-transformed both the mean species' ages and range sizes of the 58 Neotropical *Piper* species for which I had data. I found a highly significant positive relationship (y = 0.9399x + 2.6143, P < 0.001) that explains 25% (R^2 = 0.252) of the variation in range size for these *Piper* species (Fig. 2.4). Thus, my analysis supports the simple, positive relationship between species age and range size predicted by the age-and-area hypothesis. The strength of this relationship is notable in light of the various factors that can potentially obscure a positive age and area relationship.

There are some important caveats to this initial analysis of age and area in a group of tropical plants. First, my ages were based on divergence times of *Piper* species. My analysis represents only about 5-10% of the approximately 700 (Jaramillo & Manos 2001) to 1150 (Quijano-Abril et al. 2006) Neotropical *Piper* species. Taxon sampling affects age estimates, because missing taxa would alter the estimated divergence times of species if they were included in the analysis (Linder et al. 2005). Missing taxa can lead to an overestimation of ages (Chown & Gaston 2000, Jones et al. 2005, Webb & Gaston 2000). However, given the strength of the positive age and area relationship that I found based on the *Piper* sequences available, and no reason to expect an inherent bias to the species that were selected to sequence or to the locations of missing taxa on the tree, I suspect the positive age and area relationship found here will be borne out in future analyses of larger datasets.

2.5 WHAT DO OTHER TROPICAL PLANT CLADES TELL US?

Aside from *Piper*, there are very few molecular datasets available for specific clades of tropical plants that can be effectively used to assess age and area relationships. Considerable molecular data have amassed recently on tropical plant lineages and their divergence dates, but most of these data examine higher phylogenetic levels (e.g., families or higher; Davis et al. 2005, Lavin et al. 2005, Renner et al. 2001) and have focused on the origin and age of the clades and species that make up current tropical communities. These data tell an interesting story, but do not yet provide any clear expectations for the generality of the kind of age and area relationship found for *Piper*.

Species-rich genera like *Piper* have a wide range of ages, based on the available evidence

from molecular dating. *Piper* is a member of the basal angiosperms (APG 2003), and may be a rather old lineage (based on *Piper* and *Peperomia* divergence, ~ 40 Ma, Wikström et al. 2001). In contrast, analysis of the diverse legume genus *Inga* suggests that it is a young genus and many species originated on the scale of 2-10 Ma (Richardson et al. 2001). In light of evidence of the existence of rainforests from the late or mid-Cretaceous (~100 Ma; Davis et al. 2005, Morely 2004), *Inga* species must be considered quite young (Bermingham & Dick 2001). Despite its relatively recent origin, this clade has spread throughout the forests of South and Central America, and at many sites *Inga* species are important forest components both in terms of number and biomass (Richardson et al. 2001). In fact, legume clades in general may be remarkably young given their widespread distribution and numerical importance in tropical forests (~ 4-16 Ma, Lavin et al. 2004). Other speciose tropical clades are considerably older, such as those in the Annocaceae (e.g., *Xylopia*, *Annona*) which appear to be on the scale of ~ 15-25 Ma (Pirie et al. 2006, Richardson et al. 2004). Like *Piper*, many of these clades have pantropical or even cosmopolitan distributions; in fact, one of the most widespread tropical plant species, *Symphonia globulifera* (Clusiaceae), also ages to the mid-Tertiary (~ 28 Ma, Dick et al. 2003). In Africa, the origin of the herbaceous Begonias (Begoniaceae), is also on the scale of ~ 30 Ma, but many of the species in this group diverged relatively recently (from ~ 1-10 Ma; Plana et al. 2004).

In another widespread herbaceous genus, *Costus* (Costaceae), the Neotropical species appear to have diversified rapidly and recently (Kay et al. 2005). In the case of very recent diversification of clades like *Inga* and *Costus*, widespread species within these genera provide evidence that common members of these clades are not particularly old. However, the relationships of age and area within these and other genera have not been assessed. In a rapidly

diversifying genus, if more widespread species were found to be older, the expected slope of the age and area relationship would simply be very steep. However, finding young but common species would certainly not be surprising in light of recent evidence confirming a rare species advantage in many tropical forests, probably resulting from lower density-dependent or frequency-dependent mortality (e.g., Harms et al. 2000, Volkov et al. 2005, Wills et al. 2006). Rare species that have a fitness advantage are expected to increase in abundance much more rapidly than predicted under neutral drift, for example, resulting in younger species that have large range and population sizes. Thus, if new species do indeed start with small population and range sizes, some of these species may be expected to increase their population and range sizes rapidly. Overall, the generality of a positive age and area relationship in tropical plant species awaits future analyses, particularly of densely sampled, speciose clades.

Fortunately, there is considerable promise that in the near future we can gain a broader perspective on age and area relationships in tropical plants. For example, work on the diverse tropical herbaceous genus *Begonia* (Begoniaceae) has provided insight into the phylogenetics and timing of diversification in this pantropical genus (e.g., Forrest & Hollingsworth 2003, Plana et al. 2004). Likewise, phylogenetic work on the diverse pantropical genus *Psychotria* (Rubiaceae; Nepokroeff et al. 1999, J. Paul, *unpublished data*) promises to provide evidence from a genus that in many ways mirrors *Piper* in its species' ecology, abundance, and distribution (e.g., high local and regional species richness, numerical abundance, understory and gap habitat, etc.), although it is phylogenetically distantly related. Interestingly, Hamilton (1989a) suggested that within the Mesoamerican members of *Psychotria* subgenus *Psychotria*, species groups often contained one basal member with a large geographic range, and putatively derived members with narrow ranges.

2.6   AN AGE-AND-AREA HYPOTHESIS FOR MODERN TIMES

The strong positive age and area relationship found for Neotropical *Piper* species warrants further investigation into the generality of this relationship in tropical plants. If, in general, many rare species are found to be young species, this information may be crucial to incorporate into our understanding of the variation in range size among species, and at the local scale, variation in abundance, which often shows a positive relationship with range size (Gaston 1994). In order to effectively integrate species age information derived from molecular inference (as most future data promises to be) into our understanding of tropical forest community structure, we need to recognize the potential sources of error in these data, as well as take a broader view on the simple age-and-area hypothesis proposed by Willis (1922).

First, one of the obvious shortcomings of the traditional age-and-area hypothesis (Willis 1922) is its failure to account for old species with small ranges. Empirical evidence suggests that in some cases, the age and area relationship may be a humped-shaped relationship (Webb & Gaston 2000), where both old and young species have small ranges, and intermediate age species have the largest ranges (or the greatest degree of ecological occupancy, e.g., Foote et al. 2007). Clearly, many old species must either go through range contraction as they age, or have their ranges sizes reduced through the process of speciation. As a result, a complete age-and-area hypothesis needs to account for these species, recognizing that a positive age and area relationship may be limited to the lower end of the temporal axis. For example, if the assumption that new species start with small ranges is accepted, then the general positive relationship between species age and range size can be expected to persist until some threshold, and then the relationship will become flat or negative, as older species lose range size. Almost all of the models of post-speciation range size transformation presented in Gaston (1998), for

example, have an initial phase in which there is a roughly linear positive relationship between species age and range size. The differences in these lines is the steepness of their slope and their temporal duration; some models, such as a cyclical and stasis models, predict a rapid increase in range size post-speciation, while the traditional age and area model is depicted as a gradual increase. However, depending on the total age of a clade of interest, and the rate at which transformations occur, all of these models are similar in their initial prediction of a positive species age and range size relationship. Thus, the more important question may be, when does a positive age and area relationship cease to exist, and why? Furthermore, analyses that examine clades of species and ask if on average rare species are younger than old species, rather than simply looking for a positive slope of an age and area relationship, may be more informative.

Second, the positive age and area expectation of most models of post-speciation transformation are primarily driven by the assumption that new species start with small population sizes. But do they? It has been asserted that much speciation in tropical woody plants arises through isolation of small local populations (e.g., Ehrendorfer 1982, Leigh et al. 2004), but strong empirical evidence to support this position is generally lacking. Since the population sizes of new species cannot practically be measured, inference must be used to estimate the sizes of ranges and populations. For example, fossil evidence supports African large-mammal populations starting as small, narrowly-ranging populations (Vrba and DeGusta 2004). Unfortunately, the sparse fossil record for many taxa, particularly plants in the tropics, makes inference based on fossil evidence rare. The data presented here for *Piper* are certainly suggestive that newer species have small range sizes, as evidenced by the preponderance of young species with small range sizes and the lack of young species with large ones. Future analyses of age and area relationships in tropical plants may help to fill in the gaps of our

21

knowledge of new species population and range sizes that are unlikely to be filled by fossil evidence.

Third, a practical difficulty arises from using divergence times of species as proxies for ages. When speciation is defined as a cladogenic (splitting) event, such as on a dichotomously-branching phylogenetic tree, any speciation event yields at least two new species, both assigned the same age. These new species have range and population sizes defined by the boundaries of their newly isolated gene pools (or lineages). Thus, when speciation is viewed as a splitting process with a geographical component, new species will often have smaller range and population sizes than their direct ancestor, because the ancestral range (and the distribution of individuals defining it) is subdivided. If the relative range and population sizes of sister species are markedly skewed, there will be considerable variance in the distribution of population sizes of the new species. For example, when a new species (B) is introduced via a point-mutation model of speciation (where one individual is assigned a new species status based on some new defining character, *sensu* Hubbell 2001a), its ancestor species (A) with population size N must also be deemed a new species (C), with a population size N - 1. Since species B and C are assigned the same age, the youngest species in the community are represented by species with both small (B) and large (C) population and range sizes. In other words, when a widespread species gives rise to a narrowly-endemic sister species, but the widespread species persists essentially unchanged in its ecological and genetic attributes, both sister species are assigned the same age. This is potentially at odds with the meaning of species age in an evolutionary sense. It also clearly creates difficulty in analyzing age and area, as such a process will obscure any expectation of a positive relationship if such asymmetric range splits are commonplace in a clade. In light of this potential source of noise in the age and area relationship, it is all the more

remarkable that a positive relationship explaining a good portion of the variance in range size was found in the analysis of *Piper* species.

Finally, molecular age estimates are potentially subject to many different kinds of errors and uncertainties (Arbogast et al. 2002). For example, the model of molecular evolution used, the degree of consensus between gene trees examined and true species trees (Nichols 2001), the reliability of any fossil ages used for calibration, and success of an analytical model dealing with rate heterogeneity all can introduce potential errors in estimates of ages (Renner 2005, Sanderson et al. 2004).

In summary, future studies on age and area relationships in tropical plants have the potential to provide insight into the role that the simple explanatory variable species age can play in explaining patterns of rarity and endemism. Of course, as Willis himself recognized, age by itself cannot be the mechanistic driver of these patterns we observe. Rather, age acts as a proxy for the playing-out of various ecological interactions at different spatial and temporal scales. If a positive age and area relationship is found for a group of taxa, this finding can point to valuable lines of research for future studies (Jones et al. 2005). For example, if such a relationship is found in a 20 Ma clade of plants, could this be an indication that the range-size transformations within in this group are rather slow and potentially governed by the ecological drift? If only certain guilds of plants (e.g., understory shrubs) show a positive age and area relationship, could this be related to the potential dispersal limitations imposed on these plants through their canopy position and reliable seed dispersers? In addition, how do clades that have many old species with small ranges differ from those clades like *Piper*, which apparently lack many old species with small ranges? An updated view of the age-and-area hypothesis thus allows researchers to inquire about much more than whether the age and area relationship in a given group of

organisms is linear and positive. The shape of the relationship in a given clade can be used to infer the importance of various factors in the range transformation of species, and suggest if new species start with small range sizes.

2.7   CONCLUSIONS

An explanation for why many tropical forests species are rare and endemic may simply be the relatively young age of these species. I reviewed the limited empirical work addressing age and area relationships, none of which came from strictly tropical taxa, and showed that support for the traditional age-and-area hypothesis is equivocal. Using Neotropical *Piper* species as a case study, I conducted the first age and area analysis for a tropical plant clade, and found significant support for a positive age and area relationship that explains a quarter of the variation in range size among species. Speculation about the age and area relationships within other taxonomic groups, however, is difficult because species-level data on either ages or ranges are sparse. Although inferring species ages from molecular data and phylogenetic trees can introduce difficulties when interpreting results of age and area analyses, I predict that in the near future broader analyses of age and area will be plausible with many clades of tropical plants.

Table 2.1. Relative ages estimated by Bayesian relaxed clock analysis showing the mean ages and standard deviations (S.D.), median ages, and highest posterior density distributions.

Species	Mean Age	S.D.	Median Age	95% HPD (lower)	95% HPD (upper)
Piper aduncum	1.0E-02	2.3E-04	9.6E-03	3.2E-03	2.0E-02
P. albozonatum	3.5E-03	6.6E-05	3.0E-03	2.3E-04	7.9E-03
P. amalago	7.1E-03	1.5E-04	6.3E-03	1.4E-03	1.5E-02
P. amoenum	1.0E-02	2.3E-04	9.6E-03	3.2E-03	2.0E-02
P. appendiculatum	1.2E-02	1.6E-04	1.1E-02	4.3E-03	2.0E-02
P. arboreum	3.2E-02	6.4E-04	3.1E-02	1.9E-02	4.7E-02
P. archeri	2.3E-02	3.7E-04	2.2E-02	1.0E-02	3.6E-02
P. arieianum	3.4E-02	4.7E-04	3.3E-02	2.2E-02	5.0E-02
P. augustum	2.2E-02	3.5E-04	2.2E-02	1.2E-02	3.4E-02
P. auritum	4.2E-02	5.9E-04	4.2E-02	2.6E-02	6.2E-02
P. bartlingianum	4.8E-02	9.3E-04	4.7E-02	2.3E-02	7.5E-02
P. basilobatum	3.9E-03	6.2E-05	3.5E-03	7.6E-04	7.6E-03
P. brachypodon	1.3E-02	1.6E-04	1.2E-02	4.4E-03	2.1E-02
P. brevipedicellatum	2.9E-03	5.4E-05	2.3E-03	4.5E-05	7.6E-03
P. cajambrense	8.6E-03	1.8E-04	8.3E-03	4.1E-03	1.4E-02
P. cararense	1.9E-02	3.6E-04	1.8E-02	7.2E-03	3.3E-02
P. cavendishioides	1.8E-02	3.5E-04	1.8E-02	1.0E-02	2.6E-02

Table 2.1. (continued)

Species	Mean Age	S.D.	Median Age	95% HPD (lower)	95% HPD (upper)
Piper chuarense	9.7E-03	2.3E-04	8.9E-03	3.4E-03	1.9E-02
Piper cihuatlanense	7.1E-03	1.5E-04	6.3E-03	1.4E-03	1.5E-02
P. cinereum	7.6E-02	1.2E-03	7.5E-02	4.6E-02	1.1E-01
P. cocornanum	1.8E-02	3.1E-04	1.8E-02	4.5E-03	3.3E-02
P. colligatispicum	1.8E-02	3.1E-04	1.8E-02	4.5E-03	3.3E-02
P. confertinodum	9.7E-03	2.3E-04	8.9E-03	3.4E-03	1.9E-02
P. darienense	1.7E-02	3.5E-04	1.5E-02	4.3E-03	3.2E-02
P. filistilum	7.1E-03	1.4E-04	6.4E-03	1.4E-03	1.4E-02
P. flagellicuspe	1.2E-02	1.6E-04	1.1E-02	4.3E-03	2.0E-02
P. friedrichsthalii	4.4E-02	5.4E-04	4.4E-02	3.1E-02	5.9E-02
P. garagaranum	2.4E-02	4.7E-04	2.4E-02	7.9E-03	3.9E-02
P. gesnerioides	1.9E-02	3.9E-04	1.8E-02	7.6E-03	3.3E-02
P. hartwegianum	8.2E-03	1.2E-04	7.7E-03	2.0E-03	1.5E-02
P. hispidum	3.7E-02	5.1E-04	3.6E-02	1.8E-02	5.6E-02
P. imperiale	1.6E-02	3.0E-04	1.5E-02	5.6E-03	3.0E-02
P. longispicum	1.9E-02	3.2E-04	1.8E-02	1.1E-02	2.9E-02
P. marequitense	3.9E-02	4.8E-04	3.9E-02	2.6E-02	5.2E-02
P. marginatum	2.9E-02	6.8E-04	2.8E-02	9.1E-03	5.0E-02
P. michelianum	2.9E-03	5.4E-05	2.3E-03	4.5E-05	7.6E-03
P. multiplinervium	2.9E-02	6.8E-04	2.8E-02	9.1E-03	5.0E-02

Table 2.1. (continued)

Species	Mean Age	S.D.	Median Age	95% HPD (lower)	95% HPD (upper)
P. munchanum	1.8E-02	3.2E-04	1.7E-02	8.5E-03	2.8E-02
P. obovatum	3.9E-03	6.2E-05	3.5E-03	7.6E-04	7.6E-03
P. ottoniifolium	1.6E-02	3.3E-04	1.6E-02	8.9E-03	2.5E-02
P. oxystachyum	1.3E-02	1.6E-04	1.2E-02	4.4E-03	2.1E-02
P. parvulum	1.2E-02	2.6E-04	1.1E-02	3.1E-03	2.2E-02
P. pedunculatum	5.0E-03	9.2E-05	4.5E-03	7.8E-04	1.0E-02
P. peltatum	2.2E-02	3.1E-04	2.1E-02	1.0E-02	3.4E-02
P. perpusillum	5.0E-03	9.2E-05	4.5E-03	7.8E-04	1.0E-02
P. phytolaccifolium	1.2E-02	2.4E-04	1.2E-02	5.0E-03	2.0E-02
P. pilibracteum	1.8E-02	2.0E-04	1.7E-02	9.3E-03	2.7E-02
P. pulchrum	1.6E-02	3.0E-04	1.5E-02	5.6E-03	3.0E-02
P. reticulatum	4.4E-02	7.4E-04	4.3E-02	2.2E-02	6.9E-02
P. sabaletasanum	1.4E-02	2.4E-04	1.4E-02	7.6E-03	2.2E-02
P. schuppii	1.6E-02	3.3E-04	1.6E-02	8.9E-03	2.5E-02
P. sp1maj674	1.8E-03	3.4E-05	1.5E-03	3.1E-05	4.5E-03
P. sp2maj689	1.0E-02	2.0E-04	9.7E-03	4.7E-03	1.6E-02
P. spoliatum	7.4E-03	1.6E-04	7.1E-03	3.3E-03	1.2E-02
P. subglabribracteatum					
	1.3E-02	1.9E-04	1.3E-02	5.4E-03	2.2E-02
P. subpedale	2.4E-02	4.7E-04	2.4E-02	7.9E-03	3.9E-02

Table 2.1. (continued)

Species	Mean Age	S.D.	Median Age	95% HPD (lower)	95% HPD (upper)
P. terryae	7.1E-03	1.3E-04	6.0E-03	7.1E-04	1.7E-02
P. tomas-albertoi	8.2E-03	1.2E-04	7.7E-03	2.0E-03	1.5E-02
P. trianae	7.5E-03	1.4E-04	6.9E-03	1.8E-03	1.4E-02
P. tricuspe	1.2E-02	2.4E-04	1.2E-02	5.0E-03	2.0E-02
P. tuberculatum	3.2E-02	6.4E-04	3.1E-02	1.9E-02	4.7E-02
P. ubatubense	1.7E-02	3.5E-04	1.5E-02	4.3E-03	3.2E-02
P. umbellatum	2.2E-02	3.1E-04	2.1E-02	1.0E-02	3.4E-02
P. umbricola	1.2E-02	2.6E-04	1.1E-02	3.1E-03	2.2E-02
P. unispicatum	7.1E-03	1.3E-04	6.0E-03	7.1E-04	1.7E-02
P. villosum	2.1E-02	3.6E-04	2.0E-02	1.0E-02	3.4E-02
P. yanaconasense	1.8E-03	3.4E-05	1.5E-03	3.1E-05	4.5E-03
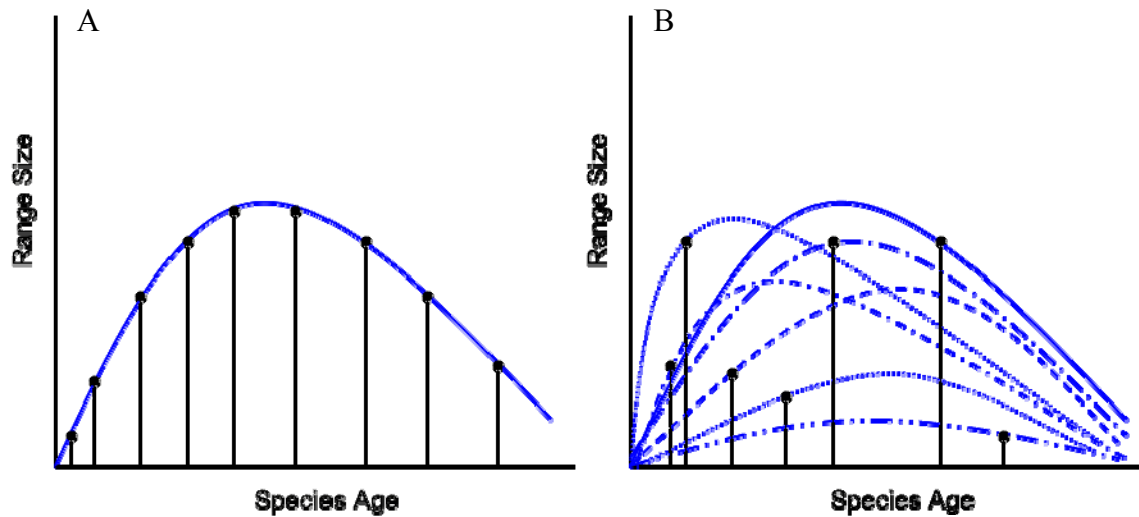
Figure 2.1. Graphical depictions of range size trajectories of species over time. Black dots indicate sampling points in time. Ideally, fossil analyses can allow the range size of a species to be assessed at multiple time points (A), effectively sampling over the life-span of a species. When using molecular estimates of ages, species can usually only be sampled at a single point in time (B). By sampling multiple species (different lines on the graph), a general relationship between species age and range size can be inferred. However, even if all species show roughly the same shape for an age and area relationship (e.g., humped-shaped), if they follow varying range transformation trajectories, sampling single points over time will introduce considerable variation into the species age and range size relationship and make inferring general trends more difficult.
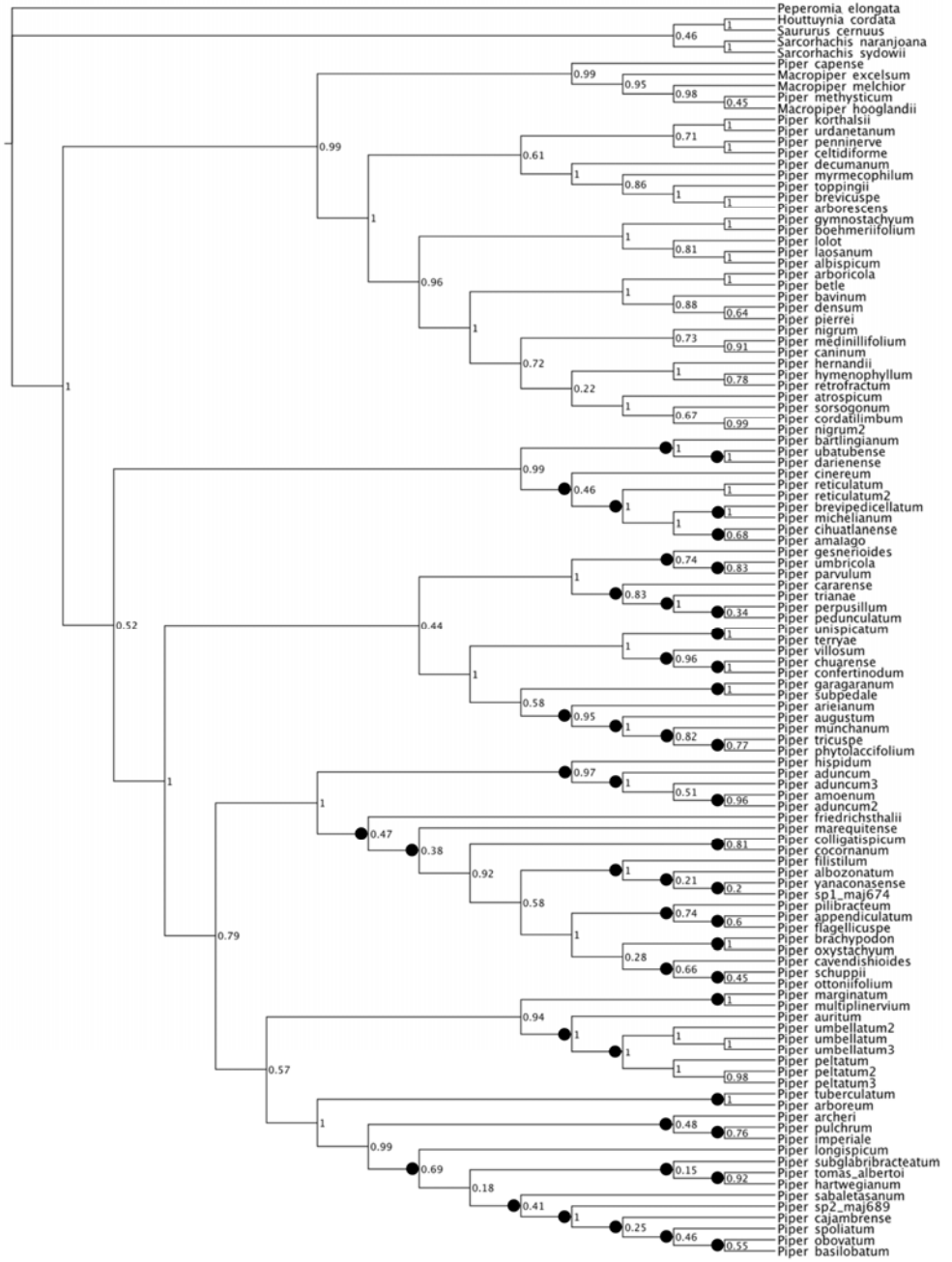
Figure 2.2. A phylogenetic hypothesis of *Piper* species relationships inferred by a Bayesian analysis of ITS sequences. Posterior probabilities of clades are shown at the nodes. Black dots depict the nodes for which relative ages were calculated in a separate Bayesian analysis in which this tree topology was used (see text for details).
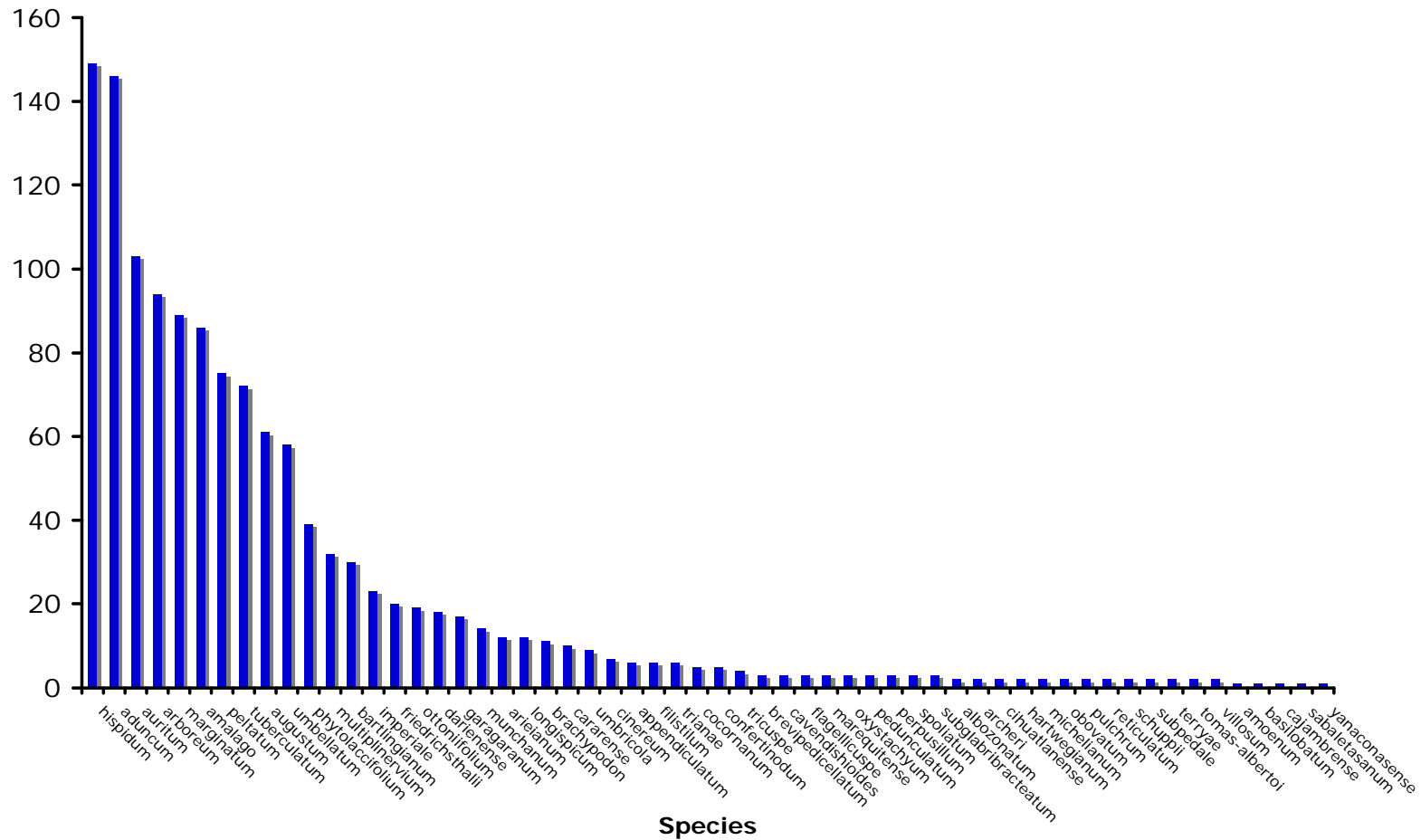
Figure 2.3. The distribution of ranges sizes (number of 1° x 1° latitude-longitude squares) of the Neotropical *Piper* species used in the analysis of age and area.
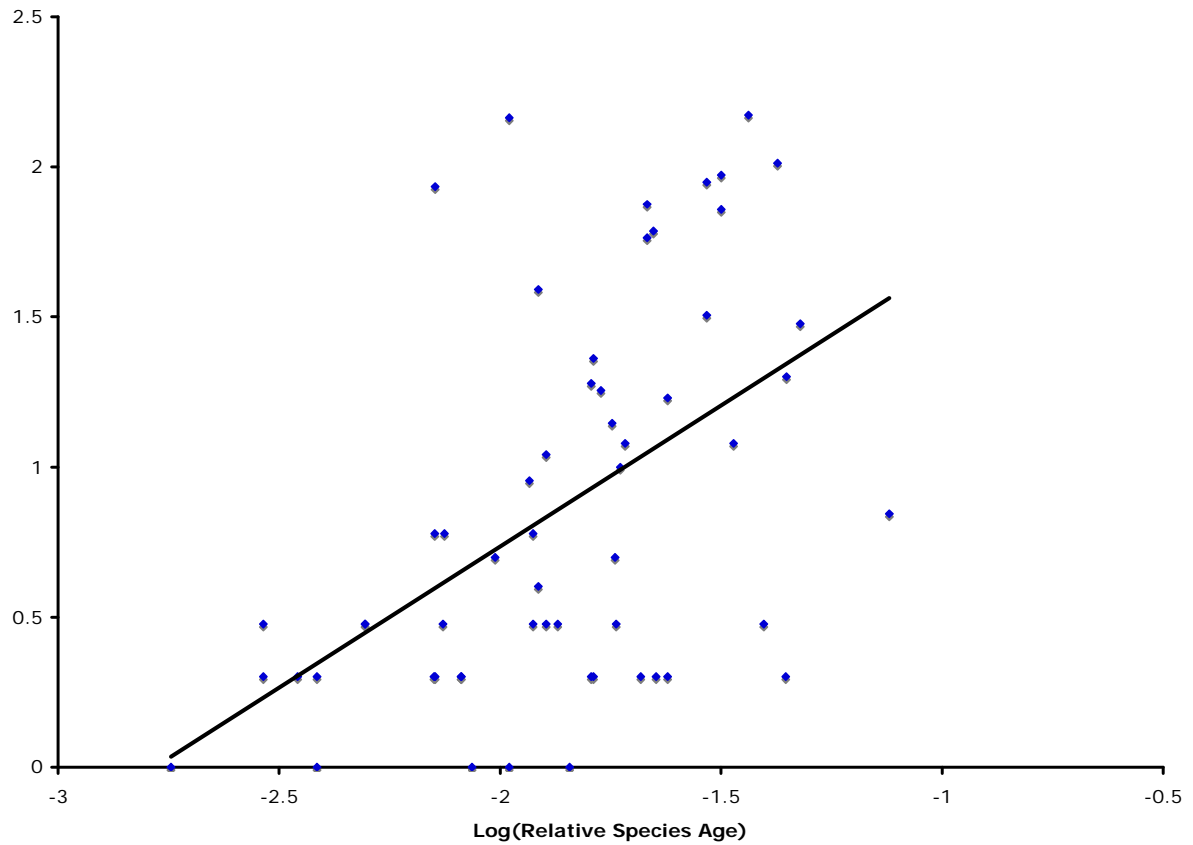
Figure 2.4. Linear regression of log-transformed relative species age and log-transformed range size; $y = 0.9399x + 2.6143$, $R^2 = 0.252$, $P < 0.001$.

3.0 EVOLUTIONARY TIME FOR DISPERSAL LIMITS THE EXTENT, BUT NOT THE OCCUPANCY OF SPECIES' POTENTIAL RANGES IN THE NEOTROPICAL PLANT GENUS *PSYCHOTRIA*

3.1 ABSTRACT

Explaining the diversity in geographic range sizes among species is a central goal of ecological and evolutionary studies. I tested species age as an explanation of range size variation among a closely related group of understory shrubs in Mesoamerican (*Psychotria* subgenus *Psychotria*, Rubiaceae). *Psychotria* species vary by orders of magnitude in geographic range size, yet species appear to be generally ecologically similar, bringing into question what drives variation in range size. I sequenced the internal-transcribed spacer (*ITS*) and chloroplast *psbA-trnH* loci of a large majority of the Mesoamerican species. I used Bayesian relaxed-clock dating to estimate phylogenetic relationships and species' ages. I measured species' geographic range occupancies and range extents using herbarium collection records. Range occupancy measures how much of a geographic range is filled, and range extent measures the maximum linear distance between collection records. I used species distribution modeling to predict species' potential ranges. If species range sizes are limited by time for dispersal, I hypothesized that older species should have 1) larger *realized range occupancies* and *realized range extents* than younger species, 2) filled a greater proportion of their *potential range occupancies*, and 3) colonized a greater

proportion of their *potential range extents*. I found 1) a significant but weak, positive relationship between species age vs. both realized range occupancy and realized range extent in *Psychotria*. Furthermore, I found 2) no relationship between species age and filling of potential range occupancies, but 3) older species had colonized a significantly greater proportion of their potential range extents than younger species. However, within *Psychotria,* species are nested in two strongly supported clades that diverged ~15 Mya. When analyzed separately, older species in one clade had colonized a significantly greater proportion of their potential range extents than younger species, explaining a third of the variance. Species age did not explain proportional range extent in the other clade, or occupancy of potential ranges in either clade. Despite the divergent evolutionary history of the clades, I found no significant differences in average geographic or elevation range attributes of species between clades and no differences in the phenotypic characteristics measured. However, younger species in the clade where species age was not predictive of proportional range extent had larger fruit volumes than older species, suggesting that larger frugivorous birds may enhance these species' dispersal. Our results indicate a time-for-dispersal effect may limit the extent of species' ranges, but not necessarily their occupancy.

3.2   INTRODUCTION

Even among closely related species, geographic range size can vary over many orders of magnitude (Brown et al. 1996, Gaston 2003). Understanding what factors best explain variation in geographic range size among species is a central question at the interface of ecology and evolution. Range expansions are driven by dispersal, so variation in dispersal ability has been

predicted to explain much of the variation in range sizes among species (Hanski et al. 1993, Gaston 2003), with a general expectation that species with superior dispersal abilities attain larger range sizes more quickly (Brown et al. 1996, Hanski et al. 1993). Despite the perceived importance of dispersal, only limited empirical evidence supports this conjecture, and a recent review even suggests that dispersal ability may not be particularly important in driving range size variation in many species (Lester et al. 2007). Dispersal ability, however, is only one side of the coin, since dispersal that expands a species' range is not an instantaneous process; the time available for dispersal can also play a central role in explaining range size variation. For example, even a species with very poor dispersal abilities may attain a large geographic range size, given sufficient time. Similarly, a group of species that shows little variation in dispersal ability may have drastically different range sizes, simply because the time that has been available for dispersal differs among the species. Hence, when attempting to explain the variation in range size among species, *time may be a critical limiting factor*, particularly if the species of interest show no obvious differences in their dispersal potentials.

Temporal dispersal limitation, although not explicitly stated as such, forms the underpinning of theory that predicts a positive relationship between species age and range size (e.g., Willis 1922). If species start with small population sizes and restricted geographic ranges, many species with restricted geographic ranges could simply be young species. This was one of the key predictions of John Willis' "Age and Area Hypothesis" (1922), and a similar prediction is made by Hubbell's neutral theory (2001), a dispersal-assembly theory. The premise is simple and built on three key assumptions: 1) new species have restricted geographic ranges, 2) species with small geographic ranges are extinction prone (and thus most young species never attain either older ages or larger ranges), while 3) species with large geographic ranges are buffered

from extinction (Johnson 1998, Payne and Finnegan 2007). When these assumptions are met, there is a general expectation that on average, young species will have smaller ranges than old species (Paul and Tonsor 2008). The majority of studies of the age and area relationship have simply tested if there is a positive linear relationship between some metric of species age and range size. These tests have been largely equivocal (Jones et al. 2005). Evidence for a positive relationship between species age and range size has been found for some mollusk species (Jablonski 1987, Miller 1997), but only early in their evolutionary history, after which ranges appear to stabilize in their size. In Cenozoic mollusks, species occupancy of fossil assemblages (the proportion of collections in which a species is present) shows a hump-shaped distribution, with species attaining their maximum occupancy for a brief time in the approximate middle of their species' lifetimes (Foote et al. 2007). Studies on large diverse groups of taxa, such as all New World bird species (Gaston and Blackburn 1997) or all mammals or carnivores (Jones et al. 2005) find no consistent relationship between species age and range size. These tests used taxa in which the species have diverse and broadly different ecological niches. When species with more similar ecological requirements have been compared, for example, in six clades of birds (Webb and Gaston 2000), the relationship between species age and range size is variable and clade-specific. In the case of the *Sylvia* warblers (Böhning-Gaese et al. 2006), age is a significant factor explaining variation in range size (although age was strongly correlated with dispersal ability). Likewise, Paul and Tonsor (2008) examined a genus of ecologically similar tropical plants (*Piper*) and found that species age explained 25% of the variation in range size in this group, with young species having smaller ranges than old species.

Two important components have been missing from previous tests of age and area. The first is the biological reality that all area outside a species' range is not actually habitable, due to

a species' specific physiological and ecological requirements. As a result, previous tests of age and area have not accounted for one major potential driver of range size variation, the ecological tolerances of species. Better tests would evaluate the area that a species could occupy, given its ecological constraints, termed its *potential range* (Gaston 1994, Gaston 2003), relative to the *realized range* (current, observed range). The ratio of the realized range (R) to the potential range (P), can be used to assess to what degree species' occupy their potential ranges ('range filling', Gaston 2003, Svenning and Skov 2004). Species distribution modeling (i.e., Elith et al. 2006) provides a method to estimate the potential range of a given species (in the absence of dispersal limitation), based on a set of biologically relevant variables and georeferenced records of presence localities.

The second component that has not been adequately addressed in previous studies is that both realized and potential range sizes can be measured in two general ways, as the *area of occupancy* (the number of locations with a presence record for a species) or as the *extent of occurrence* (the maximum linear distance between locations with a presence record for a species, Gaston 1994b). While these measures can be correlated, they can also be decoupled (Gaston 1994b). For many applications of range size data, such as studies regarding conservation biology, the area of occupancy is the preferred measure, as it gives a better idea of where specifically on the landscape a species is likely to be found. Only area of occupancy measures have been used for previous tests of age and area (e.g., Webb and Gaston 2000, Jones et al. 2005). However, if the predictions of age and area are viewed as a result of the process of temporal dispersal limitation, then an extent of occurrence measure may be more appropriate. The time available for dispersal could limit how far a species has colonized into its potential range, but have little impact on its occupancy within its range. For example, a species that is a

poor competitor, but has superior dispersal abilities could have a large range extent but only limited occupancy within its range. Interestingly, the only study incorporating species age as a factor in range filling (Schurr et al. 2007) found no evidence of an effect on the area of occupancy of potential ranges in South African Proteaceae species (extent of occurrence was not measured).

To address these two limitations of previous studies I developed range size metrics that specifically incorporate species' potential ranges, as well as *area of occupancy* and *extent of occurrence* measures. I define species' *realized range occupancy* (R_O) as the number of occupied locations (e.g., grid cells) and the *realized range extent* (R_E) as the maximum linear distance between the locations of records of occurrence. *Potential range occupancy* (P_O) and *potential range extent* (P_E) are defined the same way as for realized ranges, except modeled potential locations are used (Fig. 3.1). I define the degree to which species occupy their potential ranges as the ratio of *realized range occupancy* to *potential range occupancy* (*range occupancy ratio*, R_O/P_O), and the degree to which they have colonized their potential range extents as the ratio of *realized range extent* to *potential range extent* (*range extent ratio*, R_E/P_E).

I tested the general hypothesis of a positive relationship between species age and these four metrics of geographic range size: 1) realized range occupancy (R_O), 2) realized range extent (R_E), 3) range occupancy ratio (R_O/P_O), and 4) range extent ratio (R_E/P_E). I examined the impact of temporal dispersal limitation on range size variation in a clade of closely related, ecologically similar species of tropical understory shrubs in the genus *Psychotria* (Rubiaceae). I predicted that older species have greater realized range occupancies and realized range extents, greater range occupancy ratios, and greater range extent ratios. I used species distribution modeling (using MAXent, Phillips et al. 2006.) to estimate the potential ranges of species and Bayesian

relaxed-clock dating (using BEAST, Drummond et al. 2006) of a molecular phylogeny to estimate the 'tip-ages' (*sensu* Roy and Goldberg 2007) of species. I focused on species in one clade within *Psychotria* (subgenus *Psychotria*) in one biogeographic region (Mesoamerica) that has been well collected and in which the taxonomic work has been recently updated (see Flora Mesoamericana, www.mobot.org/mobot/fm/). Mesoamerican *Psychotria* subgenus *Psychotria* is a valuable model group because the species vary by over three orders of magnitude in their range sizes yet are broadly ecologically similar.

3.3   METHODS

3.3.1   Study Taxa

Psychotria (Rubiaceae) is one of the most speciose angiosperm genera consisting of approximately 1600 species (Hamilton 1989a). In a detailed monograph of *Psychotria* subgenus *Psychotria* in Mesoamerica, Hamilton (1989a,b,c) suggested that it consisted of eight sections of species, each with a widespread, assumed ancestral species, and many narrowly endemic taxa assumed to be its descendents. This proposed pattern suggests that evolutionary history plays a vital role in shaping the current range size distribution of species in *Psychotria*. *Psychotria* species are primarily found in wet to seasonal forests pantropically, with a few species occupying dryer habitats. *Psychotria* species vary markedly in both their range size and local abundance (J. Paul *unpub. data*) and make up a significant proportion of species and stems in many tropical forests understories (Gentry 1990). The majority of *Psychotria* species are similar in their general growth form (small trees and shrubs), most are obligate outcrossing species

pollinated by insects (Stone 1995), and the seeds of most species are dispersed by frugivorous birds (Loiselle et al. 1995). Molecular phylogenetic work by Nepokroeff et al. (1999) and Andersson (2002) have largely confirmed the systematic relationships outlined by Taylor (1996) that Neotropical species of *Psychotria* form two distinct groups; *Psychotria* subgenus *Psychotria*, related to the other members of the subgenus in Africa and Asia, and subgenus *Heteropsychotria*, closely related to and polyphyletic with respect to species of the diverse genus *Palicourea*. *Notopleura*, a third Neotropical group that was formerly included in *Pychotria* is now considered a separate genus (Taylor 2001). There are ~78 recognized *Psy.* subgenus *Psychotria* taxa in Mesoamerica (Taylor, *Flora Mesoamericana;* www.mobot.org/MOBOT/fm/). In this paper, I use *Psychotria* to refer to *Psy.* subgenus *Psychotria*.

3.3.2   Realized and Potential Geographic Range Estimates

3.3.2.1 Collection Records

I used the Missouri Botanical Garden's (MBG) W3 Tropicos database of collection records to estimate the range sizes of species. I queried the database on September 9, 2006 for all collection records of the Mesoamerican *P.* subgenus *Psychotria* species (including records from South America). Species determinations in *Psychotria*, like many Rubiaceae taxa, can be challenging. All species determinations at MBG have been made or checked by one of us (C. Taylor), thus affording a high degree of consistency to the species identifications. Furthermore, MBG has one of the largest and most extensive Rubiaceae collections from Mesoamerica. Therefore, I chose to limit my geographic estimates to the MBG database in order to preserve the consistency of the species identifications.

3.3.2.2 Species Distribution Modeling

To model the potential geographic range sizes of species, I used the program MAXent (Phillips et al. 2006). Species distribution modeling uses presence-only data and a set of environmental variables to predict the probability of a species' occurrence across a landscape. MAXent uses a maximum entropy approach to species distribution modeling (Phillips et al. 2004) and has been shown to perform better than many other species distribution modeling programs (Elith et al. 2006), particularly for species with a small number of collection records (Hernandez et al. 2006). For each species, geo-referenced collection records were input into MAXent along with 20 environmental variables (e.g., altitude, precip. of the warmest quarter, etc., from the WorldClim database, www.worldclim.org; see Appendix B for full list). The model was then run using 75% of the data for 'training' the model and 25% for testing the model. Finally, the model was run with all collections used for training. Details on the species distribution modeling are presented in Appendix B.

3.3.2.3 Geographic Range Size Estimates

Range occupancy, for the R_O and R_O/P_O analyses, was calculated as the number of occupied (or predicted to be occupied) grid cells (Fig. 3.1). Range extent, for the R_E and R_E/P_E analyses, was calculated as Feret's diameter in ImageJ (Rasband 1997), the largest distance between two occupied (or predicted to be occupied) grid cells (Fig. 3.1). I calculated R_O and R_E for all species and R_O/P_O and R_E/P_E of all species with sufficient collection records (> 4 unique collection localities). Details of my methods to assess range sizes are presented in Appendix B.

3.3.3   Molecular Methods

I used both field-collected samples and herbarium sheets as the basis for DNA extractions. Leaf samples for DNA extraction (stored in 15 ml centrifuge tubes with silica gel) and corresponding voucher specimens were collected in 2005 in Costa Rica. Vouchers were field identified by J. Paul, and C. Taylor and J. Paul made final determinations of the specimens at MBG. Vouchers were deposited at MBG, the Carnegie Museum of Natural History Herbarium (CM), and the Universidad de Costa Rica Herbarium (USJ). To sequence many of the rare and endemic *Psychotria* species for which collecting was unfeasible, specimens of 73 of the 78 recognized taxa of Mesoamerican *P*. subgenus *Psychotria* were loaned from MBG to C. Morton at CM. I extracted DNA from spare leaf material. Some specimens yielded only highly degraded DNA, resulting in partial or missing sequence data for some species (Appendix C, Table 3B1). I used nuclear ribosomal internal-transcribed spacer sequences (*ITS*) and chloroplast *psbA-trnH* intron sequences for phylogenetic inference. *ITS* is one of the most extensively used loci for species level phylogenetic work in angiosperms (Mort et al. 2007). I also used the chloroplast intron *psbA-trnH* to attain a separate estimate of phylogenetic relationships within *Psychotria*. This intron was tested in three species of *Heteropsychotria* by Kress et al. (2005) and showed considerable variation at the species level. Details of laboratory techniques and protocols and justification of my molecular marker choices are provided in Appendix B.

3.3.4   Phylogenetic Inference and Divergence Time Estimation

I used a Bayesian relaxed-clock approach as implemented in the program BEAST v.1.4.5 (Drummond and Rambaut 2003) to concurrently estimate the phylogenetic relationships of species and their divergence times (Renner 2005). This method has been shown to provide robust estimates of both phylogeny and divergence times (Drummond et al. 2006). Details of alternative phylogenetic methods I used to analyze the data are presented in Appendix B. For my purposes here, relative ages of species are sufficient, but I used fossil evidence to guide a prior distribution on the root age of the tree, in order to make the ages more easily interpretable. Fossil dating estimates a minimum age of Rubiaceae as 53 million years old (Mya, Magallón et al. 1999), while a molecular analysis of angiosperm-wide divergence dates (Wikström et al. 2001) estimates Rubiaceae to be 61-64 Mya. *Psychotria* is a fairly basal group in the basal subfamily Rubioideae (Bremer and Manen 2000). Fossil pollen of the genus *Faramea* has been dated to 40 Mya (Graham 1985). *Faramea* is within the Coussareeae, which is closely related to the Psychotrieae alliance (Bremer and Manen 2000). Using these data as a guide, I gave the root age of the tree a gamma prior distribution with a median of ~ 46 Mya (zero offset = 37.0, Shape($\propto$) = 4.0, Scale(ß) = 2.5; Table 3.1). I included Hawaiian *Psychotria* species in the analyses (Nepokroeff et al. 2003), and used the time-to-most-recent-common-ancestor (tmrca) of these species to assess the validity of the absolute age estimates. I took a total evidence approach, running analyses on the combined *ITS-psbA* dataset (including all taxa with some missing data). Analyses including missing data can be robust and at times help break up long branches that would exist without including taxa with only partial data (Wiens 2006).   I ran analyses with and without partitioning the two loci and using separate models of molecular evolution. Results were similar, except partitioned analyses had parameter estimates with very

low effective sample sizes (ESS), indicating that the partitioned analyses were over-parameterized. Hence, the results presented here are from non-partitioned analyses. A summary of the priors and model parameters for each analysis is given in Table 3.1. I ran each BEAST analysis for 3 x 10^7 generations, sampling trees every 10^3 generations. For analyses, the burn-in was determined by looking for stabilization of parameter estimates and tree log-likelihoods using the program Tracer (Rambaut and Drummond 2004) and stabilization of clade probabilities using the web program Are We There Yet? (Wilgenbusch et al. 2004). Since the MCMC sampling of BEAST is stochastic, I ran each analysis three times, each with a different starting tree but otherwise identical parameters. All runs had parameter ESS values greater than 100. I used TreeAnnotator (Rambaut and Drummond 2006) to find the maximum clade credibility tree (MCC tree; the tree that maximizes the product of clade probabilities). This tree and its divergence time estimates were used for all further analyses.

3.3.5 Statistical Analysis

I performed least-squares linear regression analyses and one-way ANOVAs using SAS 8.2 (SAS Institute 2001). Variables were checked for normality using the SAS protocol 'proc univariate' and transformed as necessary to meet the assumptions of regression and ANOVA. Phylogenetic analyses revealed a strongly supported, basal divergence within the study species that was relevant to interpretation of the results. As a result, I conducted range size analyses on both the total set of species (termed all species) and the two clades separately (termed clades 1 and 2). Regression was used to examine relationships between species age and R_O, R_E, R_O/P_O, and R_E/P_E, as well as between species age and the morphological characters of fruit volume and plant height (estimated from flora descriptions). ANOVA was used to compare the mean values

of R_O, R_E, R_O/P_O, R_E/P_E, fruit volume, plant height, mean species age, median latitudinal position, elevation range, and elevation midpoint between species in Clades 1 and 2.

3.4   RESULTS

3.4.1   Phylogenetic Relationships and Divergence Times

Mesoamerican *Psychotria* species have primarily diversified in the last 16 Mya (Fig. 3.3), with most lineages diversifying within the last 12 Mya. The tmrca of the Hawaiian *Psychotria* species was estimated to be 10.06 Mya (95% HPDs: 5.97 and 14.63). The Bayesian MCC tree had a highly supported split of Mesoamerican *Psy.* subgenus *Psychotria* species into two distinct clades (labeled Clades 1 and 2; e.g., Fig. 3.2). These clades had not been identified previously based on any morphological, biogeographic, or ecological characters, but they were also recovered in parsimony and maximum likelihood searches of the full *ITS* and *ITS-psbA* datasets (results not shown). The relaxed-clock dating analysis estimates the divergence of these two clades at 15.43 Mya (95% HPD: 10.14 – 20.66 Mya) with posterior probability > 0.95. Clade 1 includes 27 taxa and Clade 2 includes 35 taxa.

3.4.2   Species Age and Geographic Range Size Metrics

When I analyzed all species, the relationship between species age and R_O was a significantly positive relationship (R^2 = 0.06, *P* = 0.02, df = 57; Fig. 3.3A), as was the relationship between species age and R_E (R^2 = 0.08, *P* = 0.02, df = 57; Fig. 3.3B). I found no

significant relationship between species age and R_O/P_O (Fig. 3.3C) but did find a significant, positive relationship between species age and R_E/P_E ($R^2 = 0.15$, $P < 0.01$, df = 57; Fig. 3.3D). When analyzing clades 1 and 2 separately, there was a significant, positive relationship between species age and R_E/P_E for Clade 1 ($R^2 = 0.30$, $P = 0.03$, df = 17; Fig. 3.3E), but not for Clade 2 (Fig. 3.3F). Neither clade had significant relationships between species age and R_O, R_E, or R_O/P_O.

3.4.3 Potential Explanatory Differences between Clade 1 and Clade 2

Species in Clade 1 and Clade 2 did not significantly differ in their average stature, fruit volume, elevation range size, elevation range midpoint, or in their average R_O, R_E, P_O, P_E, R_O/P_O, or R_E/P_E (Table 3.2). Species in clade 1 had significantly more southern latitudinal median realized range occupancies than species in clade 2 (8° 2' 24" vs 12° 56' 60"; F = 11.42, P = 0.001). Species in clade 1 are significantly older (back-transformed xbar = 2.71, S.D. = 2.56) than species in clade 2 (xbar = 1.51, S.D. = 2.27; $F_{1,54}$ = 6.10 , P = 0.017). Within clade 2, older species are significantly smaller in height ($R^2 = 0.13$, $P = 0.04$, df = 32) and have significantly smaller fruit volume ($R^2 = 0.38$, $P < 0.001$, df = 30) than young species. No such relationships were found in clade 1.

3.5   DISCUSSION

3.5.1   Species Age and Geographic Range Size Metrics

The results provide evidence of a positive relationship between species age and various measures of range size, supporting the central prediction of the age-and-area hypothesis. The strength of this relationship, while significant, was weak when either R_O or R_E were used for the analyses, explaining only a small fraction of the variances in range sizes. When I accounted for the environmental limitations of where a species can be expected to live by modeling species' potential ranges, the explanatory power of species age doubled in the R_E/P_E analysis (explaining 15% of the variance), but there was no relationship with R_O/P_O. Since some species had too few collections to accurately model potential ranges, the analyses using potential ranges had smaller samples sizes and less power. This may explain why the weak relationship found between species age and R_O was not recovered in the R_O/P_O analysis, but also suggests range occupancy is not influenced by species age as much as range extent. In contrast, the importance of accounting for potential range extent was clear, as the R_E/P_E relationship with species age was stronger despite the smaller sample size. Only one other study has looked for a relationship between species age and the occupancy of species' potential ranges. Schurr et al. (2007) found that species age had no effect on the proportion of species' potential ranges that were filled in a clade of South African Proteaceae (equivalent to my R_O/P_O metric, but no range extent metric was examined). They argue that processes acting on ecological timescales are largely responsible for the degree that species fill their potential ranges. Taken together, these results suggest that time-for-dispersal can be an important factor limiting how far within species' potential ranges individuals disperse and colonize, but not limit the density of occupancy of a

geographic area once colonizing populations have been established. In a 45-year experiment in Belgium (van der Veken 2007), transplanted populations of the forest herb *Hyacinthoides non-scripta* (Hyacinthaceae) have remained established but grown very slowly. As a result, small populations distant from the source population persist, but the occupancy of much of the suitable range between populations remains without individuals. The geographic spread of this slowly dispersing forest herb may exhibit properties similar to *Psychotria* species, which are primarily dispersed by understory birds (Nepokroeff et al. 2003). Rare long-distance dispersal events could establish distant populations intermittently, with species that have more time-for-dispersal (older species) colonizing farther into their potential ranges. However, if average population spread is slow, these species will fail to occupy large portions of their potential ranges. In particular, population spread can be strongly influenced by biotic interactions regulated by the other members of the newly colonized habitat. In areas where natural enemies or superior competitors are absent, populations could flourish, while in other areas population growth can be strongly regulated by the presence of these same factors. As a result, the occupancy of species ranges will be highly variable and show no relation to the time available for dispersal.

3.5.2 Species Age Estimates and Potential Explanatory Differences between Clade 1 and Clade 2

Mesoamerican *Psychotria* subgenus *Psychotria* species arose within the last 17 Mya, and two well-supported clades diverged approximately 15 Mya. Although these two clades have not been previously identified based on their morphology or ecology, the genetic data clearly indicate they have had separate evolutionary trajectories. Clade 1 species had significantly more southern ranges and species were on average older than clade 2 species. The species in these two

clades did not differ in their average geographic realized range occupancies or extents, or in their average range occupancy or extent ratios. Furthermore, species in the two clades did not differ in the average size of their elevation ranges (min. to max difference) or in their average elevation midpoint. It is striking that two groups of species that are so superficially similar could be so divergent in the impact of species age on their current range size distributions. Species age explained a full third of the variance in range extent ratio (R_E/P_E) of clade 1 species, but was not significant in clade 2. The morphological character that I examined that directly relates to dispersal, fruit volume, also did not differ significantly between the clades. However, within clade 2, younger species had significantly larger fruits than older species (as well as larger stature, which is likely a correlated character, e.g., Wright et al. 2007). Interestingly, clade 2 species did not show a significant relationship between species age and R_E/P_E. This could indicate that younger species in clade 2 have on average greater dispersal ability, and as a result, have been able to colonize farther into their potential range extents than expected if dispersal ability were a neutral character within the clade. Additional evidence would need to be garnered to address this hypothesis, but it is interesting note that for species with animal-dispersed seeds, larger fruit size can correlate with seed-dispersal by larger-bodied frugivores (Wheelwright 1985, Jordano 1995), which often have larger home ranges and greater average dispersal distances (Howe and Smallwood 1982, Holbrook and Smith 2000).

3.5.3 Robustness of Results

The age estimates for the Hawaiian *Psychotria* species are reasonable (tmrca of ~ 10 Mya), indicating they predated the present islands (oldest ~ 5 Mya), similar to Hawaiian *Drosophila* (~ 10 Mya, Thomas and Hunt 1991). Hence I have confidence that my age estimates

are a reasonable approximation, and even if they are refined with more data, they will not be dramatically altered. More importantly, the absolute ages estimated here are less important than the ages relative to one another, in terms of the comparative tests for which I used them. My geographic range estimates are conservative in that I limited the data to collections where I have confidence in the species identifications. I also concentrated on Mesoamerica as a biogeographic region because the of the relatively high collection intensity for a tropical genus like *Psychotria*. Predicted range sizes took into account many environmental variables, but including other sorts of data, such as edaphic factors, would likely strengthen the predictions. Furthermore, an aspect lacking from most ecological niche modeling studies is the exclusion of data on biological interactions (Phillips et al. 2006). Clearly, biotic interactions can limit species distributions, and incorporating maps of other species presence and absence, if geographically accurate data could be amassed, could refine the estimates I make of potential geographic distributions. However, similar to the age estimates, these sources of error should not be biased in their placement or magnitude among species, thus my estimates of realized and potential geographic ranges are highly suitable for the comparative framework in which are using them.

3.5.4 Other Factors Impacting Species Age and Range Size Relationships

Ultimately, the utility of species age as a predictor of range size rests on the assumption that various ecological and evolutionary processes (Gaston 1998) do not obscure the simple pattern predicted if species start with small range sizes, are prone to extinction, and transform their ranges at a relatively equal rate (e.g., Hubbell 2001, 2003, Hubbell and Lake 2003). A general positive age and area relationship may not be found if young species attain large geographic range sizes quickly, or if old species maintain small geographic range sizes. For

example, old species that once had large range sizes could decline in range size by failing to adapt to changing environmental or ecological conditions (Murray and Hose 2005). On the other hand, the process of speciation could generate young species that start their existence with large ranges. Since speciation is predicted to split range sizes under many models of geographic speciation, new species derived from ancestral species with large ranges have some probability of starting their existence with a large range size. This probability will relate to the nature of the geographic speciation event for a given species, specifically how asymmetrical it is (e.g., Waldron 2007). In clades where asymmetrical range splitting at speciation is commonplace, the set of new species would include species starting with both relatively large and small range sizes (Paul and Tonsor, 2008).   However, as demonstrated in the range extent ratio analysis, and particularly of clade 1 species, species age clearly impacts range extents of these species. This result indicates that species age and range size are related and the signal is detectable, despite potentially complicating factors. Furthermore the lack of signal in clade 2 does not suggest that species age is unimportant in this group. Rather, other processes may simply have a greater relative impact on range size variation in these species. The results of this study, in conjunction with an analysis of age-and-area in Neotropical *Piper* species (Paul and Tonsor 2008), suggests that the impact of species age may be particularly noticeable in species that have limited dispersal abilities and relatively homogeneous habitats. Indeed, it was on the tropical island of Ceylon (now Sri Lanka) that Willis first made his observations leading to the age and area hypothesis. Despite a tropical flora being the fodder for the hypothesis, and even his critics acknowledging age-and-area may be a more reasonable hypothesis for the tropics, *Psychotria* and *Piper* are the only two tropical plant genera in which age-and-area has been tested, and both support Willis' conjecture.

3.5.5 Conclusions

In summary, species age can be a significant predictor of range size variation in plant species. My results indicate that a time-for-dispersal effect may limit the extent, but not necessarily the occupancy of species' potential ranges. Although range expansions can occur rapidly in some cases (e.g., Clark et al. 1998), my results demonstrate that time can be a limiting factor to dispersal, much like time limits the rates of processes thought to be much slower than dispersal, like speciation. For example, Stephens and Wiens (2003) showed that a "time-for-speciation effect" is central in explaining the species diversity gradient seen in North American emydid turtles – areas where this turtle lineage has been present the longest have the most species. The time-for-speciation effect has also been implicated in explaining highland-lowland diversity patterns in Mesoamerican treefrogs (Smith et al. 2007). Furthermore, the time that a clade has existed (clade age), rather than diversification rate, is the most important predictor of clade species richness in animals (McPeek and Brown 2007). Time may be an important factor limiting the range sizes of many groups of species, particularly among taxa that have limited dispersal potential. I expect that the effects of species age on range size variation will be clade specific as I found here and as has been demonstrated elsewhere (Webb and Gaston 2000). Many studies use the genus as a level of comparison, yet genera vary greatly in their age, phylogenetic diversity, and ecological breadth. I suggest that the greatest benefit of phylogenetic comparative methods will come from careful study and consideration of levels of comparison.

Table 3.1. Prior parameter values for the BEAST *ITS-psbA* relaxed-clock analysis; [a]Lower and Upper 2.5% quantiles of distribution.

Parameter	Distribution	Lower Bound	Upper Bound
Root Height	Gamma	39.7[a]	58.9[a]
GTR Substitutions	Uniform	0	100
Gamma shape	Uniform	0	100
Proportion of invariant sites	Uniform	0	1
Lognormal relaxed-clock mean	Uniform	0	100
Lognormal relaxed-clock standard deviation	Uniform	0	10
Yule speciation process birth rate	Uniform	0	1^{-6}
Mean rate of evolution across tree	Uniform	-	-
Variation in rate of evolution across tree	Uniform	-	-
Covariation in rate of lineage and ancestral lineage	Uniform	-	-

Table 3.2. One-way analysis of variance results for morphological and geographic character comparisons between clades 1 and 2. [a]Analyses on log-transformed data. [b]Analyses on arcsine-square-root transformed data.

Variable	Source	DF	Sum of Squares	Mean Square	F Value	Pr > F	R-Square
Fruit volume[a]	Model	1	0.54	0.54	3.08	0.09	0.06
	Error	53	9.21	0.17			
Plant stature[a]	Model	1	0.51	0.51	1.09	0.30	0.02
	Error	55	25.71	0.47			
Elevation range	Model	1	429773.70	429773.70	0.97	0.33	0.02
	Error	55	24475047.35	445000.86			
Elevation midpoint	Model	1	3558.74	3558.74	0.01	0.90	0.00
	Error	55	13080075.47	237819.55			
Latitude midpoint	Model	1	359.79	359.79	11.61	0.001	0.17
	Error	57	1795.39	31.50			
R_O[a]	Model	1	0.31	0.31	0.13	0.72	0.00
	Error	54	127.24	2.36			

Table 3.2. (continued)

Variable	Source	DF	Sum of Squares	Mean Square	F Value	Pr > F	R-Square
$R_E{}^a$	Model	1	2.71	2.71	1.07	0.31	0.02
	Error	54	136.36	2.53			
$P_O{}^a$	Model	1	0.69	0.69	0.34	0.56	0.01
	Error	43	87.88	2.04			
$P_E{}^a$	Model	1	0.03	0.027	0.42	0.52	0.01
	Error	43	2.79	0.07			
$R_O/P_O{}^b$	Model	1	0.06	0.06	1.53	0.22	0.07
	Error	41	1.71	0.04			
$R_E/P_E{}^b$	Model	1	0.08	0.08	0.97	0.33	0.02
	Error	41	3.40	0.08			
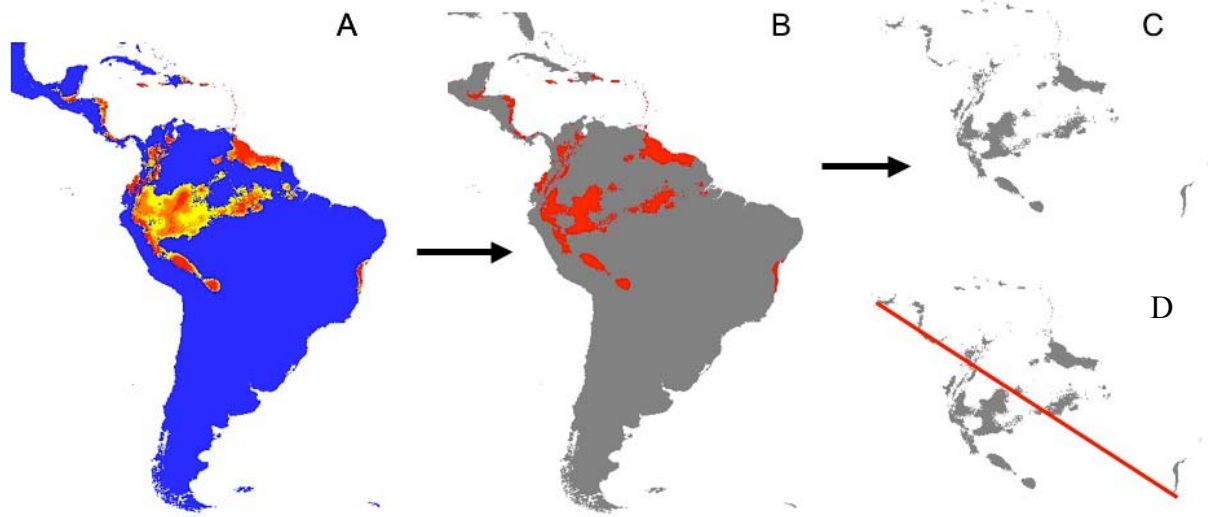
Figure 3.1. Measuring species' predicted and realized range occupancy and range extent. A) The potential distribution is modeled using a maximum entropy approach, B) the high probability areas are extracted, and C) the potential range occupancy (P_O, # of pixels predicted to be occupied), and D) potential range extent (P_E, the maximum linear extent between predicted occupied pixels, shown by red line) are measured. For realized ranges, only steps C and D are used, with the realized range occupancy (R_O) given by the number of occupied pixels (collection records) and the realized range extent (R_E) is given by the maximum linear extent between collections.
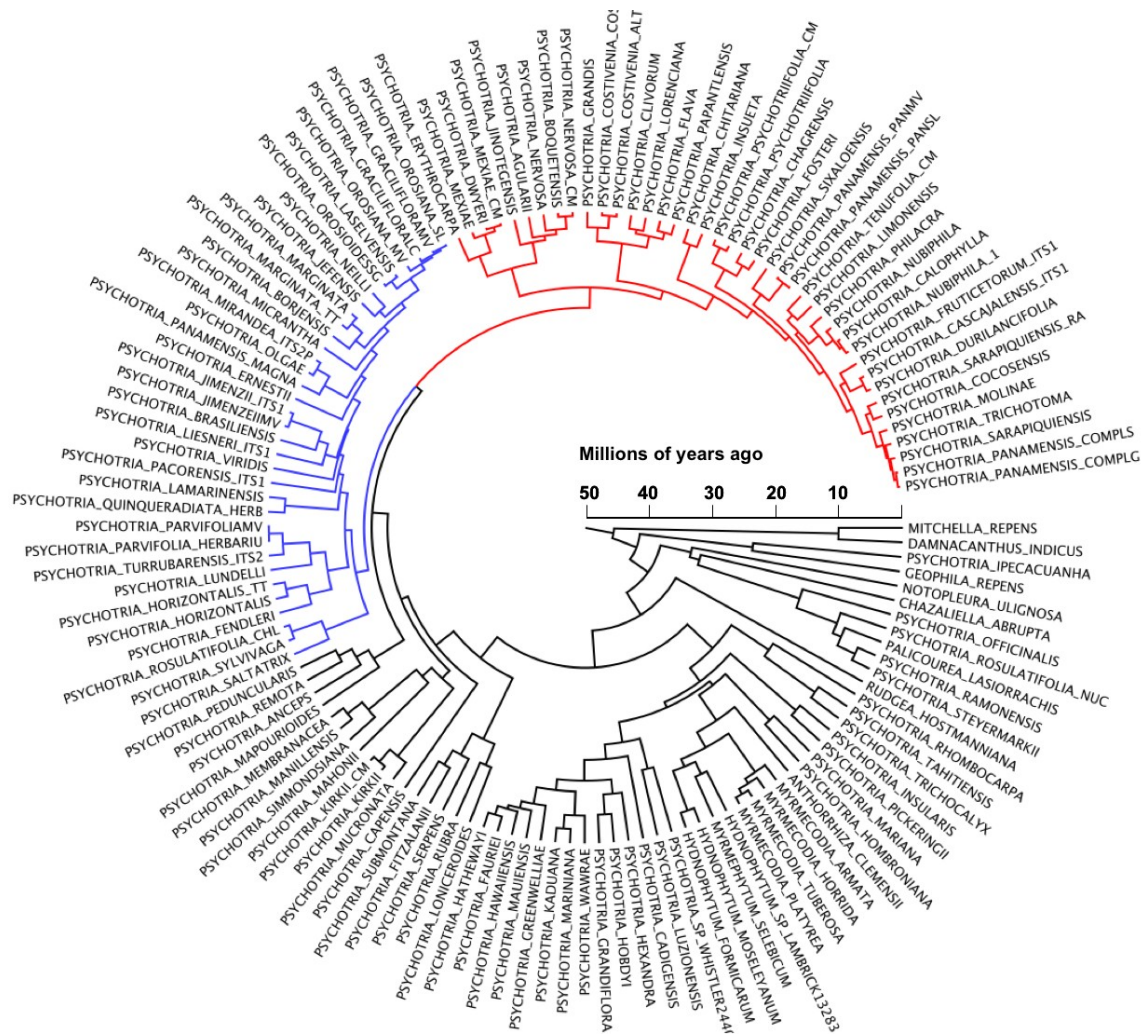
Figure 3.2. Bayesian relaxed-clock *ITS-psbA* MCC circle-chronogram. Scale gives time in Mya. Clade 1 with blue branches and clade 2 with red branches (see text for explanation).
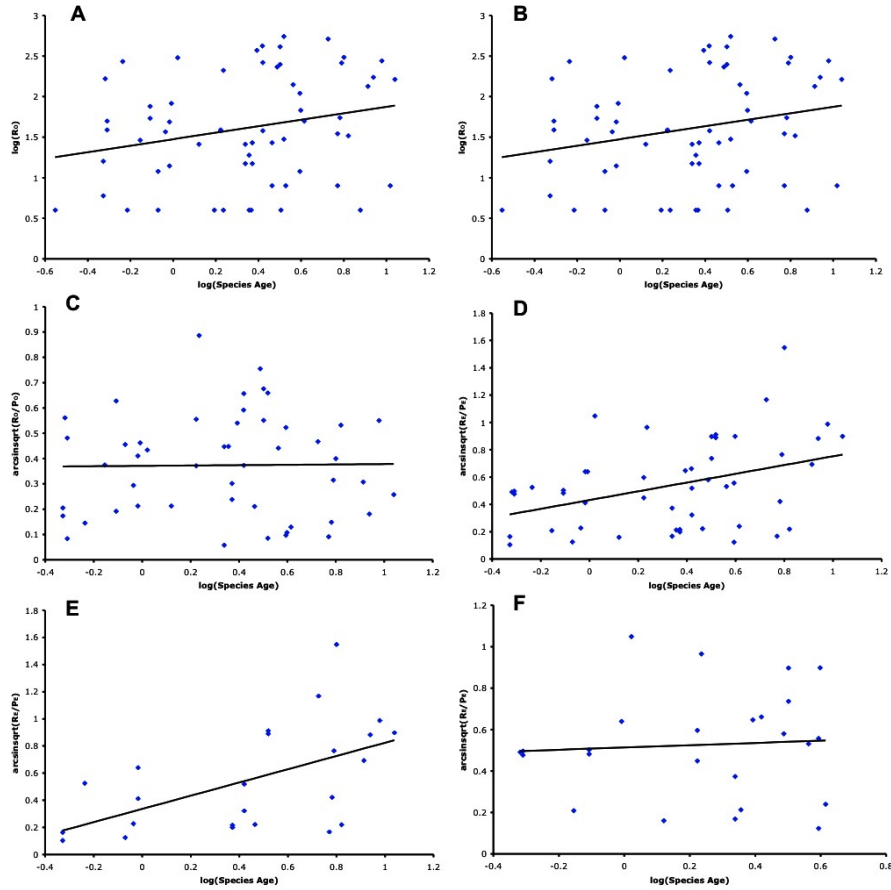
Figure 3.3. The relationship between species age and range size metrics in *Psychotria* subgenus *Psychotria* species in Mesoamerica. A) Regression of species age and realized range occupancy for all species (R_O; $y = 0.4x + 1.4747$, $R^2 = 0.06$, $P = 0.02$, df = 61). B) Regression of species age and realized range extent for all species (R_E; $y = 0.5093x + 1.4473$, $R^2 = 0.08$. $P = 0.02$, df = 61). C) Regression of species age and range occupancy ratio for all species (R_O/P_O; $y = 0.0069x + 0.3711$, $R^2 = 0.0002$, $P > 0.05$, df = 48). D) Regression of species age and range extent ratio for all species (R_E/P_E; $y = 0.3207x + 0.4313$, $R^2 = 0.15$, $P = 0.006$, df = 61). E) Regression of species age and range extent ratio for clade 1 species (R_E/P_E; $y = 0.4873x + 0.3364$, $R^2 = 0.30$, $P = 0.029$, df = 23). F) Regression of species age and range extent ratio for clade 2 species (R_E/P_E; $y = 0.0546x + 0.5136$, $R^2 = 0.0044$, $P > 0.05$, df = 24).

58

4.0 PHYLOGENETIC STRUCTURE, CO-OCCURRENCE, AND ABUNDANCE IN PSYCHOTRIEAE (RUBIACEAE) SPECIES ASSEMBLAGES IN COSTA RICA

4.1 ABSTRACT

Understanding how the phylogenetic relatedness of species impacts community assembly, co-occurrence, and abundance is a burgeoning discipline at the interface of ecology and evolution. Phylogenetically related species are expected to co-occur if their ecological niches are evolutionarily conserved, because these species will share many traits that allow them to live in similar habitats. However, if co-occurring species are too similar in their ecological niches, competition for resources is expected to repel such species from co-occurring or result in divergence of their niches. Thus, phylogenetic niche conservatism and limiting similarity provide opposing predictions for how communities or assemblages of related organisms will be phylogenetically structured. Using two genera (*Psychotria* and *Palicourea*) in the clade Psychotrieae (Rubiaceae), I examined the impact of phylogenetic relatedness on the co-occurrence and variation in abundance among these species in the forests of Costa Rica, Central America. I used co-occurrence and abundance data collected on 240 transects nested in seven Psychotrieae assemblages across Costa Rica and a phylogenetic hypothesis of species relationships using DNA sequence data to examine the phylogenetic structure of Psychotrieae assemblages. I found that Psychotrieae assemblages are overall significantly phylogenetically

overdispersed, indicating that co-occurring species are less related than expected by chance. Examining the seven assemblages individually, I found that the phylogenetic structure of assemblages differed significantly, with some assemblages overdispersed and others clustered (more related than expected by chance). Phylogenetic diversity also differed significantly across assemblages, but was often positively associated with species richness. Abundance was not a conserved trait across the phylogenetic tree of species found on transects, although species' geographic range characteristics were phylogenetically conserved. Species with high local abundances did not have larger geographic or elevational ranges. On the whole, species found at higher elevations sites were more abundant. Within one heavily sampled assemblage, I found an inverse relationship between the phylogenetic relatedness of species on transects and the variation in abundance among species on transects, indicating that closely related species are more similar in their abundances. However, when species are found on phylogenetically clustered transects, their average abundances are lower than when they are found on phylogenetically overdispersed transects. When the relationship between phylogenetic relatedness and variation in abundance was examined across assemblages, they were positively correlated, in opposition to the trend found at the local scale. I interpreted these results in light of the opposing pressures exerted by the ecological process of limiting similarity and the evolutionary process of phylogenetic niche conservatism.

4.2   INTRODUCTION

Ecologists have long been interested in the processes that govern the assembly of ecological communities and the mechanisms that maintain diversity with communities (Chesson

2000, Connell 1971, Grinnell 1917, Grinnell 1924, Hubbell 2001a, Hutchinson 1959, Janzen 1970, Johnson 1910, MacArthur 1960, MacArthur and Levins 1967, Paine 1966, Ricklefs 2004, Tilman 1982, Whittaker et al. 1975). Although incorporating evolutionary relationships among species into the understanding of community assembly dates back over sixty years (e.g., genus-to-species ratios, Elton 1946), only recently has an explicit connection been made between the patterns of phylogenetic relationships of co-occurring species and the processes that potentially drive these patterns (Webb et al. 2002). Recently, investigations of the phylogenetic structure of communities (e.g., Webb 2000, Kembel and Hubbell 2006) and assemblages of closely-related species (Cavender-Bares et al. 2004, 2006, Helmus et al. 2007, Slingsby and Verboom 2006, Vamosi and Vamosi 2007) have begun to shed light on how the phylogenetic relatedness of species impacts co-occurrence and abundance, by asking if the phylogenetic structure of communities is significantly different from random.

In the bulk of studies to date, researchers use patterns of phylogenetic structure to draw inferences about what processes are most important to assembling communities. A central axiom of evolutionary biology predicts that closely related species generally share many characteristics due to common ancestry and modification by decent (Darwin 1859).   When related organisms share a set of ecological traits (traits that dictate the combined abiotic and biotic conditions in which they can maintain stable populations, Hutchinson 1957), they are deemed to have similar ecological niches (e.g., Peterson et al. 1999). The ecological niches of related organisms can remain similar over evolutionary time via niche conservatism (Wiens and Graham 2005). When species' niches are evolutionary conserved, closely related species should be found living together in similar environments, since they share a set of traits that allow them to pass through abiotic ecological filters (Weiher et al. 1998) imposed by a particular

61

environment (e.g., temperature or precipitation levels). As a result of phylogenetic niche conservatism and ecological filtering, related species can exhibit clustering in space (phylogenetic clustering, when co-occurring species are more related than expected by chance, Webb 2000, Webb et al. 2002). In contrast, theory focused on species interactions (Hutchinson 1959, MacArthur and Levin 1964) suggests that for complete competitors to coexist they must have diverged along some important ecological niche axis. As a result, competitive exclusion (Gause 1934) predicts that close relatives should not be found in the same habitat and that co-occurring species should be less related than expected by chance (phylogenetic overdispersion, Webb 2000, Webb et al. 2002; also called phylogenetic evenness, see Kraft et al. 2007). Hence, when species' ecological niches are evolutionarily conserved, the processes of ecological filtering and limiting similarity impose contradictory expectations for phylogenetic community structure (Webb et al 2002). Empirical data have demonstrated that in the cases examined thus far, communities are phylogenetically nonrandomly structured, and that the direction of structure (clustered or overdispersed) is both taxonomically and spatially scale dependent (Cavender-Bares et al. 2006, Swenson et al. 2007). Specifically, at lower taxonomic scales (e.g., within genera) and smaller spatial scales (e.g., $10^1 - 10^3$ m^2) species tend to be phylogenetically overdispersed, while at higher taxonomic scales (e.g., within families or orders) and larger spatial scales (e.g., $> 10^4$ m^2) species tend to be phylogenetically clustered (Cavender-Bares et al. 2006, Kembel and Hubbell 2006, Swenson et al. 2007).

Despite the short history of investigations of phylogenetic spatial structure, the signal of phylogeny is now well documented in nonrandom patterns of co-occurrence (Cavender-Bares et al. 2004, Helmus et al. 2007, Slingsby and Verboom 2006, Swenson et al. 2007, Webb 2000). In contrast, the impact of phylogeny on patterns of abundance has been much less studied (e.g.,

Andersen et al. 2004, Cavender-Bares et al. 2006, Silvertown et al. 2006), although the phylogeny-abundance relationship is fundamental to understanding how phylogeny impacts the structure of communities (Webb et al. 2002). In meadow communities in Great Britain, Silvertown et al. (2006) found no evidence of an effect of phylogeny on the abundance of meadow plants or on their degree of niche overlap or divergence. In contrast, in yeast communities of decaying cacti, Andersen et al. (2004) found that in one cactus species, abundant yeast species tended to be phylogenetically overdispersed, while the rare species tended to be close relatives. Similarly, Vamosi and Vamosi (2007) found that common predaceous diving beetles tended to be overdispersed phylogenetically, while rare species showed no such pattern.

A clear expectation for the relationship between phylogeny and abundance is hard to ascertain, because predictions rely on knowing which of two processes is paramount in structuring abundances in a given community, niche conservatism or limiting similarity. A heuristic table of the potential relationship between phylogenetic structure and the coefficient of variation in abundance among species is presented in Figure 4.1. Consider the case of species showing ecological niche conservatism. If a set of species are similar in their traits, and hence also similar in their ability to extract resources from the environment and interact with other species, they are expected to be similar in their predicted equilibrial abundances (i.e., abundance is a conserved trait in these species; Figure 4.1, upper left). Conversely, phylogenetically distant species will differ in their traits, and as a result, they will also differ in their abundances, depending on which species' set of traits are best matched to local conditions (Figure 4.1, upper right). Now consider the case where competitive interactions are deemed most important in structuring abundance distributions (e.g., Tilman 1988), but co-occurring species are phylogenetically clustered, because habitat filtering leads to phenotypic attraction (Webb et al.

2002). Limiting similarity predicts that poorer competitors should be driven from equilibrium communities (MacArthur and Levin 1967). However, the process of driving species from communities takes time, and communities may not be at an equilibrium state. Hence, while species do coexist and compete for resources, all species' population growth rates will be negatively affected. Since species differ in their competitive abilities within a given niche space, poorer competitors' population growth rates will be depressed, resulting in considerable variation in abundance among co-occurring species (Figure 4.1, lower left). Alternatively, if co-occurring species are phylogenetically distant and experience (presumably) weak competition due to trait divergence, their abundances may be similar, because competition with related taxa is relaxed, and abundances of all species can be driven by general site conditions (Figure 4.1, lower right). However, for all these situations, if an unmeasured factor is more important than either competition or niche conservatism, there may be no consistent relationship between the phylogenetic structure of assemblages and the variation in abundance in those assemblages. An unpredictable relationship between phylogeny and abundance is also predicted by neutral theory in ecology (Hubbell 2001a), where the identities of species (and individuals) and hence their phylogenetic relatedness are inconsequential to their abundance. Based on both limited empirical information and the conflicting predictions resulting from different theoretical frameworks, a general expectation for how phylogenetic structure should impact abundance distributions requires empirical investigation.

Here, I investigate the spatial phylogenetic structure of co-occurring species and how phylogenetic structure impacts abundance in the diverse angiosperm clade Psychotrieae (Rubiaceae) using data collected from 240 transects located in moist tropical forests in Costa Rica. I focus on ecologically similar species in the clade that are woody understory sub-shrubs,

shrubs, and small trees, *Psychotria* and *Palicourea*. Two central goals of this study are: 1) to use species co-occurrence data to determine how assemblages of Psychotrieae species are phylogenetically structured, and 2) to draw an explicit link between phylogenetic structure and variation in abundance among species in these assemblages. I inferred the phylogenetic relationships of Psychotrieae species using molecular sequence data. Using a phylogenetic tree of species relationships, I assessed the phylogenetic structure of Psychotrieae assemblages using two measures that provide different information about the phylogenetic level of structuring of assemblages. The net relatedness index (NRI) uses the average phylogenetic distance among all co-occurring species to assess the degree to which these species are clustered or dispersed across a phylogenetic tree (Webb et al. 2002). The nearest taxon index (NTI) uses the average phylogenetic distance of each species to the phylogenetic nearest relative with which it co-occurs, to assess the degree that closest relatives co-occur (Webb et al. 2002). Hence, NRI provides information about the general phylogenetic structure of an assemblage while NTI more specifically indicates the degree to which closely-related species co-occur in communities. These statistics have been used in a number of phylogenetic structure studies (e.g., Kembel and Hubbell 2006, Valmosi and Valmosi 2007, Webb 2000, Weiblen et al. 2006) and their utility and power have been addressed in a simulation study (Kraft et al. 2007). I also calculated the phylogenetic diversity (PD; Faith 1992) of each transect, which is the proportion of the phylogeny represented by species in a transect and provides an estimate of the amount of unique evolutionary history (e.g., sum of unique branch lengths) found in a transect. I investigated the phylogenetic structure of communities using three approaches. First, using all 240 transects, I asked if the phylogenetic relatedness of co-occurring Psychotrieae species is structured significantly differently from random. Second, I grouped transects into seven geographically

65

distinct assemblages. I asked if the mean phylogenetic structure of each of these assemblages was significantly different from random, and if the mean phylogenetic structure of assemblages differed across assemblages. Furthermore, to address the relationship between phylogenetic structure and relative abundance of species, I correlated measures of phylogenetic structure with the variation in abundance among species on transects. I calculated the coefficient of variation in abundance (CV_A) among co-occurring species on each transect, and the species richness of each transect. I then correlated these measures with NRI, NTI, and PD. I also analyzed these data by grouping transects and examining correlations within and across assemblages. Finally, I examined the relationship between species' regional distribution measures (e.g., geographic range size, elevation range size) and both their abundance and frequency as inferred from the transect data, asking if species' abundance and distribution characters are phylogenetically conserved.

Based both on theoretical expectations regarding the interplay of niche conservatism, habitat filtering, and limiting similarity (Webb et al. 2002) and on empirical evidence (Cavender-Bares et al. 2004, Slingsby and Verboom, Webb 2000) I made the following predictions: 1) On average, Psychotrieae assemblages will be phylogenetically overdispersed at the transect scale, where limiting similarity may constrain the co-occurrence of close relatives; 2) Psychotrieae assemblages will be phylogenetically clustered at the assemblage scale, where habitat filtering and phylogenetic niche conservatism are expected to result in closely related species with many shared traits, but sufficient spatial separation of individuals can allow coexistence of ecologically-similar phylogenetically-related species; 3) Variation in abundance among species on transects (CV_A) will be inversely related to the phylogenetic relatedness of species on transects, following the conventional assumption that similarity in phenotypic and ecological

traits, and hence abundance, will increase with phylogenetic similarity. There has been little previous comparative phylogenetic structuring work (Webb et al. 2002), and it is not clear how phylogenetic structure of assemblages may be impacted by the various abiotic and biotic factors that can vary between assemblages. Hence, I did not have *a priori* expectations for the consistency of phylogenetic structuring across assemblages, as measured by NRI, NTI, and PD.

4.3 METHODS

4.3.1 Study Taxa

Species within the clade Psychotrieae are highly diversified in the New World tropics, comprising over 1500 species (Taylor 1996). Mesoamerican members of this clade are drawn primarily from two clades (*Heteropsychotria/Palicourea* and *Psychotria* sensu stricto, Nepokroeff et al. 1999) that have diverged ~ 40 million years ago (Mya; Paul, unpublished). Each of these clades contains two subclades that diverged ~ 15 Mya ago (Paul, Chapter 3). Psychotrieae species are ideal candidates for studies of phylogenetic assemblage structure, because many species co-occur in local communities, species have diverse geographic and elevation range placement and sizes, species vary by orders of magnitude in their abundance, yet the majority of species are restricted to moist or wet tropical forests habitats (Hamilton 1989a, Taylor 1996). In this paper, I use Psychotrieae as a general term to refer to *Psychotria* and *Palicourea* species, *Psychotria* to refer to *Psychotria* subgenus *Psychotria*, and *Heteropsychotria* to refer to *Psychotria* subgenus *Heteropsychotria* and *Palicourea* species.

4.3.2 Transect Surveys

I surveyed 240 belt transects nested in seven locations in Costa Rica (Table 4.1) to assess the presence and abundance of Psychotrieae species. Each transect was 50 m long and 4 m wide, for a total survey area of 200 m^2 per transect and 48000 m^2 total across all transects. I noted all *Psychotria* and *Palicourea* individuals > 20 cm tall on each transect, recording their location to the nearest meter on the 50 m axis of the transect. I also recorded the elevation and the approximate latitude and longitude of each transect using a global positioning system (GPS) when possible. Transects were located in lowland and premontane wet forests in Costa Rica, which is a Central American center of diversity for both *Psychotria* and *Palicourea* species (Hamilton 1989a, Taylor 1989). I surveyed 141 transects, all located in La Selva in 2003. I used La Selva's trail system and grid of georeferenced location poles to establish transects. All transects in La Selva started at a randomly chosen pole location and were surveyed along the horizontal axis of the La Selva grid system. In 2005, I surveyed 99 additional transects, 14 in La Selva and the rest at six other sites, with 12 to 18 transects representing each site. At each assemblage location, I used established trail systems and prior knowledge of the forest to choose general areas to place transects. Within these areas, the exact starting point and direction of transects was randomly chosen. Transects were situated to avoid crossing large trails and rivers, but were surveyed regardless of the difficulty in traversing the site (e.g., steep aspect, vine tangles) to get unbiased estimates of Psychotrieae presence and abundance. I identified species in the field and collected voucher specimens for each representative species. Vouchers were deposited in three herbaria (Missouri Botanical Garden, MBG; Carnegie Museum of Natural History Herbarium, CM; and Universidad de Costa Rica Herbarium, USJ). Charlotte Taylor

(MBG, Rubiaceae curator) and I made final determination of the voucher specimens. I also collected leaf tissue stored in silica gel for each voucher specimen for DNA extraction.

4.3.3   Molecular and Phylogenetic Inference Methods

I combined published internal-transcribed spacer (*ITS*) sequences from GenBank with sequences that I generated using field collected samples. The list of species names and voucher specimen information is presented in Appendix D and the details of my molecular methods are presented in Appendix E. I aligned 311 sequences from operational taxonomic units (OTUs) representing all of the available Psychotrieae *ITS* sequences as of April 2008.   Preliminary analysis showed strong geographic structure in Psychotrieae (Paul, unpublished) so I reduced the OTUs to 187 to include all the Neotropical species. Aligned sequence data was analyzed by maximum likelihood inference using the programs GARLI (Zwickle 2006) and PAUP* (Swofford 2002). ModelTest (Posada and Crandall 1998) was used to determine that the appropriate model of molecular evolution, which was the general time reversible model with gamma distributed rates and a proportion of invariant sites (GTR+I+G). I inferred the highest likelihood tree using GARLI, and then optimized the tree in PAUP*. I then used the program PATHD8 (Britton et al. 2007) to rescale my likelihood tree to a time calibrated ultrametric tree. PATHD8 estimates node ages by estimating mean path lengths from nodes to the tips, while accounting for molecular clock deviations from calibrated nodes (Britton et al. 2007). I placed three constraints on the nodes of my tree based on previous work (Paul, Chapter 3) and fossil evidence (Graham 1985, Magallón et al. 1999). Using previous estimates of the age of the Rubiaceae, I constrained the root node of my tree to 46 Mya, and the crown ages of two *Psychotria* subgenus *Psychotria* clades to be 13.4 Mya and 12.1 Mya. I found that these time

constraints provided reasonable estimates of the absolute crown age of the Hawaiian *Psychotria* diversification (Paul, Chapter 3), and hence I used these constraints for this analysis as well. The resulting ultrametric tree was then pruned of species until only species represented in the transect dataset were present. This method of pruning a larger tree is preferable to inferring the tree only using sequences from species found on transects, because of the greater resolution and better branch length estimates obtained by having more taxa in an analysis (Hillis 1998). I was not able to obtain sequence data for some species found on transects. Since the purpose of this paper is to explore the impact of phylogeny on co-occurrence and abundance rather than a general description of transect species distributions, I excluded these missing species (N = 9) from all analyses. One species, *Psychotria graciliflora*, had sequences from two populations (MVEB and LC) that were divergent, and field observations suggested these two populations are morphologically different (Paul. pers. obs.), so I included these two sequences as separate taxa for my analyses. Similarly, I included two subspecies of *Psychotria panamensis*, subspecies *compressicaulis* and *panamensis* as separate taxa, based on clear phylogenetic and geographic evidence that these two subspecies are probably reproductively isolated species, and may not actually be sister taxa (see phylogeny in Paul, Chapter 3). For simplicity and clarity, I refer to the 39 OTUs used in my phylogenetic analysis as species for the duration of this paper.

4.3.4   Phylogenetic Structure Methods

I used the program Phylocom (Webb et al. 2007) to assess phylogenetic structure. Randomized datasets (see description of the randomization procedures below) are used for significance testing and calculating NRI and NTI, which are the standardized effect sizes of phylogenetic structure (Webb et al. 2007). Phylocom calculates two principle values that are

used to derive summary statistics of phylogenetic structure. To assess the average phylogenetic clustering of co-occurring species, the pairwise phylogenetic distance of each species in a given sample is calculated, and the mean pairwise phylogenetic distance (MPDsample) is recorded for each sample. The same calculations are made on repeatedly randomized data (999 iterations) and the mean (rndMPDsample) and the standard deviation (sd_rndMPDsample) of these randomization-based null hypothesis values are recorded for each sample. The net relatedness index, NRI, is then calculated as [-1 * (MPDsample –rndMPDsample)/(sd_rndMPDsample)] (Webb et al. 2007). This measure gives a standardized effect size of the average relatedness of species in samples compared to a random expectation, with negative numbers indicating overdispersion and positive numbers indicating clustering (Webb et al. 2007). To assess the degree that closest relatives show phylogenetic clustering, the phylogenetic nearest neighbor distance is calculated for each species in a given sample, and the mean phylogenetic nearest neighbor distance (MNND) is recorded (this same measure has also been referred to as the mean nearest-taxon distance, MNTD, Kraft et al. 2007). The same calculations are made on randomized data and the mean (rndMNNDsample) and standard deviation (sd_rndMNNDsample) are recorded for each sample. The nearest taxon index, NTI, is then calculated as [-1 * (MNNDsample –rndMNNDsample) /(sd_rndMNNDsample)] (Webb et al. 2007). This measure gives a standardized effect size of the average phylogenetic distances between the closest relatives found in samples compared to a random expectation, with negative numbers indicating overdispersion and positive numbers indicating clustering (Webb et al. 2007). Phylocom offers four different methods of randomization that vary in the specifics of how the randomization is conducted and what species are included (Webb et al. 2007). I used two methods that differ in their details but both randomize species identities across samples, and

71

only include species from the phylogeny in the randomization that are present in at least one sample (methods 1 and 3 in Webb et al. 2007). The results of these two methods were nearly identical; hence I only present the results using method 1. For further details on the randomization methods see Kembel and Hubbell (2006). For an assessment of the power and sensitivity of NRI and NTI see Kraft et al. (2007). I also calculated the phylogenetic diversity (PD, Faith 1992) for each transect. Phylogenetic diversity is a measure of the proportion the total tree length represented by summing the unique branch lengths of all species found in a transect. If many species from different parts of a phylogeny are represented in a sample, then PD will be high, while if many close relatives occur together, PD will be low because these species also share a large proportion of their total branch length with one another. This measure has been widely used in the conservation literature and I include it here for the benefit of comparisons with other studies. For a broader overview of Phylocom's various applications see Webb et al. (2007).

4.3.5   Grouping of Transects for Phylogenetic Structure Estimates

I assessed phylogenetic structure of the transect data at a number of different scales. For the analysis across all transects (*all transect analyses*) I calculated NRI and NTI for each transect. For the analysis of distinct assemblages (*within assemblage analyses*) transects were grouped based on their geographic location (see Table 4.1) and the mean and standard deviation of NRI and NTI were calculated for each assemblage. These values were also used to compare across assemblages (*across assemblage analyses*). I used one-way analysis of variance (ANOVA) to compare mean NRI, NTI, and PD among transects. For all phylogenetic structure analyses, transects that had no species (N= 7) or only one species (N = 28) were excluded.

72

4.3.6 Variation in Abundance and Species Richness

To address the relationship between phylogenetic structure and variation in abundance among species, I calculated the coefficient of variation in abundance (CV_A) for each transect. If closely-related species are similar in many phenotypic traits, they may also be similar in their abundance, due to phylogenetic niche conservatism. I used CV_A to determine if variation in abundance scaled with the degree of phylogenetic relatedness on transects. I also calculated a standardized abundance (StdAbun) for each species as the deviation of the species' within-transect abundance from its mean abundance across transects, divided by the standard deviation of the species' abundance across transects. If a species' StdAbun is negative for a given transect, that species is less abundant than its transect-wide average; if StdAbun is positive, it is more abundant. I used StdAbun to ascertain if the average standardized abundance of species within transects changes with species richness, CV_A, or phylogenetic structure. In addition, I calculated the mean number of species per transect and the mean number of species per transect per assemblage. I used correlation rather than regression to assess the relationships between phylogenetic structure, CV_A, and species richness because the direction of causality between these variables is unclear. I conducted these analyses on all transect data combined (*all transect analyses*, only at the 50 m scale) and the assemblage data (*within and across assemblage analyses*), using mean values or NRI, NTI, CV_A, and species richness per transect and assemblage for analyses. These results are presented along with other results from a given scale of analysis (i.e., *across assemblage analyses*).

73

4.3.7 Phylogenetic Signal of Abundance and the Relationship between Regional and Local Abundance

I examined the relationship between species' regional characteristics (e.g., geographic range size and elevation range size) and their mean abundance and frequency on transects. Using collections data from the Missouri Botanical Garden's web interface (Tropicos, www.tropicos.org), I calculated each species' range occupancy (Gaston 2003) as the total number of unique collections points, and each species' range extent as the maximum linear distance between two collection points. I calculated each species' elevational range size using published estimates from *Flora Costaricensis Family # 202 Rubiaceae* (Burger and Taylor 1993) and *Flora Mesoamericana* (Taylor 2007, Rubiaceae, provided by C. Taylor). I used these data to define the regional species pool for each assemblage, defined as the number of species at a given assemblages' mean elevation that could potentially be found at that elevation. The elevational zonation of Costa Rican *Psychotria* and *Palicourea* species that make up the full regional species pool is presented graphically in Appendix F. From the transect data, I calculated the mean abundance of each species across all transects (calculated using only transects in which a given species was actually present), the frequency of occurrence on transects (calculated as the number of transects in which a given species was present, divided by the total number of transects in which a species could possibly be present, based on their elevation limitations). I used correlations to examine the relationship between species' mean transect abundance, transect frequency, range occupancy, range extent, elevational range size, and median elevational range. To assess the phylogenetic signal of local and regional abundance, I used Phylocom's "Analysis of Traits" (AOT) function (Webb et al. 2007). I calculated the phylogenetic signal of mean abundance, transect frequency, range occupancy, and range extent. Phylogenetic signal is

calculated by AOT using independent contrast calculations that follow Blomberg and Garland (2002) and Blomberg (2003). If phylogenetic signal is conserved, divergences between species in trait values (e.g. mean abundance) will be small, while divergences will be large if the trait is evolutionarily labile (Webb et al. 2007). Significance is tested using randomization. I used this measure to assess if abundance and distribution were relatively labile or conserved traits across the tree of species represented on transects.

4.3.8   Statistical Analyses

Randomization tests of phylogenetic structure were made is Phylocom as described above. For analyses where the mean and standard deviation of NRI and NTI values were used, I tested significance with a one-sample t-test (Kembel and Hubbell 2006). All correlation and ANOVA analyses were conducted in SAS 9.1.3. (2005), using the SAS 'Corr' and 'Glm' procedures, respectively. I examined variable distributions using the SAS 'univariate' procedure to look for departures from normality. I transformed non-normal variables in SAS for analyses when needed. I calculated the coefficient of variation in abundance using the SAS 'Means' procedure.

4.4   RESULTS

4.4.1   Phylogenetic Relationships and Descriptive Statistics of Transects

I found a total of 48 species on the transect surveys. Of these, I had sequence data from 39 species to use in the phylogenetic analysis. The phylogenetic relationships of the species found on transects is shown in Fig. 4.2. In all, the species found on transects were represented by 13 species in the *Psychotria* clade and 26 species in the *Heteropsychotria* clade. Over half of the species in the *Heteropsychotria* clade diverged from a common ancestor within the last ~ 5 million years (Fig. 4.2). La Selva had the greatest number of total species (but also the highest sampling effort), followed by the two Monteverde sites. Rara Aves had the greatest mean number of species per transect and San Gerardo had the fewest (Table 4.1). Assemblages had from 14% to 50% of their potential regional species pool species represented on transects.

4.4.2   All Transect Analyses

When all transects were analyzed together, transects varied in their individual phylogenetic structure estimates, with NTI values having a broader range than NRI, particularly in the negative numbers, indicating overdispersion (Fig. 4.3). The mean phylogenetic structure, as measured by the co-occurrence of closest relatives, was significantly overdispersed (NTI = -0.31, df = 204, P < 0.0001), but not significantly different than random across the tree as a whole (NRI = -0.04, df = 204, P = 0.48).   Species richness and CV_A were significantly positively correlated (Table 4.2). Mean standardized abundance of species was also significantly positively correlated with both CV_A and species richness.   The correlations between CV_A or species

76

richness and NRI or NTI were not significant. Both species richness and CV_A were significantly positively correlated with PD (Table 4.2).

4.4.3   Within Assemblage Analyses

When transects were grouped into seven discrete geographic assemblages, all assemblages were significantly phylogenetically structured, as assessed by at least one of the two phylogenetic structure metrics (NRI or NTI; Fig 4.4).   In all, three assemblages showed significant tree-wide overdispersion (negative NRI) and two showed phylogenetic clustering (significantly positive NRI).   Three assemblages had significantly overdispersed nearest relative measures (negative NTI) and three were significantly clustered (positive NTI).   In all locations spare one (San Luis), the sign of the NRI estimates matched the sign of NTI estimates (Fig. 4.4). In La Selva, the correlations between CV_A and both NRI and NTI were significantly negative, indicating that the variation in abundance of species increased as phylogenetic relatedness decreased (Table 4.2).   In addition, CV_A and species richness were positively correlated, and both of these variables were significantly positively correlated with mean standardized abundance (Table 4.2).   For all other assemblages, sample sizes were too small for detection of correlations of the magnitude estimated for the all-transect analysis or the within La Selva analysis.   Hence, there were no significant correlations detected between CV_A or species richness and NRI and NTI.   For all assemblages except MVEB, species richness and PD exhibited correlations of $r > 0.72$, and the correlations were significant in four of the seven assemblages (Table 4.2).

4.4.4 Across Assemblage Analyses

Phylogenetic structure (NRI and NTI), species richness, and phylogenetic diversity (PD) differed significantly across the seven assemblages (Table 4.3). Examining the relationships of mean assemblage values of species richness, CV_A, NTI, NRI, and PD across assemblages, mean assemblage CV_A was positively correlated with mean NTI ($r = 0.74$, df $= 6$, $P = 0.06$) and mean NRI ($r = 0.73$, $P = 0.07$; Fig. 4.5). Despite the high magnitudes of the correlation coefficients, both relationships were only marginally significant, due to the small sample size of this analysis. Species richness was significantly positively correlated with mean NRI ($r = 0.79$, $P = 0.04$). No other significant correlations with species richness or CV_A were found for any other variable.

4.4.5 Phylogenetic Signal of Abundance and Relationship with Species' Regional Characteristics

When testing if abundance was conserved across the phylogeny, neither of the two measures of local abundance of species across transects, mean species abundance and transect frequency, were significantly different from random expectation ($P = 0.45$ and $P = 0.09$, respectively). In contrast, both measures of regional abundance were significantly conserved across the tree (range occupancy: $P = 0.01$; range extent: $P = 0.01$). I found significant positive correlations between elevation range median and mean species abundance, range occupancy, and range extent (Table 4.4). In addition, range occupancy and range extent were strongly positively correlated. No other correlations were significant (Table 4.4).

4.5   DISCUSSION

4.5.1   Patterns of Phylogenetic Structure

Assemblages of Psychotrieae in Costa Rican wet forests are significantly phylogenetically structured.   When all transects are considered together, species are phylogenetically overdispersed, as measured by the nearest taxon index (NTI < 0). This result indicates that species' closest relatives tend not to be found in the same assemblage. There was no evidence of tree-wide phylogenetic structuring at this scale (NRI ≈ 0), indicating that the higher-level clade membership of the species in a given assemblage does not differ from random. These results are similar to other work on a single clade of plants, such as the oak assemblage in Florida (Cavender-Bares et al. 2004), where co-occurring oaks were found to be phylogenetically overdispersed. Similarly, Slingsby and Verboom (2006) found that co-occurring South African sedge species were overdispersed. Studies on single lineages of animals have also found similar results in lizards (Losos et al. 2003), birds (Lovette and Hochachka 2006), and fishes (Helmus et al. 2007).   However, my results contrast with another study that examined the phylogenetic structure of a specific lineage across a large geographic area. Vamosi and Vamosi (2007) examined the phylogenetic structure of predaceous diving beetles (Dytiscidae: Coleoptera) across 53 lakes in Alberta, Canada. Using a supertree of phylogenetic relationships, museum collection records of occurrence, and the same statistical analyses as I have used here, they found that beetle assemblages were on average phylogenetically clustered (NRI > 0).   The authors interpreted this trend as evidence that habitat filtering plays a stronger role in regulating community assembly in these beetles than competition.

Assemblages differed significantly in mean phylogenetic structure measures (NRI, NTI, and PD), as well as in mean species richness (Table 4.3). Interestingly, the three assemblages showing significantly overdispersed phylogenetic structure (LS, LC, and SG; Fig. 4.4) had the three lowest mean species richness values (Table 4.1). If limiting similarity is indeed leading to overdispersion, then this suggests that species are being excluded from these habitats that may otherwise exist there. To my knowledge, there have not been previous studies using replicated samples from multiple sites to compare phylogenetic structure of a clade of species. Other studies have looked at multiple sites across a large geographic range (Slingsby and Verboom 2006, Vamosi and Vamosi 2007), but they did not have replicate measures at sites, so no statistical inferences could be made regarding the differences in sites. The lack of consistency in both the sign and magnitude of phylogenetic structure suggests that different process may be most important in regulating co-occurrence at each site. However, without knowing the various ways in which one site is different from another, it is hard to determine the specific causes of the differences in structure. Furthermore, the regional species pool at each site may itself be phylogenetically structured in a manner that influences the phylogenetic structure of these assemblages.

4.5.2 The Relationship between Regional Characteristics and Local Abundance of Species

Abundance is not a phylogenetically conserved trait among Costa Rican wet forest Psychotrieae. Mean abundance and frequency of occurrence on transects is positively correlated but only weakly, and the relationship was not significant after accounting for multiple comparisons. However, both range occupancy and range extent are significantly conserved, indicating that closely-related species are similar in their general distributional attributes. Not

surprisingly, range occupancy and range extent are strongly positively correlated, as these two measures both quantify aspects of geographic range size (although they need not be correlated and can be influenced by different attributes of geographic ranges, see Paul, Chapter 3). There was no relationship between the local abundance of species, measured either as a species' mean abundance across transects or frequency on transects, and any of the regional range size estimates of species. This is in contrast to the general trend of a positive relationship between local abundance and range size found for many organisms (Gaston 1996). I did find that species' mean abundance and median range elevation were positively correlated, indicating that higher elevation species were found at higher average abundance. Furthermore, on average higher elevation species have significantly smaller geographic range sizes (Table 4.2), and these species tend to be more abundant. These patterns help explain why there is not a relationship between local abundance and geographic range sizes in these species. However, understanding why the higher elevation species reach higher mean abundances is more of a mystery.

4.5.3 Phylogenetic Structure of Assemblages and Variation in Abundance

Taken in concert, the patterns of co-occurrence and abundance of Mesoamerican Psychotrieae paint a complicated picture of how phylogenetic relatedness, species richness, and variation in abundance interact and change with the scale of analysis. Abundance was not phylogenetically conserved among species. However at the local scale of the La Selva assemblage, when species in a transect are more closely related, their abundances are also more similar, as predicted by the simple model of species' abundances being a product of their net phylogenetic and ecological similarity via niche conservatism. The sample size for La Selva was much larger compared to other assemblages and hence I had more power to detect relationships.

The relationships between CV_A and the phylogenetic structure metrics were negative and intermediate in strength (for NRI, $r = -0.182$ and for NTI $r = -0.397$), indicating that as the variation in abundance among co-occurring species increases, the relatedness of species decreases. Species richness and CV_A were positively correlated ($r = 0.482$), meaning that as the species richness of a transect increases, the variation in abundance among species increases as well. However, the mean standardized abundance of species was also positively correlated with both CV_A and species richness at La Selva, indicating that as the number of species and CV_A increases, species on average become more abundant. On the whole, these results indicate that the Psychotrieae assemblage at La Selva is on average phylogenetically overdispersed, but there is considerable variation among individual transects. Transects that are more phylogenetically clustered have species' abundances that are more similar as predicted by niche conservatism, yet on average these species' abundances are lower than their assemblage-wide average. These results suggest that when closely-related Psychotrieae species co-occur at the local scale, their abundances are impacted, likely because of increased competitive pressures, while on transects where species are less related than expected by chance, competition may be relaxed, as predicted by limiting similarity, and on average species' respond with increased abundances in relation to their assemblage-wide average abundance. Hence, phylogenetic relatedness does impact the abundance of Psychotrieae species in La Selva, but the effects are subtle and nuanced.

When examined across assemblages, the relationship between phylogenetic structure, CV_A, and species richness takes an entirely different form. At this scale, when assemblages contain species that are less related, CV_A among those species is small, and when assemblages contain species that are more related, the variation in abundance is large (for both NRI and NTI; Fig. 4.5). These trends are opposite the results found at the local scale in La Selva. Furthermore,

NRI is positively correlated with species richness, also an opposite result to La Selva. Clearly, either different processes are responsible for these contradictory trends, or similar processes whose results are borne out differently across spatial scales. At the local scale, my results suggest that species that are similar phylogenetically are also similar in their abundances, pointing to the importance of niche conservatism at this scale (Fig. 4.1, upper half). At the regional scale, when phylogenetically similar species make up an assemblage, species are more divergent in their abundances than assemblages of less-related species. This pattern, of a higher CV_A associated with phylogenetic clustering, and a lower CV_A associated with phylogenetic overdispersion, is predicted if limiting similarity is the paramount processes structuring relative abundances (Fig. 4.1, lower half).

Perhaps the patterns at the assemblage scale point to a longer-term equilibrium among species within an assemblage. An equilibrium could be produced by generations of competition against a background of niche conservatism (e.g., Lovette and Hochachka 2006). In phylogenetically clustered assemblages, closely related species co-occur due to habitat filtering (and maybe also the geography of speciation, see below), yet these species' abundances are regulated by competition for resources. Although these assemblages are phylogenetically clustered, species are not identical in their traits, and the long-term consequences of competitive interactions are borne out in the variable abundances of species at this scale (see Fig. 4.1, lower left). In contrast, phylogenetically overdispersed assemblages include species that are divergent in their ecological niches, and therefore the resources and habitats on which they specialize. These phylogenetically overdispersed species are relatively released from the negative effects of competition with other clade members and species abundances are more strongly regulated by the quality and quantity of the resources available at a given site. As a result, species

abundances are more similar, and hence CV_A is lower, because species are on average responding similarly to general site conditions (Fig. 4.1, lower right).

Although the line of reasoning presented above fits with my proposed model of the phylogenetic structure-abundance relationship (Fig. 4.1), if and how these species actually compete for resources cannot be addressed with my data and this model remains untested. Finally, it is intriguing to note that the positive relationship between species richness and phylogenetic clustering (NRI) found in the across assemblage analysis seems to support the integration of niche conservatism and limiting similarity at this scale. MacArthur and Levins (1967) list three ways that more species can be packed into communities regulated by limiting similarity. Species richness can be increased by 1) an increase in potential niche spaces (the niche dimensionality), 2) a decrease in species' niche breadths, or 3) *species' carrying capacities being uniform* (MacArthur and Levins, 1967, pg. 381). In Psychotrieae assemblages where species richness is higher, species are also more closely related. According to niche conservatism, species' ecological niches, and therefore their equilibrial carrying capacities, should be very similar as well. Hence, the higher species richness of phylogenetically clustered assemblages can potentially be explained by the interplay of phylogenetic niche conservatism and limiting similarity at this scale.

4.5.4   Consideration of the Processes Leading to Patterns of Phylogenetic Structure

The idea that congeneric species should be each other's strongest competitors dates back to Darwin (1859) and has been a central focus of phylogenetic structuring studies to date (e.g., Cavernder-Bares et al. 2004). However, some assumptions must be met if the predictive framework is to hold power. First, if close relatives are indeed stronger competitors than non-

related species, this implies that the related species have not diverged in some key trait that is important to resource use. However, even minimally-genetically diverged species can show considerable quantitative trait divergence (e.g., Yang et al. 1996), throwing into question the general validity of this assumption. Hence, even close relatives that differ very little in most traits may differ substantially in a key trait. This is particularly likely when speciation is driven by ecological factors and can lead to stable coexistence and weakened competition, despite overall genetic similarity. Under these circumstances, species' co-occurrence and local abundances may be driven by factors other than competition for resources, and as a result, the co-occurrence or variation in abundance among species may not be influenced by phylogenetic distance. Second, there are circumstances where close relatives are not expected to be each other's strongest competitors. In the case of convergence of trait values, relatively distantly related species may co-occur and compete for shared resources (Webb et al. 2002). Such a situation was found for assemblages of oaks (*Quercus* spp.) in Northern Florida (Cavender-Bares et al. 2004), where co-occurring species were phylogenetically overdispersed, yet species were similar in many trait values, due to convergence to a similar habitat. Under these circumstances, the phylogenetic structure of assemblages is overdispersed, yet the variation in abundance among species is expected to be high given the similarity in traits among species and assuming that there is a hierarchy of competitive abilities among species that is reflected in species' relative abundances. This prediction is opposite to the view championed by Darwin, yet follows the same logic and only differs in what species are expected to be each other's strongest competitors.

Although Darwin's musings regarding the competitive interactions of congeners and the predictions of limiting similarity have dominated the theoretical framework of phylogenetic structuring thus far, integrating our knowledge of the process and timing of speciation into this

framework may be beneficial. Since all species are the result of a process of divergence from ancestral entities, understanding how different modes of divergence potentially impact trait distributions vital to coexistence is of central importance. McPeek (2007) presents a model of species competing in a metacommunity in which speciation can produce new species that fall on a gradient from ecologically similar to ecologically divergent from their progenitor. If new species are ecologically divergent from their progenitor species, they can coexist in the same community because they specialize on different components of the niche space, and hence competitive exclusion is relaxed. Under these conditions, the phylogenetic structure of such a community would be clustered, yet the assumption that close relatives are each other's closest competitors would be violated. On the other hand, if species are ecologically very similar to their progenitor, they can still co-occur for long periods of time, despite the fact they would not be predicted to coexist in equilibrium communities (McPeek 2007). This is because when the competitive differences between species are slight (such as when speciation is recent and driven by non-ecological processes, e.g. sexual selection), inferior competitors are predicted to be driven from communities at a glacial pace, leading to a large number of "transient" minimally-diverged species in communities (McPeek 2007). Finally, if speciation is recent and is not allopatric, and if species require ample time to expand their ranges and become members of communities distant from where there originated, as is potentially the case in *Psychotria* (Paul, chapter 3), assemblages might be phylogenetically clustered because of the geography of speciation. In the two highest elevation assemblages in this study (MVCF and MVEB), assemblages were significantly phylogenetically clustered (Fig 4.5), and the species represented in these assemblages include many recently diverged *Heteropsychotria* species (Fig. 4.2). Furthermore, these higher elevation species tend to have smaller geographic ranges sizes (Table

4.4) meaning that they are less likely to be found as members of distant assemblages. In contrast, if speciation is allopatric, the most closely related species, by definition, will not be found in the same assemblage (assuming the delimitation of assemblages follows the same geographic barriers that led to speciation), and hence assemblages of species in which allopatric speciation is commonplace are expected to be overdispersed, barring a great deal of secondary contact of sister species over time. In both of these situations, the geography of speciation, rather than the processes of niche conservatism or limiting similarity *per se*, would be the causes of the observed phylogenetic structure (but see Wiens 2004 on how phylogenetic niche conservatism can promote speciation). Accounting for the dominant mode of speciation and the recency of speciation in assemblages in which phylogenetic structure is being measured will provide a more complete picture of the potential causal factors of phylogenetic structure.

4.5.5   Conclusions

The research presented here demonstrates that Psychotrieae assemblages in Costa Rica are significantly phylogenetically structured. Different patterns emerge when all transects are analyzed together versus comparing across transects, and assemblages differ significantly from one another in their phylogenetic structure. This study may be the first to quantify phylogenetic structure in replicated sample units from a number of different sample locations. I found that interesting patterns emerge when, for example, the relationship between phylogenetic structure and variation in abundance is compared within an assemblage versus across a set of assemblages. In order to broaden our understanding of the processes that drive patterns of phylogenetic community assembly, researchers will need to relate phylogenetic structure of communities to a larger array of potential causal factors by assessing structure both within and across ecological

gradients. I found that while abundance is not a conserved trait when examined across species, abundance did nonetheless exhibit significant relationships with phylogenetic structure. The impact of phylogeny on the variation in abundance of species was subtle, yet I found patterns that are consistent with an interplay between niche conservatism and limiting similarity at both the local and across assemblage scales. This study represents one of the first attempts to explicitly link abundance and phylogenetic structure. Hopefully more researchers will bring data to bear on this complex interaction, and we can gain a better understanding of how phylogenetic structure and abundance interact. Furthermore, investigations of experimental communities in which the phylogenetic relatedness of co-occurring species can be manipulated in concert with external variables (e.g., strength of competition, the regional species pool, available resources) will undoubtedly provide new insight to the growing study of phylogenetic community ecology.

Table 4.1. Characteristics and descriptive statistics of the seven assemblages surveyed in this study. The assemblages are: La Selva (LS), Rara Aves (RA), San Luis (SL), Las Cruces (LC), San Gerardo (SG), Monteverde Cloud Forest Reserve (MVCF), and Monteverde Estacion Biologia (MVEB). [a]Holdridge life zone classification: TWF = Tropical wet forest, PMWF = Premontane wet forest, LMWF = Lower Montane Wet Forest. [b]Side of the continental divide in which a n assemblage is located. [c]Species richness of assemblage (total number of species found on transects). [d]The mean (± S.D.) number of species found per transect. [e] The number of species in Costa Rica that have elevational ranges overlaping with an assemblage (regional species pool) and the percent represented in transects.

Assemblage	Transects (N)	Elevation (m)	Life Zone[a]	Geography[b]	Num. Spp.[c]	Mean (± S.D.) Spp.[d]	Reg. Spp. Pool (%)[e]	
LS	155	92	TWF	Caribbean	18	3.78 (±1.52)	53	(34%)
RA	14	640	TWF	Caribbean	11	6.00 (±1.47)	46	(24%)
SL	19	1170	PMWF	Pacific	9	4.33 (±1.41)	27	(33%)
LC	12	1200	PMWF	Pacific	5	3.67 (±0.98)	35	(14%)
SG	13	1235	LMWF	Caribbean	6	3.08 (±0.76)	34	(17%)
MVCF	12	1523	LMWF	Pacific	12	4.83 (±1.47)	37	(32%)
MVEB	15	1694	LMWF	Pacific	12	4.86 (±1.41)	24	(50%)

Table 4.2. Pearson correlation coefficients (r) for mean values of CV_A, species richness, standardized abundance (StdAbun), NRI, NTI, and PD for the all transect and within-assemblage analyses. Correlations in bold were significant after sequential Bonferroni correction (N = 72 comparisons; starting alpha < 0.00069). Values with an asterisk indicate a P < 0.01.

	Richness	StdAbun	NRI	NTI	PD
All Transects (N = 205)					
CV_A	**0.43**	**0.33**	0.02	-0.10	**0.39**
Richness	---	**0.28**	0.21	-0.05	**0.78**
La Selva (N = 122)					
CV_A	**0.48**	**0.45**	**-0.18**	**-0.40**	-0.18
Richness		**0.30**	-0.04	-0.61	**0.92**
Rara Aves (N = 14)					
CV_A	0.22	0.43	-0.06	-0.02	0.18
Richness	---	-0.08	-0.31	-0.18	**0.86**
San Luis (N = 18)					
CV_A	0.22	0.16	0.05	0.08	0.14
Richness	---	-0.06	-0.22	0.54	**0.89**
Las Cruces (N = 12)					
CV_A	0.36	0.07	0.28	0.30	0.15
Richness	---	0.63	0.73*	0.60	**0.85**

Table 4.2. (continued)

	Richness	StdAbun	NRI	NTI	PD
San Gerardo (N = 13)					
CV_A	0.47	0.17	0.29	0.06	0.42
Richness	---	0.50	0.74*	0.39	0.72*
Monteverde EB (N = 14)					
CV_A	-0.04	-0.06	-0.03	-0.09	0.01
Richness	---	-0.29	0.68*	0.77*	0.24
Monteverde CF (N = 12)					
CV_A	0.27	0.02	0.17	0.13	0.23
Richness	---	0.05	0.05	0.49	0.74*

Table 4.3. One-way analysis of variance results for NRI, NTI, PD, and mean species richness between assemblages. [a]Analysis on arcsin-squareroot transformed data.

Variable	Source	DF	Sum of Squares	Mean Square	F Value	Pr > F	R-Square
NRI	Assemblage	6	44.21	7.37	14.29	0.0001	0.48
	Error	90	46.39	0.52			
NRI	Assemblage	6	63.89	10.65	19.00	0.0001	0.56
	Error	90	50.44	0.56			
PD[a]	Assemblage	6	0.14	0.02	4.63	0.0004	0.24
	Error	90	0.45	0.01			
Richness	Assemblage	6	70.06	11.68	5.96	0.0001	0.28
	Error	90	176.18	1.96			

Table 4.4. Pearson correlation coefficients (*r*) of mean abundance per transect (Mabun), frequency on transects (Freq), range occupancy, range extent, elevation range, and elevation median. Both Mabun and Freq were log transformed prior to analysis. Correlations in bold were significant after sequential Bonferroni correction (N = 14 comparisons; starting alpha < 0.0036). Values with an asterisk indicate a P < 0.01.

Mean	Freq	Occupancy	Extent	Elev. Range	Elev. Med.
Abundance	0.42*	-0.18	-0.13	0.22	**0.49**
Frequency	---	-0.19	-0.15	-0.25	0.31
Occupancy	---	---	**0.76**	0.15	**-0.44**
Extent	---	---	---	0.15	**-0.44**

	Phylogenetic Structure	
	<u>**Clustered**</u>	<u>**Overdispersed**</u>
<u>**Process Regulating Abundance**</u>		
Niche Conservatism Paramount	CV_A Low	CV_A High
Limiting Similarity Paramount	CV_A High	CV_A Low

Figure 4.1. The expected effect of phylogenetic structure on species' abundances depends on which process, ecological niche conservatism or limiting similarity, is the strongest factor governing abundances. Variation is expressed as the coefficient of variation in abundance of co-occurring species (CV_A). When <u>Ecological Niche Conservatism</u> is most important (i.e., abundance is a conserved 'trait' similar to, or the product of a set of other conserved phenotypic traits), variation in abundance among phylogenetically clustered species is expected to be low. However, if species are phylogenetically overdispersed (i.e., they differ in many traits), variation in abundance is expected to be high. The opposite pattern is predicted if <u>Limiting Similarity</u> primarily structures communities. Competitive interactions are predicted to regulate species' abundances, thus variation in abundance among phylogenetically clustered species are expected to be high, as these species share similar traits and are predicted to be strong competitors. When species are phylogenetically overdispersed, variation in abundance is expected to be low, because competition is relaxed due to the divergent trait distributions of co-occurring species. Under this scenario, abundance may be regulated by general site conditions, and species will respond similarly to good or poor conditions. This diagram assumes traits are phylogenetically conserved, rather than a product of trait convergence by phylogenetically distant species.
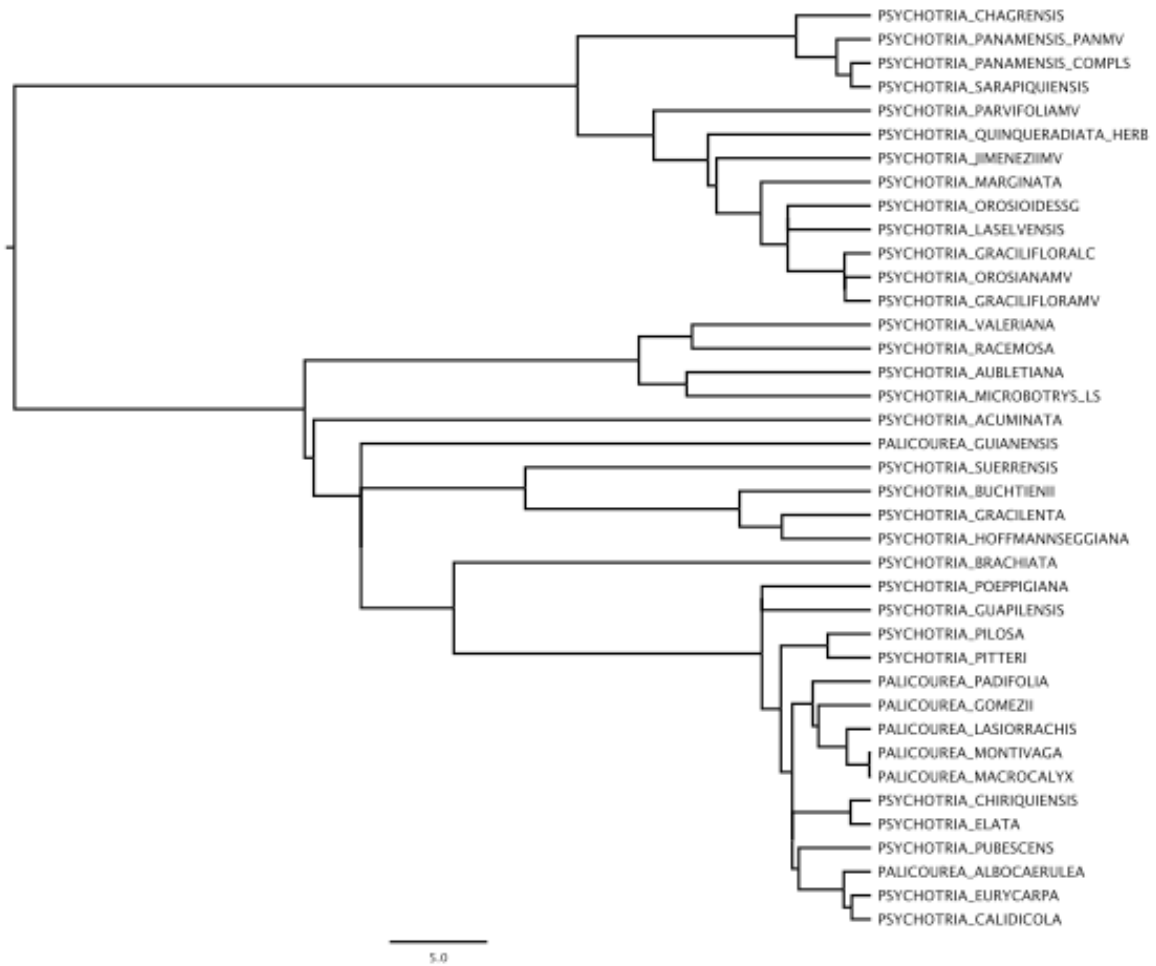
Figure 4.2. An ultrametric tree showing the phylogenetic relationships of the species found on transects, as inferred by maximum likelihood. This tree is pruned from a much larger tree (N > 300 taxa) that was used to get accurate estimates of phylogenetic relationships and branch lengths. Scale bar shows time in millions of years.
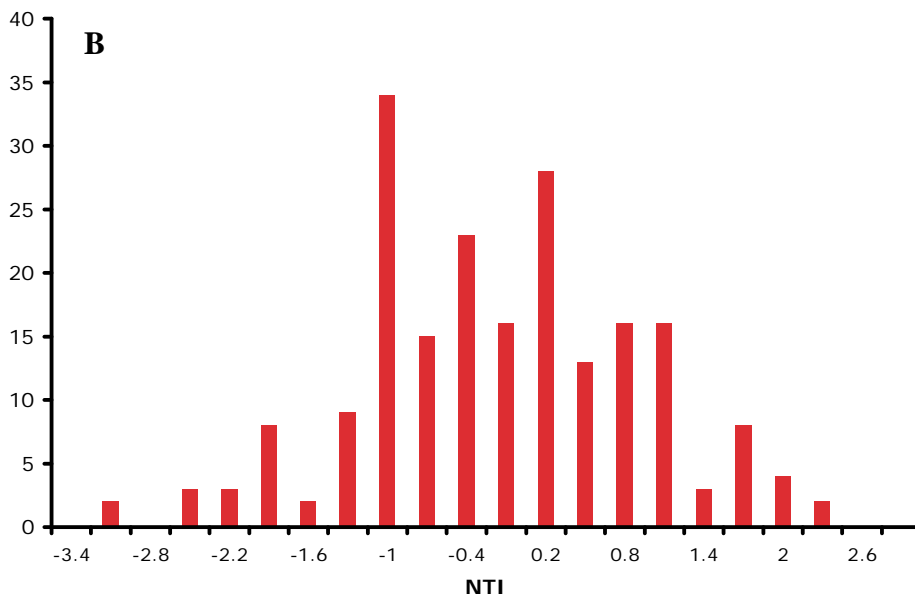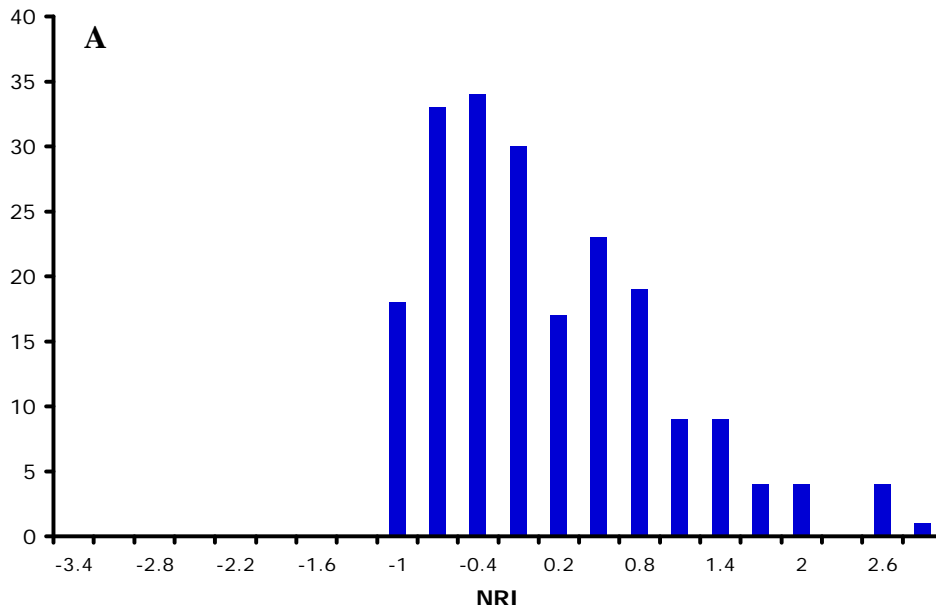
Figure 4.3. Frequency histograms showing the distribution of net relatedness index (NRI; A) values and nearest taxon index (NTI; B) values across all transects (N = 205).
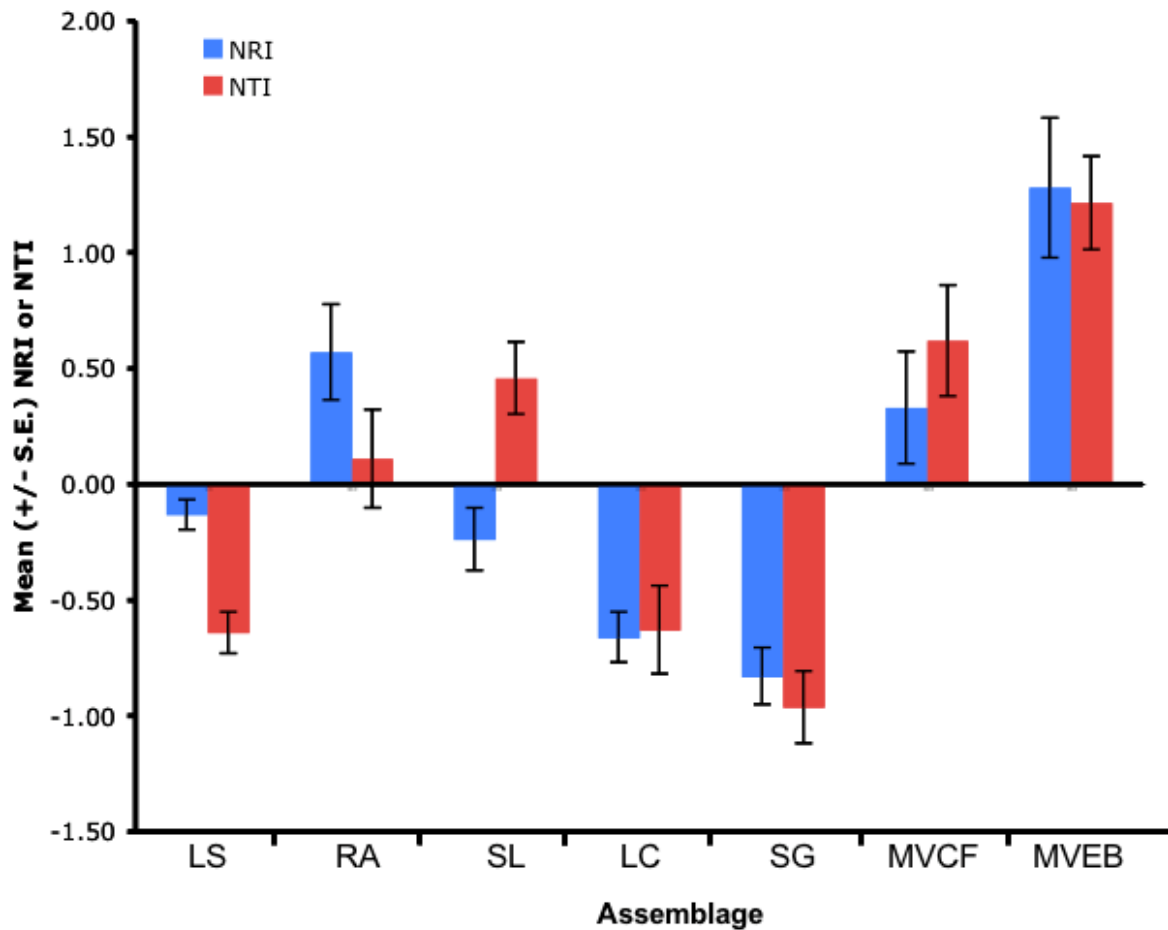
Figure 4.4. The phylogenetic structure of seven Psychotrieae assemblages. Assemblages are ordered from the lowest elevation site to the highest elevation site. La Selva (LS): NRI = -0.13, df = 121, P = 0.048; NTI = -0.64, df = 121, P < 0.001. Rara Aves (RA): NRI = 0.57, df = 13, P = 0.02; NTI = 0.11, df = 13, P = 0.61. San Luis (SL): NRI = -0.24, df = 17, P = 0.09; NTI = 0.45, df = 17, P = 0.009. Las Cruces (LC): NRI = -0.66, df = 11, P < 0.001; NTI = -0.63, df = 11, P = 0.007. San Gerardo (SG): NRI = -0.83, df = 12, P < 0.001; NTI = -0.96, df = 12, P < 0.001. Monteverde Cloud Forest (MVCF): NRI = 0.33, df = 11, P = 0.20; NTI = 0.62, df = 12, P = 0.025. Monteverde Estacion Biologia (MVEB): NRI = 1.28, df = 13, P < 0.001; NTI = 1.22, df = 13, P < 0.001.
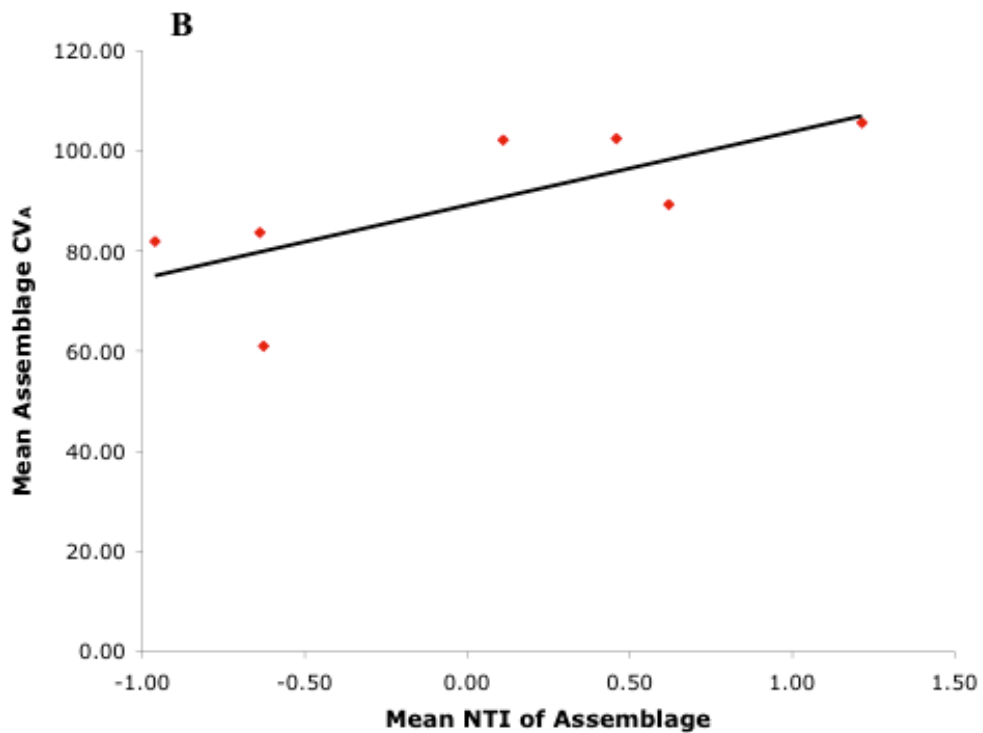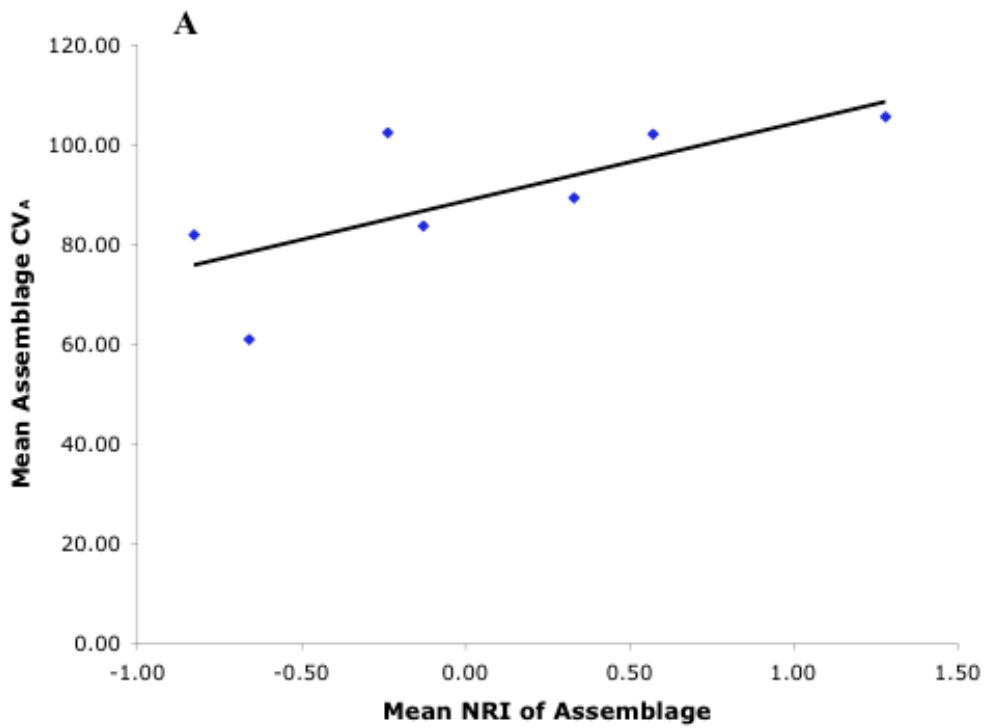
Figure 4.5. The relationship between mean assemblage CV_A and mean assemblage NRI ($r = 0.73$, df = 6, P = 0.065; A) and NTI ($r = 0.74$, df = 6, P = 0.057; B).

5.0   CONCLUSIONS AND FUTURE DIRECTIONS

This dissertation demonstrates the importance of taking evolutionary history into account when investigating the variation in abundance and distribution of plants. I demonstrate that species age can explain a significant proportion of the variation in range sizes of species, and that phylogenetic relatedness of co-occurring species has subtle but measurable effects on the local abundance of plant species. By taking an explicitly phylogenetic approach to addressing classic questions in ecology, this work sheds new light on the potential causal mechanisms behind variation in species range sizes, such as a time-for-dispersal effect. Furthermore, this dissertation sets the stage for a multitude of future research directions that will address the impact of phylogeny on ecological patterns and processes.

To my knowledge, this dissertation includes the first two tests of age and area hypothesis using tropical plant species. This is surprising in light of the fact that John Willis' hypothesis was developed based on his observations of the tropical flora of Ceylon (now Sri Lanka), and even his critics acknowledged that age-and-area might be more important in the tropics (Gleason 1924), which were seen as stable and relatively homogenous. To date, tests of age and area have focused on large clades of animals that might be expected to have high dispersal abilities and have many ecologically unique clades nested within them (e.g., mammals, Jones et al. 2005; birds, Gaston and Blackburn 1997). Tests of specific clades of animals have found many different patterns (e.g., in birds, Webb and Gaston 2000), indicating that combining across

numerous clades will likely swamp out interesting patterns. Unsurprisingly, these studies have failed to find any strong or consistent pattern of an age and area relationship. Other tests using animals have been largely based on fossil assemblages of mulloscs (e.g., Miller 1997) and have found results that are at least partially consistent with age and area. My tests with two tropical plant lineages (chapters 2 and 3) both find significant support for the age and area hypothesis, and show that species age can explain as much as a third of the variation in geographic range sizes among species. These results are fairly remarkable given the age of these clades (~ 20 - 40 millions years old) and the list of potential factors that can obfuscate the simple pattern predicted by age and area. The question remains if the signal of species age will only be detected in tropical plants that have limited dispersal abilities, or if these results might apply more generally. Of course, given the immense diversity of angiosperms in the tropics (Gentry 1982), even if age and area only applies to tropical understory shrubs, this could potentially include hundreds of lineages.

My work also demonstrates the importance of carefully considering the assumptions of age and area (chapter 2), as well as defining range sizes in ways that are most likely to be informative to the question at hand (chapter 3). Specifically, by examining two different range size metrics (occupancy and extent), I was able to demonstrate that the relationship between species age and range size is likely a product of temporal dispersal limitation, and that species age may be expected to impact range extents more than range occupancies. Furthermore, by addressing the reality that species have ecological limitations to where they can live, I conducted a better test of age and area by predicting where species should potentially be found (using species distribution modeling) and asking if young species had have colonized less of their potential ranges than old species. Finally, my work draws a clear link between Willis'

hypothesis and the currently hotly debated predictions of neutral theory in ecology (Hubbell 2001), bringing age and area into a modern light, and helping to renew interest in the macroecological work of the beginning of the 20th century. Future students of ecology and evolution will greatly benefit by carefully reading these works and discovering the many yet untested and insightful ideas they present.

Charles Elton would likely be happy with the progress that has been made in understanding how the relatedness of species affects community structure, since this was the central interest of his work on genus-to-species ratios in communities (Elton 1946). My work on the phylogenetic structure of Psychotrieae assemblages (chapter 4) adds insight to the renewed interest in phylogenetic community structure. I found that Psychotrieae assemblages are on the whole phylogenetically overdispersed, indicating that co-occurring species are less related than expected by chance. This result is similar to the results of the few other studies that examined this question in single plant lineages (e.g., Cavender-Bares 2004, Slingsby and Verboom 2006). Furthermore, my work takes a unique step forward by linking the phylogenetic structure of assemblages to the variation in abundance among species. I show that at the local scale, closely-related species are similar in their abundances, as predicted if niche conservatism (Wiens and Graham 2005) is important in structuring communities. However, I also found that these co-occurring, close relatives are on average depressed in their abundances compared to when they co-occur with less related species, suggesting that limiting similarity is also playing a role in regulating the relative abundances of species in these assemblages. Finally, I found that the scale of analysis had strong effects of the patterns of phylogenetic structuring and its relation to variation in abundance. The seven assemblages I examined differed significantly in their phylogenetic structure. When I examined the relationship between variation in abundance and

phylogenetic structure across assemblages, I found that the pattern was opposite of what I found at the local scale. I provide some speculative comments on why this may be the case, but definitive answers will have to be left for future research.

The integration of ecology and evolution is in its infancy and will likely see many new and exciting directions in the near future. As more genomes are sequenced and we gain a better understanding of how organism's phenotypes are built from their genetic material, we will have the opportunity to make a stronger link between the processes that drive ecological patterns but also shape phenotypes and guide the course of evolution. Similarly, as we gain a better understanding of how the evolutionary relationships of organisms influence their ecological interactions, and how evolutionary history shapes patterns of diversity, we will have the opportunity to truly integrate ecology and evolution into a more complete and thorough explanation of how organisms arise, interact, and build the remarkable patterns of diversity, distribution, and abundance that we find on Earth.

5.1   EPILOGUE

John Willis' ideas obviously had a strong influence on development of this dissertation. Recently, Willis' age and area hypothesis has been written off as nothing more than quaint anachronism (Brown et al. 1996); this dissertation challenges and rejects that narrow view. In fact, I would argue that Willis' work is currently receiving more support than is even acknowledged. The focus of this dissertation has been on one half of Willis' work (1922); specifically on the influence of species age on range sizes. However, the other half of his book was an argument for another beautifully simple prediction: older groups of organisms should

have more species than younger groups. Two recent studies support this prediction. Brown and McPeek (2007) found that across the animals, species richness of clades is best explained by the age of clades, rather than differential rates of diversification. Another study on the vast diversity of the beetles (Coleoptera) found that the diversity of beetles is best explained by the clades very old age and high survival of lineages over time (Hunt et al. 2007). Finally, every day, phylogeneticists like me use the Yule pure-birth model of diversification to build starting trees for Bayesian analyses and various other aspects of our work. As noted in chapter 2, Yule (1925) based his pure-birth model of diversification on the ideas that John Willis presented in Age and Area (1922). Clearly, Willis' ideas are expanding their influence as they age.

SPECIES AND ACCESSION NUMBER OF SPICIMENS USED IN THE PIPER STUDY

Table A1.1. Species and GenBank accession numbers used in this study.

Species	GenBank Accession Numbers
Houttuynia cordata	AF275211
Macropiper excelsum	AF275193
Macropiper hooglandii	AF275192
Macropiper melchior	AF275191
Peperomia elongata	AF275213
Piper aduncum	AF275159
Piper aduncum2	AF275158
Piper aduncum3	AF275157
Piper albispicum	AY572317
Piper albozonatum	AY326195
Piper amalago	AF275186
Piper amoenum	AF275160

Table A1.1 (continued)

Species	GenBank Accession Numbers
Piper appendiculatum	AY326196
Piper arborescens	AF275202
Piper arboreum	AF275180
Piper arboricola	AY572319
Piper archeri	AF275178
Piper arieianum	AF275163
Piper atrospicum	AY572318
Piper augustum	AF275165
Piper auritum	AF275175
Piper bartlingianum	AF275183
Piper basilobatum	AY326197
Piper bavinum	AF275199
Piper betle	AF275201
Piper boehmeriifolium	AF275204
Piper brachypodon	AY326198
Piper brevicuspe	AY572321
Piper brevipedicellatum	AF275189
Piper cajambrense	AY326199
Piper caninum	AF275195
Piper capense	AY326200
Piper cararense	AY326201

Table A1.1 (continued)

Species	GenBank Accession Numbers
Piper cavendishioides	AF275153
Piper celtidiforme	AF275205
Piper chuarense	AY326202
Piper cihuatlanense	AF275187
Piper cinereum	AF275190
Piper cocornanum	AY326203
Piper colligatispicum	AY326204
Piper confertinodum	AF275166
Piper cordatilimbum	AY572323
Piper darienense	AF275181
Piper decumanum	AF275203
Piper densum	AY615963
Piper filistilum	AF275155
Piper flagellicuspe	AF275154
Piper friedrichsthalii	AY326205
Piper garagaranum	AF275162
Piper gesnerioides	AY326206
Piper gymnostachyum	AY572325
Piper hartwegianum	AY326207
Piper hernandii	AY572324
Piper hispidum	AF275156

Table A1.1 (continued)

Species	GenBank Accession Numbers
Piper hymenophyllum	AY572327
Piper imperiale	AF275176
Piper korthalsii	AF275208
Piper laosanum	AY572326
Piper lolot	AY326208
Piper longispicum	AY326209
Piper marequitense	AY326210
Piper marginatum	AY326211
Piper medinillifolium	AY667455
Piper methysticum	AF275194
Piper michelianum	AF275188
Piper multiplinervium	AF275168
Piper munchanum	AF275164
Piper myrmecophilum	AY572328
Piper nigrum	AF275198
Piper nigrum2	AF275197
Piper obovatum	AY326212
Piper ottoniifolium	AY326213
Piper oxystachyum	AF275152
Piper parvulum	AF275167
Piper pedunculatum	AY326214

Species	GenBank Accession Numbers
Piper peltatum	AF275171
Piper peltatum2	AF275170
Piper peltatum3	AF275169
Piper penninerve	AF275206
Piper perpusillum	AY326215
Piper phytolaccifolium	AY326216
Piper pierrei	AF275200
Piper pilibracteum	AY768829
Piper pulchrum	AF275177
Piper reticulatum	AF275185
Piper reticulatum2	AF275184
Piper retrofractum	AF275196
Piper sabaletasanum	AY326217
Piper schuppii	AY326218
Piper sorsogonum	AY572320
Piper sp1 maj674	AY326219
Piper sp2 maj689	AY326230
Piper spoliatum	AF275179
Piper subglabribracteatum	AY326220
Piper subpedale	AF275161
Piper terryae	AY326221

Table A1.1 (continued)

Species	GenBank Accession Numbers
Piper tomas-albertoi	AY326222
Piper toppingii	AY572322
Piper trianae	AY326224
Piper tricuspe	AY326225
Piper tuberculatum	AY326223
Piper ubatubense	AF275182
Piper umbellatum	AF275174
Piper umbellatum2	AF275173
Piper umbellatum3	AF275172
Piper umbricola	AY326226
Piper unispicatum	AY326227
Piper urdanetanum	AF275207
Piper villosum	AY326228
Piper yanaconasense	AY326229
Sarcorhachis naranjoana	AF275210
Sarcorhachis sydowii	AF275209
Saururus cernuus	AF275212

APPENDIX B

DETAILED METHODS FOR CHAPTER 3

B.1 DETAILS OF SPECIES DISTRIBUTION MODELING

I included 20 layers (altitude and 19 'bioclimatic' variables) in the distribution modeling (WorldClim database, www.worldclim.org). I used 30' (~ 0.83-0.83 km^2 resolution) layers of Mesoamerica and South America for the following variables: altitude, annual mean temp., mean diurnal temp. range, isothermality, temp. seasonality, max. temp. of the warmest month, min. temp. of the coldest month, temp. annual range, mean temp. of wettest quarter, mean temp. of driest quarter, mean temp. warmest quarter, mean temp. coldest quarter, annual precip., precip. of wettest month, precip. of driest moth, precip. seasonality, precip. of wettest quarter, precip of driest quarter, precip. of warmest quarter, and precip. of coldest quarter. Further information on the specifics of the variables and the data formats are available at the WorldClim website. For each species, all geo-referenced collections in the W3 Tropicos database for the *Psychotria* of interest were input in MAXent for analysis. I did extensive model testing, following similar protocols to those outlined in Phillips et al. (2006). For a subset of species, I ran MAXent with variable proportions of training and testing data (used to detect omission rates and look for violations of model assumptions), varying combinations of bioclimatic variables, and varying

110

combinations of 'features types' (linear, quadratic, product, hinge and threshold; see Phillips et al. 2006 for a detailed description). Essentially, when sufficient samples are available, higher-order feature types allow the model distribution to be constrained in more ways, and ultimately provide a better estimate of the maximum entropy distribution. After initial testing, for each species I modeled the predicted distribution using 50% of the records for testing, all bioclimatic variables, and the default feature type settings (restricts distribution modeling of species with smaller sample sizes: N = 2-9, linear only; N = 10-79, linear + quadratic; N = ≥80, linear + quadratic + product; N = ≥80 threshold; N ≥15, hinge). Finally, I modeled each species using all samples for training and projection, all bioclimatic variables, and the default feature and iteration settings. In addition, for species with large samples sizes, I also modeled the distribution using only linear features, to make them directly comparable to modeled results of species with small sample sizes. Some species had four geo-referenced collection records or less, and could not be reliably modeled.

B.2   DETAILS OF GEOGRAPHIC RANGE SIZE ESTIMATES

To calculate range occupancy and range extent from the collection records and the predicted distributions, I used the image analysis programs ImageJ (Rasband 1997) and Photoshop (Adobe 2005), and the geographic information system program DIVA-GIS (Hijmans et al. 2001). For predicted distributions, I imported grid files created by MAXent into DIVA-GIS and visualized the maps. All grid cells with a cumulative probability ≥ 1.0 (see MAXent manual for an explanation) were coded in one color and all other values were not shown. I then took a bitmap image of the full extent map and opened the image in Photoshop. I selected the

pixels in the appropriate color range and copied and pasted these pixels into a standard-sized new document, 'flattened' the image and saved it as a .tiff file. This file was opened in ImageJ, 'threshold' was applied, the pixels were selected, and then various attributes of the selection were recorded with ImageJ's measurement tool. Area of occupancy, for the R_O and R_O/P_O analyses, was calculated as the number of colored pixels. Range extent, for the R_E and R_E/P_E analyses, was calculated as Feret's diameter, the largest distance between two occupied (or predicted to be occupied) pixels. This entire procedure was repeated for each map image that I processed. The procedure was the same for realized ranges, except collection data points were imported as shape files in DIVA-GIS and given one color for analysis. The four bordering pixels for each collection record were also included to account for expected spatial autocorrelation of occupancy with recorded collections.

B.3   DETAILS OF MOLECULAR MARKER CHOICES

I choose *ITS* in order to have a marker with species level resolution that could be amplified from poor quality and degraded DNA, as is often found in herbarium samples. In addition, the two spacers of *ITS* (*ITS*1 and *ITS*2) can easily be amplified separately because they are flanked on one side of the conserved 18S gene and on the other by the conserved 5.8S gene (Baldwin et al. 1995). The use of *ITS* has been criticized by some authors because it has multiple copies in the genome and the complex nature of the mechanism behind the concerted evolution that the loci are proposed to experience is not well understood (Álvarez and Wendel 2003). However, suitable markers from the nuclear genome without the potential problem of paralogy are limited in number, and a well-studied region like *ITS* may actually have benefits over little

studied loci in which the possibility of paralogous copies is unknown (e.g., Feliner and Rosselló 2007). *ITS* has proven to be an effective species-level marker in *Psychotria* (Nepokroeff et al. 1999) as well as a number of other Rubiaceae taxa (e.g., Malcomber 2002). Finally, two recent studies using Rubiaceae taxa show that while multiple divergent *ITS* copies are present, the divergence does not transcend species boundaries (Razafimandimbison et al. 2004, Malcomber 2002) and thus *ITS* is an appropriate species-level phylogenetic marker. The cloning of one species also found minimally diverged copies that did not transcend species boundaries in phylogenetic analyses.

B.4   DETAILS OF LABORATORY TECHNIQUES AND PROTOCOLS

DNA was extracted from fresh material using Qiagen® DNeasy Plant Mini Kits. Herbarium samples were extracted following a standard CTAB protocol, except extractions were left for two or more weeks in isopropanol to allow the maximum amount of DNA to precipitate out of solution. Many of the herbarium extractions were cleaned prior to amplification using Qiagen® MiniElute columns. The *ITS* and *psbA* loci were amplified using PCR primarily in the CMNH biosystematics lab and secondarily in the molecular lab of S. Kalisz at the University of Pittsburgh. For *ITS*, I used the same primers as Nepokroeff et al. (1999; LEU, ITS4, ITS3B), as well as 5.8s (for reverse strand amplification of ITS1), and a new primer I developed, very similar to ITS3B, ITS3C (5'-3':GATATCTAGGCTCTCGCATC; for forward strand amplification of ITS2). For *psbA*, I used the primers used by Kress et al. (2005). Standard 50 µl reactions consisted of 35 µl sterile H_2O, 5 µl 10x buffer, 5 µl Mg, 1 µl BSA (10mg/ml), 1 µl DMSO, 1 µl DNTPs (10mM), 0.5 µl 5' 20 µM primer, 0.5 µl 3' 20 µM primer, 1 µl genomic

DNA. Standard *ITS* PCR amplification started with 94.0°C for 2 m, then 40 cycles of 94.0°C for 30 s, 48.0°C for 1 m, 72°C for 1 m, and a final elongation step at 72°C for 7 m. Standard *psbA* PCR amplification started with 94.0°C for 3 m, then 34 cycles of 94.0°C for 1 m, 51.4°C for 1 m, 72°C for 1 m, and a final elongation step at 72°C for 7 m. PCR reaction and amplification protocols were slightly modified for difficult to amplify taxa. All DNA sequencing was performed at the Davis Sequencing facility (www.davissequencing.com). Both strands were sequenced using the same primers used for amplification. Sequence strands were assembled using Sequencher 4.5 (Gene Codes 2005). I aligned sequences using ClustalX (Thompson et al. 1997) and made manual adjustments using Se-Al (Rambaut 1996).

B.5   DETAILS OF ALTERNATIVE APPROACHES TO PHYLOGENETIC INFERENCE

Molecular data were also analyzed with maximum parsimony using PAUP* (Swofford 2002) and maximum likelihood using GARLI (Zwickle 2006) to compare the tree topologies of these methods with those inferred by BEAST. I examined the following datasets: the full *ITS* alignment, the full combined *ITS-psbA* alignment (including taxa with missing data), the reduced *ITS-psbA* alignment (only species with both *ITS* and *psbA* data), and the reduced *ITS-psbA* alignment, but analyzing the *ITS* and *psbA* partition separately.   Parsimony searches used tree bisection-reconnection (TBR) branch swapping, 1000 random-addition replicates, with 10 optimal trees held for each replicate. The resulting saved trees were then used as starting trees for additional branch swapping to fill out the optimal tree space search. Parsimony statistics were recorded for the optimal set of trees. For likelihood analyses, the best-fit model of

nucleotide substitution was estimated using ModelTest (Posada and Crandall 1998) and the best model was chosen via the Akaike information criterion (AIC) test. I compared the resulting trees from each analysis. Support for phylogenetic trees was assessed with Bayesian posterior probabilities.

PSYCHOTRIA SUBGENUS *PSYCHOTRIA* SPECIES SPECIMENS SAMPLED FOR GENETIC DATA

Table C1.1 *Psychotria* subgenus *Psychotria* species specimens sampled for genetic data. Accession numbers for the complete *ITS*

locus, *ITS1*, *ITS2*, and *psbA* loci. Specimens for which a GenBank accession number has not yet been assigned are denoted by XX.

[a]Numbers are Missouri Botanical Garden accession numbers; except those starting with a letter are GenBank accession number.

Species	*ITS*(full)	*ITS1*	*ITS2*	*psbA*	Voucher Specimen[a]
Psychotria aguilarii Standl. & Steyerm.	XX	-	-	XX	04910704
Psychotria alfaroana Standl.	-	-	-	-	-
Psychotria bakeri Dwyer	-	-	-	-	-
Psychotria balancanensis C.W. Ham.	-	-	-	-	-

Table C1.1. (continued)

Species	*ITS*(full)	*ITS1*	*ITS2*	*psbA*	Voucher Specimen[a]
Psychotria boquetensis Dwyer	-	XX	XX	-	1172505
Psychotria calophylla Standl.	-	XX	XX	XX	5161725
Psychotria carthagenensis Jacq.	-	-	-	-	-
Psychotria cascajalensis C.W. Ham.	-	XX	XX	XX	2901097
Psychotria cerrocoloradoensis Dwyer ex C.M. Taylor	-	-	-	-	-
Psychotria chagrensis Standl.	AF072051	-	-	-	AF072051
	XX	-	-	XX	3006992
Psychotria chiriquina Standl.	-	-	-	-	-
Psychotria chitariana Dwyer ex C.W. Ham.	-	XX	XX	-	4297951
Psychotria clivorum Standl. & Steyerm.	XX	-	-	-	04933220
Psychotria cocosensis C.W. Ham.	-	XX	XX	XX	4297840
Psychotria costx: altorum (Standl. & Steyerm.) C.W. Ham.	XX	-	-	XX	3616103
Psychotria costx: costivenia	XX	-	-	XX	4060820
Psychotria dressleri (Dwyer) C.W. Ham.	-	-	-	-	-
Psychotria durilancifolia Dwyer	-	XX	XX	-	3613928

117

Table C1.1. (continued)

Species	*ITS*(full)	*ITS1*	*ITS2*	*psbA*	Voucher Specimen[a]
Psychotria dwyeri C.W. Ham	XX	-	-	XX	05072166
Psychotria erythrocarpa Schltdl.	XX	-	-	XX	3386117
Psychotria fendleri Standl.	-	XX	XX	XX	216000
Psychotria flava Oerst. ex Standl.	-	XX	XX	-	3099963
Psychotria fosteri C.W. Ham.	-	XX	XX	-	04966653
Psychotria fruticetorum Standl.	-	XX	XX	XX	2368260
Psychotria graciliflora Benth.	XX	-	-	XX	2998512
Psychotria grandis Sw.	-	XX	XX	-	3207997
Psychotria hamiltoniana C.M. Taylor	-	-	-	-	-
Psychotria hammelii Dwyer	-	-	-	-	-
Psychotria horizontalis Sw.	AF072047	-	-	-	AF072047
	EF667971	-	-	-	EF667971
Psychotria hornitensis Dwyer ex C.W. Ham.	-	-	-	-	-
Psychotria insignis Standl.	-	-	-	-	-
Psychotria insueta (Dwyer) C.W. Ham.	-	XX	XX	-	2981490

Species	*ITS*(full)	*ITS1*	*ITS2*	*psbA*	Voucher Specimen[a]
Psychotria jefensis Dwyer ex C.M. Taylor	-	XX	XX	XX	2785746
Psychotria jimenezii Standl.	XX	-	-	XX	04963228
Psychotria jinox: jinotegensis	XX	-	-	XX	3719956
Psychotria jinox: morazanensis C.W. Ham.	-	-	-	-	-
Psychotria lamarinensis C.W. Ham.	-	XX	XX	XX	5342778
Psychotria laselvensis C.W. Ham.	XX	-	-	XX	2998511
Psychotria liesneri Dwyer	-	XX	XX	-	1842296
Psychotria limonensis K. Krause	AF072052	-	-	-	AF072052
Psychotria lorenciana C.M. Taylor	XX	-	-	XX	5167031
Psychotria lundellii Standl.	-	-	-	XX	3029509
Psychotria marginata Sw.	XX	-	-	XX	2998499
	EF667972	-	-	-	EF667972
Psychotria matagalpensis C.M. Taylor	-	-	-	-	-
Psychotria mexiae Standl.	XX	-	-	XX	4282350
	XX	-	-	XX	2945604

Table C1.1. (continued)

Species	*ITS*(full)	*ITS1*	*ITS2*	*psbA*	Voucher Specimen[a]
Psychotria micrantha HBK	AF072048	-	-	-	AF072048
Psychotria mirandae C.W. Ham.	-	-	XX	-	05072106
Psychotria molinae Standl.	XX	-	-	XX	04591327
Psychotria monsalveae C.M. Taylor	-	-	-	-	-
Psychotria monteverdensis Dwyer & C.W. Ham.	-	-	-	-	-
Psychotria neilli C.W. Ham.	-	XX	XX	-	5167060
	-	-	-	XX	5315537
Psychotria nervosa Sw.	AF072046	-	-	-	AF072046
	XX	-	-	XX	2945599
Psychotria nubiphila Dwyer	XX	-	-	XX	2999147
	XX	-	-	XX	5307000
Psychotria olgae Dwyer & M.V. Hayden	-	XX	XX	-	2601725
Psychotria orosiana Standl.	XX	-	-	XX	2999142
.	XX	-	-	-	2999144
Psychotria orosioides C.M. Taylor	XX	-	-	XX	3007106

120

Table C1.1. (continued)

Species	ITS(full)	ITS1	ITS2	psbA	Voucher Specimen[a]
Psychotria pacorensis C.W. Ham.	-	XX	-	-	2892016
Psychotria panx: compressicaulis (K. Krause) C.W. Ham.	XX	-	-	-	2998531
	XX	-	-	XX	2998528
Psychotria panx: magna (Standl.) C.W. Ham.	-	XX	XX	-	04641754
Psychotria panx: panamensis	XX	-	-	-	2998535
Psychotria papatlensis (Oerst.) Hemsl.	-	XX	XX	-	3030658
Psychotria parvifolia Benth.	XX	-	-	XX	3108553
	XX	-	-	XX	2998508
	XX	-	-	-	2998514
Psychotria philacra Dwyer	-	XX	XX	XX	4324653
Psychotria pisonioides Standl.	-	-	-	-	-
Psychotria pleuropoda Donn. Sm.	-	-	-	-	-
Psychotria psychotriifolia (Seem.) Standl.	XX	-	-	XX	2998501
	XX	-	-	XX	2945595
Psychotria quinqueradiata Pol.	XX	-	-	XX	3007017

Table C1.1. (continued)

Species	*ITS*(full)	*ITS1*	*ITS2*	*psbA*	Voucher Specimen[a]
Psychotria quinqueradiata Pol.	XX	-	-	-	4306007
Psychotria remota Benth.	AF149403	-	-	-	AF149403
Psychotria rosulatifolia Dwyer	-	XX	XX	XX	3607831
Psychotria rufiramea Standl.	-	-	-	-	-
Psychotria saltatrix C.M. Taylor	-	XX	XX	XX	3752882
Psychotria sarapiquiensis Standl.	XX	-	-	XX	5727845
	-	-	-	XX	2998519
	-	-	-	XX	2998521
Psychotria sixaolensis C.W. Ham.	-	XX	XX	XX	4297832
Psychotria sylvivaga Standl.	XX	-	-	XX	2999146
Psychotria tenuifolia Sw.	XX	-	-	XX	2945600
	AF072050	-	-	-	AF072050
Psychotria trichotoma M. Martens & Galeotti	-	XX	XX	XX	3610927
Psychotria turrubarensis W. Burger & Q. Jimenez	-	-	XX	XX	04963197
Psychotria viridis Ruiz & Pav.	-	XX	XX	-	04568832

122

APPENDIX D

NAMES AND AUTHORITIES OF SPECIES IN CHAPER 4

D.1   PSYCHOTRIA SUBGENUS PSYCHOTRIA SPECIES

Psychotria chagrensis Standl.

Psychotria graciliflora Benth.

Psychotria jimenezii Standl.

Psychotria laselvensis C.W. Ham.

Psychotria marginata Sw.

Psychotria orosiana Standl.

Psychotria orosioides C.M. Taylor

Psychotria panamensis compressicaulis (K. Krause) C.W. Ham.

Psychotria panamensis panamensis Standl.

Psychotria parvifolia Benth.

Psychotria quinqueradiata Pol.

Psychotria sarapiquiensis Standl.

D.2   HETEROPSYCHOTRIA/PALICOUREA SPECIES

Palicourea albocaerulea C.M. Taylor

Palicourea gomezii C.M. Taylor

Palicourea guianensis Aubl.

Palicourea lasiorrhachis Oerst.

Palicourea macrocalyx Standl.

Palicourea montivaga Standl.

Palicourea padifolia (Willd. ex Roem. & Schult.) C.M. Taylor & Lorence

Psychotria acuminata Benth.

Psychotria aubletiana Steyerm.

Psychotria brachiata Sw.

Psychotria buchtienii (H.J.P. Winkl.) Standl.

Psychotria calidicola C.M. Taylor

Psychotria chiriquiensis (Standl.) C.M. Taylor

Psychotria cyanococca Seem. ex Dombrain

Psychotria elata (Sw.) Hammel

Psychotria eurycarpa Standl.

Psychotria gracilenta Müll. Arg.

Psychotria guapilensis (Standl.) Hammel

Psychotria hoffmannseggiana (Willd. ex Roem. & Schult.) Müll. Arg.

Psychotria microbotrys Ruiz ex Standl.

Psychotria pilosa Ruiz & Pav.

Psychotria poeppigiana Müll. Arg.

Psychotria pubescens Sw.

Psychotria racemosa (Aubl.) Raeusch.

Psychotria suerrensis Donn. Sm.

Psychotria valeriana Standl.

APPENDIX E

DETAILED METHODS FOR CHAPTER 4

E.1 DETAILS OF LABORATORY TECHNIQUES AND PROTOCOLS

DNA was extracted from fresh material using Qiagen® DNeasy Plant Mini Kits. Herbarium

samples were extracted following a standard CTAB protocol, except extractions were left for two

or more weeks in isopropanol to allow the maximum amount of DNA to precipitate out of

solution. Many of the herbarium extractions were cleaned prior to amplification using Qiagen®

MiniElute columns. *ITS* was amplified using PCR primarily in the CMNH biosystematics lab

and secondarily in the molecular lab of S. Kalisz at the University of Pittsburgh. I used the same

primers as Nepokroeff et al. (1999; LEU, ITS4, ITS3B), as well as 5.8s (for reverse strand

amplification of ITS1), and a new primer I developed, very similar to ITS3B, ITS3C (5'-

3':GATATCTAGGCTCTCGCATC; for forward strand amplification of ITS2). Standard 50 µl

reactions consisted of 35 µl sterile H$_2$O, 5 µl 10x buffer, 5 µl Mg, 1 µl BSA (10mg/ml), 1 µl

DMSO, 1 µl DNTPs (10mM), 0.5 µl 5' 20 µM primer, 0.5 µl 3' 20 µM primer, 1 µl genomic

DNA. Standard *ITS* PCR amplification started with 94.0°C for 2 m, then 40 cycles of 94.0°C for

30 s, 48.0°C for 1 m, 72°C for 1 m, and a final elongation step at 72°C for 7 m. PCR reaction

and amplification protocols were slightly modified for difficult to amplify taxa. All DNA sequencing was performed at the Davis Sequencing facility (www.davissequencing.com). Both strands were sequenced using the same primers used for amplification. Sequence strands were assembled using Sequencher 4.5 (Gene Codes 2005). I aligned sequences using ClustalX (Thompson et al. 1997) and made manual adjustments using Se-Al (Rambaut 1996).

APPENDIX F

THE ELEVATIONAL RANGES OF *PSYCHOTRIA* AND *PALICOUREA* SPECIES

FOUND IN COSTA RICA

Figure F1.1 The elevational ranges of *Psychotria* and *Palicourea* species found in Costa Rica.

These ranges were used to establish the regional species pool for each assemblage.

BIBLIOGRAPHY

Abrams, P.A. 2001. A world without competition. Nature 412:858-859.

Adobe Photoshop. 1989-2007. Adobe Systems Incorporated.

Álvarez, I., and J.F. Wendel. 2003 Ribosomal ITS sequences and plant phylogenetic inference. Molecular Phylogenetics and Evolution 29:417-434.

Anderson, M.T., M.A. Lachance, and W.T. Starmer. 2004. The relationship of phylogeny to community structure: the cactus yeast community. American Naturalist 164:709-721.

Andersson, L. 2002. Relationships and generic circumscriptions in the *Psychotria* complex (Rubiaceae, Psychotrieae). Systematic Geography of Plants 72:167-202.

Angiosperm Phylogeny Group. 2003. An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG II. Botanical Journal of the Linnean Society 141: 399-436.

Arbogast, B.S., Edwards, S.V., Wakeley, J., Beerli, P., and Slowinski J.B. 2002. Estimating divergence times from molecular data on phylogenetic and population genetic timescales. Annual Review in Ecology and Systematics 33:707-740.

Baldwin, B.G., M.J. Sanderson, J.M. Porter, M.F. Wojciechowski, C.S. Campbell, and M.J. Donoghue. 1995. The ITS region of nuclear ribosomal DNA: a valuable source of evidence on angiosperm phylogeny. Annals of the Missouri Botanical Garden 82: 247-277.

Bermingham, E., and Dick, C. 2001. The *Inga* – newcomer or museum antiquity? Science 293:2214-2216.

Blomberg, S.P., and T. Garland. 2002. Tempo and mode in evolution: phylogenetic inertia, adaptation and comparative methods. Journal of Evolutionary Biology 15:899-910.

Blomberg, S.P., T. Garland, and A.R. Ives. 2003. Testing the phylogentic signal in comparative data: behavioral traits are more labile. Evolution 57:717-745.

Böhning-Gaese, K. Caprano, T., van Ewijk, K., and Veith, M. 2006. Range size: disentangling current traits and phylogenetic and biogeographic factors. American Naturalist 167:555-567.

Bremer, B. and J.-F. Manen. 2000. Phylogeny and classification of the subfamily Rubioideae. Plant Systematics and Evolution. 225:43-72.

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

Brown JH (1995) *Macroecology*. University of Chicago Press, Chicago.

Brown, J.H., G.C. Stevens, and D.M. Kaufman. 1996. The geographic range: size, shape, boundaries, and internal structure. Annual Review of Ecology and Systematics 27:597-623.

Burger, W., and C.M. Taylor. 1993. Flora Costaricensis. Family #202 Rubiaceae. Fieldiana. Field Museum of Natural History, Chicago.

Cavender-Bares, J., D.D. Ackerly, D.A. Baum, and F.A. Bazzaz. 2004. Phylogenetic overdispersion in Floridian oak communities. American Naturalist 163:823-843.

Cavender-Bares, J., A. Keen, and B. Miles. 2006. Phylogenetic structure of Floridian plant communities depends on taxonomic and spatial scale. Ecology 87:S109-122.

Chesson, P. L. 2000. Mechanisms of maintenance of species diversity. Ann. Rev. Ecol. Syst. 31:343-366.

Chown, S.L. 1997. Speciation and rarity: separating cause from consequence. Pp. 91-109 in: *The Biology of Rarity, Causes and consequences of rare and common differences*, Kunin W.E., and Gaston, K.J. (eds.) Chapman and Hall, London.

Chown, S.L., and Gaston, K.J. 2000. Areas, cradles and museums: the latitudinal gradient in species diversity. Trends in Ecology and Evolution 15:311-315.

Clark, J.A, Fastie, C., Hurtt, G., Jackson, S.T., Johnson, C., King, G.A., Lewis, M., Lynch, M., Pacala, S., Prentice, C., Schupp, E.W., Webb, T., and Wyckoff. 1998. Reid's Paradox of Rapid Plant Migration. BioScience 48: 13-24.

Connell, J. H. 1971. On the role of natural enemies in preventing competitive exclusion in some marine animals and in rain forest trees. Pages 298-312 *in* P. J. den Boer and G. R. Gradwell, editors. Dynamics of Populations. Centre for Agriculture Publishing and Documentation, Wageningen.

Darwin, C. 1859. On the origin of species. Harvard University Press, Cambridge, MA.

Davis, C.C., Webb, C.O., Wurdack, K.J., Jaramillo, C.A., and Donoghue, M.J. 2005. Explosive radiation of Malpighiales supports a mid-Cretaceous origin of modern tropical rain forests. American Naturalist 165:E36-E65.

Dick, C.W., Abdul-Salim, K., and Bermingham E. 2003. Molecular systematic analysis reveals cryptic tertiary diversification of a widespread tropical rain forest tree. American Naturalist 162:691-703.

Dobzhansky, T. 1950. Evolution in the tropics. American Scientist 38:209-221.

Drummond, A.J. and A. Rambaut. 2003. BEAST v1.4, http://beast.bio.ed.ac.uk/.

Drummond, A.J., S.Y.W.,Ho, M.J. Phillips, and A. Ramaut. 2006. Relaxed phylogenetics and dating with confidence. PLoS Biology 4:e88:0001-0012.

Dyer, L.A., and Palmer, A.D.N. 2004 (eds.). *Piper. A model genus for studies of phytochemistry, ecology, and evolution*. Kluwer Academic, New York.

Ehrendorfer, F. 1982. Speciation patterns in woody angiosperms of tropical origin. Pp 479-509 in: *Mechanisms of Speciation*, Barrigozzi, (ed.), Liss, New York.

Elith, J. and 26 other authors. 2006. Novel methods improve prediction of species' distributions from occurrence data. Ecography 29:129-151.

Elton, C. 1946. Competition and the structure of ecological communities. Journal of Animal Ecology 15:54-68.

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

Feliner, G.N. and J.A. Rosselló. 2007. Better the devil you know? Guidelines for insightful utilization of nrDNA ITS in species-level evolutionary studies in plants. Molecular Phylogenetics and Evolution 44:911-919.

Fernald, M.L. 1924. Isolation and endemism in Northeastern America and their relation to the age-and-area hypothesis. American Journal of Botany 11: 558-572.

Foote, M., Crampton, J.S, Beu, A.G., Marshall, B.A., Cooper, R.A., Maxwell, P.A. and Matcham, I. 2007. Rise and fall of species occupancy in Cenozoic fossil mollusks. Science 318:1131-1134.

Foote, M., J.S. Crampton, A.G. Beu, B.A. Marshall, R.A. Cooper, P.A. Maxwell, and I. Matcham. 2007. Rise and fall of species occupancy in Cenozoic fossil mollusks. Science 318:1131-1134.

Forrest, L.L., and Hollingsworth, P.M. 2003. A recircumscription of *Begonia* based on nuclear ribosomal sequences. Plant Systematics and Evolution 241:193-211.

Fuentes, M. 2004. Slight differences among individuals and the unified neutral theory of biodiversity. Theoretical Population Biology 66:199-203.

Gause, G. F. 1934. The struggle for existence. Williams and Williams, Baltimore.

Gaston, K.J. 1994. *Rarity*. Chapman and Hall, London.

Gaston, K.J., 1996. The multiple forms of the interspecific-abundance distribution relationship. Oikos 75:211-220.

Gaston, K.J. 1998. Species-range size distributions: products of speciation, extinction and transformation. Philosophical Transactions of the Royal Society of London B 353:219-230.

Gaston, K.J. 2003. *The structure and dynamics of geographic ranges*. Oxford University Press, Oxford.

Gaston, K.J., and T.M. Blackburn. 1997. Age, area, and avian diversification. Biological Journal of the Linnean Society. 62:239-253.

Gene Codes. 1991-2007. Sequencher 4.5. Gene Codes Corporation, Ann Arbor, MI.

Gentry AH (1982) Patterns of neotropical plant-species diversity. Evolutionary Biology 15: 1-85

Gentry, A.H. 1990. Floristic similarities and differences between southern Central America and central Amazonia. Pp. 141-157 in: *Four Neotropical Rainforests*, Gentry, A.H. (ed.), Yale University Press, New Haven.

Gleason, H.A. 1924. Age and area from the viewpoint of phytogeography. American Journal of Botany 11: 541-546.

Graham, A. 1985. Studies of Neotropical paleobotany. II. The Eocene communities of Panama. Annals of the Missouri Botanical Garden 72:504-534.

Grinnell, J. 1917. Field tests of theories concerning distributional control. American Naturalist 51:115-128.

Grinnell, J. 1924. Geography and evolution. Ecology 5:225 – 229

Hamilton, C.W. 1989a. A revision of Mesoamerican *Psychotria* subgenus *Psychotria* (Rubiaceae), Part I: Introduction and species 1 – 16. Annals of the Missouri Botanical Garden 76: 67-111.

Hamilton, C.W. 1989b. A revision of Mesoamerican *Psychotria* subgenus *Psychotria* (Rubiaceae), Part II: Species 17 – 47. Annals of the Missouri Botanical Garden 76: 386-429.

Hamilton, C.W. 1989c. A revision of Mesoamerican *Psychotria* subgenus *Psychotria* (Rubiaceae), Part III: Species 48 – 61 and appendices. Annals of the Missouri Botanical Garden 76: 886-916.

Hanski, I., J. Kouki, A. Halkka. 1993. Three explanations of the positive relationship between distribution and abundance of species. Pp. 108-116 in: *Species diversity in ecological communities: historical and geographic perspectives*, R.E. Ricklefs and D. Schluter (eds). University of Chicago Press, Chicago.

Harms, K.E., Wright, S.J., Calderón, O., Hernández, A., and Herre, E.A. 2000. Pervasive density-dependent recruitment enhances seedling diversity in a tropical forest. Nature 404:493-495.

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

Hernandez, P.A., C.H. Graham, L.L. Master, and D.L. Albert. 2006. The effect of sample size and species characteristics on performance of different species distribution modeling methods. Ecography 29:773-785.

Hijmans, R.J., L. Guarino, M. Cruz, and E. Rojas. 2001. Computer tools for spatial analysis of plant genetic data: 1. DIVA-GIS. Plant genetic Resources Newsletter 127:15-19.

Hillis DM (1998) Taxonomic sampling, phylogenetic accuracy, and investigator bias. Syst. Biol. 47:3-8.

Holbrook, K.M., and T.B. Smith. 2000. Seed dispersal and movement patterns in two species of *Ceratogymna* hornbills in a West African tropical lowland forest. Oecologia 125: 249-257.

Howe HF, J. Smallwood. 1982. Ecology of seed dispersal. Annual Review of Ecology and Systematics 13:201-218.

Hubbell, S.P. 1979. Tree Dispersion, Abundance, and Diversity in a Tropical Dry Forest. Science 203:1299-1309.

Hubbell, S.P. 2001a. *The unified neutral theory of biodiversity and biogeography*. Princeton University Press, Princeton.

Hubbell, S.P. 2001b. The unified neutral theory of biodiversity and biogeography: a synopsis of the theory and some challenges ahead. Pp. 393-411 in: *Integrating ecology and evolution in a spatial context*, Silvertown, J. and Antonovics, J. (eds.), Blackwell Science, Oxford.

Hubbell, S.P. 2003. Modes of speciation and the lifespans of species under neutrality: a response to the comment of Robert E. Ricklefs. Oikos 100:193-199.

Hubbell, S. P. and R. B. Foster. 1986. Biology, chance and history and the structure of tropical rain forest tree communities. Chapter l9, pp. 314-329, in Diamond, J. and Case, T.J. eds. *Community Ecology*, Harper and Row, NY.

Hubbell, S.P., and J.K. Lake. 2003. The neutral theory of biodiversity and biogeography, and beyond. Pp 45-63 in: *Macroecology: concepts and consequences*, Blackburn, T.M., and K.J. Gaston, (eds.), Blackwell, London.

Hunt, T., J. Bergsten, Z. Levkanicova, A. Papadopoulou, O. St. John, R. Wild, P.M. Hammond, D. Ahrens, M. Balke, M.S. Caterino, J. Gómez-Zurita, I. Ribera, T. G. Barraclough, M. Bocakova, L. Bocak, A. P. Vogler. 2007. A comprehensive phylogeny of beetles reveals the evolutionary origins of a superradiation. Science 318:1913-1916.

Hutchinson, G.E. 1957. Concluding Remarks. Cold Spring Harbor Symp. Quantitative Biology 22:415-427.

Hutchinson, G.E. 1959. Homage to Santa Rosalia or why are there so many kinds of animals? American Naturalist 93:153-159.

Jablonski, D. 1987. Heritability at the species level: analysis of geographic ranges of Cretaceous molluscs. Science 238: 360-363.

Janzen, D. H. 1970. Herbivores and the number of tree species in tropical forests. American Naturalist. 104:501-528.

Jaramillo, M.A., and Callejas, R. 2004a. A reappraisal of *Trianaeopiper* Trelease: convergence of dwarf habitat in some *Piper* species of the Chocó. Taxon 53:269-278.

Jaramillo, M.A., and Callejas, R. 2004b. Current perspectives on the classification and phylogenetics of the genus *Piper* L., Pp. 179-198 in: *Piper. A model genus for studies of phytochemistry, ecology, and evolution*, Dyer, L.A., and Palmer, A.D.N. 2004 (eds.), Kluwer Academic, New York.

Jaramillo, M.A., and Manos, P.S. 2001. Phylogeny and patterns of floral diversity in the genus *Piper* (Piperaceae). American Journal of Botany 88:706-716.

Johnson, C.N. 1998. Species extinction and relationship between distribution and abundance. Nature 394:272-274.

Johnson, R.H. 1910. Determinate evolution in the color pattern of lady beetles. Carnegie Institute of Washington, Publication 122.

Jones, K.E., Sechrest, W., and Gittleman, J.L. 2005. Age and area revisited: identifying global patterns and implications for conservation, Pp.142-165 in: *Phylogeny and Conservation*, Purvis, A., Gittleman, J.L., and Brooks, T. (eds.) Cambridge University Press, Cambridge.

Jones, K.E., Sechrest, W., and Gittleman, J.L. 2005. Age and area revisited: identifying global patterns and implications for conservation, Pp.142-165 in: *Phylogeny and Conservation*, Purvis, A., Gittleman, J.L., and Brooks, T. (eds.) Cambridge University Press, Cambridge.

Jordano, P. 1995. Angiosperm fleshy fruits and seed dispersers –a comparative analysis of adaptation and constraints in plant animal interactions. American Naturalist 145:163–191.

Kay, K.M., Reeves, P.A., Olmstead, R.G., and Schemske, D.W. 2005. Rapid speciation and the evolution of hummingbird pollination in Neotropical *Costus* subgenus *Costus* (Costaceae): evidence from nrDNA ITS and ETS sequences. American Journal of Botany 92:1899-1910.

Kembel, S.W., and Hubbell, S.P. 2006. The phylogenetic structure of a Neotropical forest community. Ecology 87:S86-S99.

Kimura, M. 1968. Evolutionary rate at the molecular level. Nature 217:624-626.

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

Kress, W.J., K.J. Urdack, E.A. Zimmer, L.A. Weigt, and D.H. Janzen. 2005. Use of DNA barcodes to identify flowering plants. Proceedings of the National Academy of Science 102:8369-8374.

Lavin, M., Herendeen, P.S., and Wojciechowski, M.F. 2005. Evolutionary rates analysis of Leguminosae implicates a rapid diversification of lineages during the Tertiary. Systematic Biology 54:575-594.

Lavin, M., Schrire, B.P., Lewis, Gwilym, Pennington, R.T., Delgado-Salinas, A., Thulin, M., Hughes, C.E., Matos, A.B., and M. Wojciechowski, M.F. 2004. Metacommunity process rather than continental tectonic history better explains geographically structured phylogenies in legumes. Philosophical Transactions of the Royal Society of London Series 359:1509-1522.

Leigh, E.G. Jr. 1999. *Tropical forest ecology a view from Barro Colorado Island*. Oxford University Press, Oxford.

Leigh, E.G. Jr., Davidar, P., Dick, C.W., Puyravaud, J.-P., Terborgh, J., ter Steege, H., and Wright, S.J. 2004. Why do some tropical forests have so many species of trees? Biotropica 36:447-473.

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

Linder, H.P., Hardy, C.R., and Rutschmann, F. 2005. Taxon sampling effects in molecular clock dating: an example from the African Restionaceae. Molecular Phylogenetics and Evolution 35:569-582.

Loiselle, B.A., V.L Sork, and C. Graham. 1995. Comparison of genetic variation in bird-dispersed shrubs of a tropical wet forest. Biotropica 27: 487-494.

Losos, J.B., M. Leal, R.E. Glor, K. de Queiroz, P.E. Hertz, L. Rodriguez Schettino, A. Chamizo Lara, T.R. Jackman, and A. Larson. 2003. Niche lability in the evolution of a Caribbean lizard community. Nature 424:542-545.

Lovette, I.J., and W.M. Hochachka. 2006. Simultanous effects of phylogenetic niche conservatism and competition on avian community structure. Ecology 87:S14-S28.

MacArthur, R. H. 1958. Population ecology of some warblers of northeastern coniferous forests. Ecology **39**:599-619.

MacArthur, R.H. 1960. On the relative abundance of species. American Naturalist 94:25-36.

MacArthur, R. H., and R. Levins. 1964. Competition, habitat selection, and character displacement in a patchy environment. Proceeding of the National Academy of Sciences USA 51:1207-1210.

MacArthur, R. H., and R. Levins. 1967. The limiting similarity, convergence and divergence of coexisting species. American Naturalist 101:377-385.

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

Magallón, S. P. R. Crane and P. S. Herendenn. 1999. Phylogenetic pattern, diversity, and diversification of eudicots. Annals of the Missouri Botanical Garden 86:297-372.

Malcomber, S.T. 2002. Phylogeny of *Gaertnera* Lam. (Rubiaceae) based on multiple DNA markers: evidence of a rapid radiation in a widespread, morphologically diverse genus. Evolution 56:42-57.

Marquis, R.J. 2004. Biogeography of Neotropical *Piper*, Pp. 78-96 in: *Piper. A model genus for studies of phytochemistry, ecology, and evolution*, Dyer, L.A., and Palmer, A.D.N. 2004 (eds.), Kluwer Academic, New York.

McGill, B.J., B.A. Maurer, and M.D. Weiser. 2006. Empirical evaluation of neutral theory. Ecology 87:1411-1483.

McPeek, M.A. 2007. The macroevolutionary consequences of ecological differences among species. Palaeontology 50:111-129.

McPeek, M.A., and J.M Brown. 2007. Clade age and not diversification rate explains species richness among animal taxa. American Naturalist 169:E97-E106.

Miller, A.I. 1997. A new look at age and area: the geographic and environmental expansion of genera during the Ordovician radiation. Paleobiology 23: 410-419.

Morely, R.J. 2000. *Origin and evolution of tropical rainforests*. Chichester, UK, Wiley Press.

Mort, M.E. J.K. Archibald, C.P. Randle, N.D. Levsen, T. R. O'Leary, K. Topalov, C.M. Wiegand and D.J. Crawford. 2007. Inferring phylogeny at low taxonomic levels: utility of rapidly evolving cpDNA and nuclear ITS loci. American Journal of Botany 94:173-183.

Muellner, A.N., Savolainen, V., Samuel, R., and Chase, M.W. 2006. The mahogany family "out-of-Africa": divergence time estimation, global biogeographic patterns inferred from plastid *rbc*L DNA sequences, extant, and fossil distribution of diversity. Molecular Phylogenetics and Evolution 40:236-250.

Murray, B.R., and G.C. Hose. 2005. Life-history and ecological correlates of decline and extinction in the endemic Australian frog fauna. Austral Ecology 30:564-571.

Nee, S. 2005. The neutral theory of biodiversity: do the numbers add up? Functional Ecology 19:173-176.

Nepokroeff, M., B. Bremer, B., and K.J. Sytsma. 1999. Reorganization of the genus *Psychotria* and Tribe Psychotrieae (Rubiaceae) inferred from ITS and *rbcL* sequence data. Systematic Botany 24:5-27.

Nepokroeff, M., K.J. Sytsma, W.L. Wagner, E.A. Zimmer. 2003. Reconstructing ancestral patterns of colonization and dispersal in the Hawaiian understory tree genus *Psychotria* (Rubiaceae): a comparison of parsimony and likelihood approaches. Systematic Biology 52:820-838.

Nichols, R. 2001. Gene trees and species trees are not the same. Trends in Ecology and Evolution 16:358-364.

Paine, R. T. 1966. Food web complexity and species diversity. American Naturalist 100:65-76.

Paul, J.R. and S.J. Tonsor. 2008. Explaining geographic range size by species age: a test using Neotropical *Piper* species. In Press: *Tropical Forest Community Ecology*, Carson, W.P. and S. Schnitzer, eds. Blackwell Publishing.

Payne, J.L., and S. Finnegan. 2007. The effect of geographic range on extinction risk during background and mass extinction. Proceedings of the National Academy of Sciences (USA): 104:10506-10511.

Peterson, A.T., J. Soberón, and V. Sanchez-Cordero. 1999. Conservatism of ecological niches in evolutionary time. Science 285:1265-1267.

Phillips, S.J., M. Dudik, and R.E. Schapire. 2004. A maximum entropy approach to species distribution modeling. Proceedings of the Seventeenth Annual Conference on Machine Learning 555-662.

Phillips, S.J., R.P. Anderson, and R.E. Schapire. 2006. Maximum entropy modeling of species geographic distributions. Ecological Modelling 190:231-259.

Pirie, M.D., Chatrou, L.W., Mols, J.B., Erkens, R.H.J., and Oosterhof, J. 2006. 'Andean-centered' genera in the short-branch clade of Annonaceae: testing biogeographical hypotheses using phylogeny reconstruction and molecular dating. Journal of Biogeography 33:31-46.

Pitman, N.C.A., Terborgh, J.W., Silman, M.R., Nunez, P., Neill, D.A., Ceron, C.E., Palacios, W.A., and Aulestia, M. 2001. Dominance and distribution of tree species in upper Amazonian terra firma forests. Ecology 82:2101-2117.

Plana, V., Gascoigne, A., Forrest, L.L., Harris, D., and Pennington, R.T. 2004. Pleistocene and pre-Pleistocene *Begonia* speciation in Africa. Molecular Phylogenetics and Evolution 31:449-461.

Posada, D. and K.A. Crandall. 1998. Modeltest: testing the model of DNA substitution. Bioinformatics 14:917-818.

Quijano-Abril, M.A., Callejas-Posada, R., and Miranda-Esquivel, D.R. 2006. Areas of endemism and distribution patterns for Neotropical *Piper* species (Piperaceae). Journal of Biogeography 33:1266-1278.

Rambaut A., and A.J. Drummond. 2004. Tracer v1.3, Available from http://tree.bio.ed.ac.uk/software/tracer/.

Rambaut A., and A.J. Drummond. 2006. TreeAnnotator v1.4.5r516 Available from: http://tree.bio.ed.ac.uk/software/treeannotator/.

Rambaut, A. 1996. Se-Al: Sequence Alignment Editor. Available at http://tree.bio.ed.ac.uk/software/seal/.

Rasband, W.S., 1997-2007. ImageJ, U. S. National Institutes of Health, Bethesda, Maryland, USA, Avaialble at http://rsb.info.nih.gov/ij/.

Razafimandimbison, S.G., E.A. Kellogg, and B. Bremer. 2004. Recent origin and phylogentic utility of divergent ITS putative pseudogenes: A case study from Naucleeae (Rubiaceae). Systematic Biology 53:177-192.

Renner, S.S. 2005. Relaxed molecular clocks for dating historical plant dispersal events. Trends in Plant Science 10:550-558.

Renner, S.S., Clausing, G., and Meyer, K. 2001. Historical biogeography of Melastomataceae: the roles of tertiary migration and long-distance dispersal. American Journal of Botany 88:1290-1300.

Richardson, J.E., Chatrou, L.W., Mols, J.B., Erkens, R.H.J., and Pirie, M.D. 2004. Historical biogeography of two cosmopolitan families of flowering plants: Annonaceae and

Rhamnaceae. Philosophical Transactions of the Royal Society of London Series B, Biological Sciences 359:1495-1508.

Richardson, J.E., Pennington, R.T., Pennington, T.D., and Hollingsworth, P.M. 2001. Rapid diversification of a species-rich genus of Neotropical rain forest trees. Science 293:2242-2245.

Ricklefs, R.E. 2003. A comment on Hubbell's zero-sum ecological drift model. Oikos 100:185-192.

Ricklefs, R.E. 2004. A comprehensive framework for global patterns in biodiversity. Ecology Letters 7:1-15.

Ronquist, F. and Huelsenbeck, J.P. 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19:1572-1574.

Roy, K., and E. E. Goldberg. 2007. Origination, extinction, and dispersal: integrative models for understanding present-day diversity gradients. American Naturalist 170(suppl.):S71–S85.

Sanderson, M.J., Thorne, J.L., Wikström, N., and Bremer, K. 2004. Molecular evidence on plant divergence times. American Journal of Botany 91:1656-1665.

SAS Institute Inc. 2001. SAS Software, Release 8.2, Cary, NC: SAS Institute Inc.

SAS Institute Inc. 2005. SAS Software, Release 9.1.3, Cary, NC: SAS Institute Inc.

Schurr, F.M., G.F. Midgley, A.G. Rebelo, G. Reeves, P. Poschlod, and S.I. Higgins. 2007. Colonization and persistence ability explain the extent to which plant species fill their potential range. Global Ecology and Biogeography 16:449-459.

Silvertown, J., K. McConway, D. Gowing, M. Dodd, M. Fay, J.A. Joseph, and K. Dolphin. 2006. Absence of phylogenetic signal in the niche structure of meadow plant communities. Proceedings of the Royal Society of London B 273:39-44.

Slingsby, J.A., and G.A.Verboom. 2006. Phylogenetic relatedness limits co-occurrence at fine spatial scale: evidence from the Schoenoid Sedges (Cyperaceae: Schoeneae) of the cape floristic region, South Africa. American Naturalist 168:14-27.

Smith, S. A., A. Nieto Montes de Oca, T. W. Reeder, and J. J. Wiens. 2007. A phylogenetic perspective on elevational species richness patterns in Middle American treefrogs: why so few species in lowland tropical rainforests? Evolution 61:1188–1207.

Stebbins, G.L., and Major, J. 1965. Endemism and speciation in the California flora. Ecological Monographs 35:1-35.

Stephens, P.R. and Wiens J.J. 2003. Explaining species richness from continents to communities: the time-for-speciation effect in Emydid turtles. American Naturalist 161:112-128.

Stone, J.L. 1995. Pollen donation patterns in a tropical distylous shrub (*Psychotria suerrensis*; Rubiaceae). American Journal of Botany 82:1390-1398.

Svenning, J.-C., and F. Skov. 2004. Limited filling of the potential range in European tree species. Ecology Letters 7:565-573.

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

Swofford, D.L. 2002. PAUP*: Phylogenetic analysis using parsimony (*and other methods). Version 4.0b10. Sinauer Associates, Sunderland, MA.

Taylor, C.M. 1989. Revision of *Palicourea* (Rubiaceae) in Mexico and Central America. Systematic Botany Monographs Vol. 26. The American Society of Plant Taxonomists.

Taylor, C.M. 1996. Overview of the Psychotrieae (Rubiaceae) in the Neotropics. Opera Botanica Belgica 7:261-270.

Taylor, C.M. 2001. Overview of the Neotropical genus *Notopleura* (Rubiaceae: Psychotrieae), with the description of some new species. Annals of the Missouri Botanical Garden 88: 478-515.

Taylor, C.M. 2007. Flora Mesoamericana – Rubiaceae. Available at http://www.mobot.org/mobot/fm/.

Thomas, R.H., and J.A. Hunt. 1991. The molecular evolution of the alcohol dehydrogenase locus and the phylogeny of Hawaiian *Drosophila*. Molecular Biology and Evolution 8:687-702.

Thompson, J. D., D.G. Higgins and T.J. Gibson. 1997. ClustalW: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Research 22:4673-4680.

Thompson, J. D., Higgins, D. G. and Gibson, T. J. 1994. ClustalW: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Research 22:4673-4680.

Tilman, D. 1982. *Resource competition and community structure*. Princeton University Press, Princeton.

Tilman, D. 1988. *Plant strategies and the dynamics and structure of plant communities*. Princeton University Press, Princeton.

Trelease, W., and Yuncker, T.G. 1950. *The Piperaceae of Northern South America*. University of Illinois Press, Urbana.

Vamosi, J.C., and S.M. Vamosi. 2007. Body size, rarity, and phylogenetic community structure: insights from diving beetle assemblages of Alberta. Diversity and Distributions 13:1-10.

van der Veken, S., J. Rogister, K. Verheyen, M. Hermy, and R. Nathan. 2007. Over the (range) edge: a 45-year transplant experiment with the perennial forest herb *Hyacinthoides non-scripta*. Journal of Ecology 95:343-351.

Volkov, I., Banavar, J.R., He, F., Hubbell, S.P., and Maritan, A. 2005. Density dependence explains tree species abundance and diversity in tropical forests. Nature 438:658-661.

Vrba, E.S., and DeGusta, D. 2004. Do species populations really start small? New perspectives from the Late Neogene fossil record of African mammals. Philosophical Transactions of the Royal Society of London Series B 359:285-292.

Waldron, A. 2007. Null models of geographic range size evolution reaffirm its heritability. American Naturalist 170:221-231.

Wallace, A.R. 1878. *Tropical nature and other essays*. MacMillan, New York and London.

Webb C.O. 2000. Exploring the phylogenetic structure of ecological communities: An example for rain forest trees. American Naturalist 156: 145-155.

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

Webb. C.O, D.D. Ackerly, S.W. Kembel. 2007. Phylocom: software for the analysis of community phylogenetic structure and character evolution. Version 3.41. Available at http://www.phylodiversity.net/phylocom/.

Webb, T.J., and K.J. Gaston. 2000. Geographic range size and evolutionary age in birds. Proceedings of the Royal Society of London Series B 267: 1843-1850.

Weiblen, G.G., C.O. Webb, V. Novotny, Y. Basset, and S.E. Miller. 2006. Phylogenetic dispersion of host use in a tropical insect herbivore community. Ecology 87:S62-S75.

Weiher, E., G. Clark, and P. Keddy. 1998. Community assembly rules, morphological dispersion, and the coexistence of plant species. Oikos 81:309-322.

Wheelwright, N.T. 1985. Fruit-size, gape width, and the diets of fruit-eating birds. Ecology 66:808-818.

Whittaker, R.H., Levin, S.A., Root, R.B. 1975. Niche, habitat and ecotope. American Naturalist 107: 321 − 338.

Wiens, J.J. 2004. Speciation and ecology revisited: phylogenetic niche conservatism and the origin of species. Evolution 58:193-197.

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

Wiens, J.J., and Donoghue, M.J., 2004. Historical biogeography, ecology, and species richness. Trends in Ecology and Evolution 19:639-644.

Wiens, J.J., and C.H. Graham. 2005. Niche conservatism: integrating evolution, ecology, and conservation biology. Annual Review of Ecology, Evolution, and Systematics 36:519-539.

Wikström, N., V. Savolainen, and M.W. Chase. 2001. Evolution of the angiosperms: calibrating the family tree. Proceedings of the Royal Society of London, Series B 268:2211-2220.

Wilgenbusch J.C., D.L. Warren, D.L. Swofford. 2004. AWTY: A system for graphical exploration of MCMC convergence in Bayesian phylogenetic inference. Avaialbe at http://ceb.csit.fsu.edu/awty/.

Willis, J.C. 1922. *Age and area*. Cambridge University Press, Cambridge.

Wills, C., Harms, K.E., Condit, R., King, D., Thompson, J., He, F., Mulle-Landau, H.C., Ashton, P., Losos, E., Comita, L., Hubbell, S., LaFrankie, J., Bunyavejchewin, S., Dattaraja, H.S., Davies, S., Esufali, S., Fostre, R., Gunatilleke, N., Gunatilleke, S., Hall, P., Itoh, A., John, R., Kiratiprayoon, S., Loo de Lao, S., Massa, M., Nath, C., Noor, N.S., Kassin, A.R., Sukumar, R., Suresh, H.S., Sun, I.-F., Tan, S., Yamakura, T., and Zimmerman, J. 2006. Nonrandom processes maintain diversity in tropical forests. Science 311:527-531.

Wright, I. and 16 other authors. 2007. Relationships among ecologically important dimensions of plant trait variation in seven Neotropical forests. Annals of Botany 99:1003-1015.

Yang, R.C., F.C. Yeh, and A.D. Yanchuk. 1996. A comparison of isozyme quantitative genetic variation in *Pinus contorta* spp. latifolia by F_{ST}. Genetics 142:1045-1052.

Yu, D.W., Terborgh, J.W., and Potts, M.D. 1998. Can high tree species richness be explained by Hubbell's null model? Ecology Letters 1:193-199.

Yule, G.U. 1925. A mathematical theory of evolution, based on the conclusions of Dr. J.C. Willis, F.R.S. Philosophical Transactions of the Royal Society of London Series B, Biological Sciences 213:21-87.

Zerega, N.J.C., Clement, W.L., Datwyler, S.L., and Weiblen, G.D. 2005 Biogeography and divergence times in the mulberry family (Moraceae). Molecular Phylogenetics and Evolution 37:402-416.

Zwickl, D.J. 2006. Genetic algorithm approaches for the phylogenetic analysis of large biological sequence datasets under the maximum likelihood criterion. Ph.D. dissertation, The University of Texas at Austin. Available at http://www.bio.utexas.edu/faculty/antisense/garli/Garli.html.