Title
Ecological divergence and reproductive isolation in an Amazonian tropical tree: Protium subserratum (Burseraceae)

Author
Misiewicz, Tracy Marie

Publication Date
2014

Peer reviewed|Thesis/dissertation

Ecological Divergence and Reproductive Isolation in an Amazonian Tropical Tree: *Protium subserratum* (Burseraceae)

By

Tracy Marie Misiewicz

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

In

Biology

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Paul V.A. Fine, Chair
Professor John Huelsenbeck
Professor Gordon Frankie

Spring 2014

Abstract

Ecological divergence and reproductive isolation in an Amazonian tropical tree: *Protium subserratum* (Burseraceae)

By

Tracy Marie Misiewicz

Doctor of Philosophy in Biology

University of California, Berkeley

Professor Paul V.A. Fine, Chair

Neotropical forests have the highest tree diversity on earth, with an estimated 22,000 species. In contrast temperate North America, Europe and Asia combined support only 1,166 tree species. In the western Amazon rainforest, complex patterns of edaphic heterogeneity have been invoked as potential drivers of plant diversity through local. Numerous studies have demonstrated that physiological trade offs associated with adaptation to different habitat types leads to ecological sorting which in turn drives the spatial distribution of tropical trees along environmental gradients. If populations that are locally adapted to different environments experience a reduction in gene flow they may continue to differentiate potentially leading to new species formation.

Ecological speciation, whereby divergent natural selection results in reproductive isolation occurs as a direct result of adaptation to ecological conditions, is thought to be an important driver of plant diversification. Ecological speciation is known to occur in allopatry and a number of studies have demonstrated that it may be possible in parapatry and sympatry. However, the latter scenario, in which divergence and speciation occurs in the face of gene flow remains a point of contention. In the absence of geographic barriers to gene flow, disruptive selection must be strong enough to overcome genetic 'dilution' from neighboring populations. Environmental factors commonly associated with ecologically based disruptive selection include variation in topography, microenvironmental differences, soil heterogeneity, herbivore pressure, and pollinator differentiation, all of which can reduce gene flow, potentially leading to the evolution of reproductive barriers.

Reproductive barriers can act to isolate diverging lineages before or after fertilization. Prezygotic isolating mechanisms in plants can be intrinsic, usually entailing pollen incompatibility with the stigma or ovule, or extrinsic, where isolation is ecologically driven, as with pollinator-mediated barriers or changes in flowering phenology. Postzygotic isolating mechanisms can be intrinsic, in the case of hybrid sterility and inviability, or extrinsic, where isolation is enforced through ecologically based natural selection where hybrid offspring are less suited for survival in either parental habitat type. To date, only a handful of studies have comprehensively quantified the strength of individual reproductive isolating mechanisms among closely related plant species, the majority of which are herbaceous and found in temperate zones. Even fewer studies have focused on tropical plants and those that have focused on herbaceous lineages.

This dissertation aimed to investigate the broader role of divergent natural selection as an agent in generating and maintaining Amazonian tree diversity using edaphically divergent populations of *Protium subserratum*. Using a combination of molecular population genetics field observations and hand-pollination experiments I: (1) Developed nuclear microsatellite markers in order to evaluate population level differentiation and gene flow between populations of *P. subserratum* associated with different soil types. (2) Inferred the role of divergent natural selection in driving the genetic structure of *P. subserratum* populations found on clay, brown-sand, and white-sand soils distributed more than 100km across the Peruvian Amazon. (3) Investigated the role of habitat and distance in driving turnover in stingless bee communities, the putative effective pollinators for *P. subserratum.* (4) Systematically evaluated the role of multiple pre- and post-zygotic barriers to reproduction in maintaining population integrity between parapatric populations of white-sand and brown-sand *P. subserratum.*

I successfully developed 17 polymorphic nuclear microsattelite markers, which were subsequently used to assess population variation across populations of *P. subserratum*, found on different soil types. I found evidence that suggests that edaphic specialization has occurred multiple times in *P. subserratum* and that natural selection may be driving divergence across edaphic boundaries. I found that location and soil type play a significant role in structuring stingless bee communities in white-sand and non-white sand forest and that community turnover may be more strongly influenced by distance in white-sand habitats than non-white sand habitats. I was able to identify four active barriers to reproduction between parapatric white-sand and brown-sand populations of *P. subserratum* including ecogeographic isolation, differential pollen adhesion, differences in pollinator assemblages, and low levels of hybrid seed development. I demonstrated that a combination of pre-zygotic and post-zygotic barriers to reproduction act to maintain near complete reproductive isolation between edaphically divergent populations of the tropical tree, *P. subserratum*.

To my parents who taught me to be considerate, to work hard, and to think critically and for encouraging me to take chances and pursue my passions even when they've taken me far away from the nest.

To Tara, who has always been there for me with an ear, a shoulder or as a field assistant.

To Benjamin Carter, my (moss covered) rock, for reminding me not to overlook the small things in science and in life.

TABLE OF CONTENTS

Acknowledgements iii
List of Tables, Figures & Appendices iv

CHAPTER 1. Microsatellite primers for an Amazonian lowland tropical tree, *Protium subserratum* (Engl.) Engl. (Burseraceae)
Abstract 1
Introduction 1
Methods and Results 1
Discussion 2
Work Cited 5

CHAPTER 2. Evidence for ecological divergence across a mosaic of soil types in an Amazonian tropical tree: *Protium subserratum* (Burseraceae)
Abstract 6
Introduction 6
Methods 8
Results 13
Discussion 15
Work Cited 27

CHAPTER 3. Distance and Habitat Drive Stingless Bee Community Turnover Across Naturally Heterogeneous Forests in the Western Amazon
Abstract 33
Introduction 33
Materials and Methods 34
Results 36
Discussion 37
Work Cited 43

CHAPTER 4. The contribution of multiple barriers to reproduction between edaphically divergent lineages in the Amazonian tree *Protium subserratum* (Burseraceae)
Abstract 46
Introduction 46
Materials and Methods 47
Results 51
Discussion 53
Work Cited 61

ACKNOWLEDGEMENTS

I would like to thank my advisor, Paul Fine, for supporting my endeavors both financially and intellectually, for allowing me to work autonomously, and for always having an open office door. I also owe thanks to my Dissertation Committee members, John Huelsenbeck and Gordon Frankie, who provided invaluable feedback and advice through the years. Thanks to my Qualifying Exam Committee, Richard Dodd, Bruce Baldwin, Jim Maguire, for their stimulating conversations and for keeping me on my A-game.

This dissertation would not have been possible without the help of Laurie Hall and Abby Moore who walked me through the ins and outs of microsatellite genotyping, Julio Sanchez for his assistance in the field and his perpetual good humor and Italo Mesones for assistance with logistics and permitting in Peru. I would like to thank Steven Ruzin and Denise Schichnes at the Berkeley Biological Imaging Facility for advising me on fixation, storage, staining and fluorescent imaging of plant material. My undergraduate research assistants, Katie Pettibone and Monique Picon provided invaluable assistance in the field and the laboratory. A huge thanks to Bier Kraichak and Benjamin Carter, my R gurus, and Claus Rasmussen for identifying over 1,000 stingless bees to species.

I must thank my lab mates Seth Kauppinen, Chris DiVittorio, Juan Guevara, Gabriel Damasco, Cyrus Harp, John Lokvam and Gregory Crutsinger for editing papers and grants, providing support and feedback and always being available with an ear or for a beer. I owe a special thanks to Felipe Zapata who I had the pleasure of sharing an office with. Thank you for your advice, conversation and support throughout the years.

Thanks to Kaitlin Maguire, Eve Robinson, Adam Roddy, Bier Kraichak, Jean Alupay, Elaine Vo, Matt Guilliams, for camaraderie and dancing. Graduate school would not have been the same without you. Also, Thanks to Jill Gregory-Ames, Sami Rafi, Susan Fawcett, David Urquiza, and Angel Raygada, for friendship and for sharing their love of the Amazon.

I would like to gratefully acknowledge the Ministerio del Ambiente of Peru for providing research and export permits, Carlos Rivera of SERNANP-Allpahuayo-Mishana and the Instituto de Investigaciones de la Amazonía Peruana (IIAP) for institutional support. This research was funded by the National Science Foundation Doctoral Dissertation DEB-1311117, Sigma Xi Grant in Aid of Research, Integrative Biology Summer Grants, and the Tinker Summer Field Research Grant.

Finally, I would like to thank Nyree Zerega and Jeremie Fant from the Chicago Botanic Garden for providing me with such amazing mentorship at the beginning of my scientific career and for their ongoing support, and friendship.

TABLES, FIGURES & APPENDICES

Chapter 1

Table 1.1 Characteristics of seventeen microsatellite loci developed in *Protium subserratum*. 3

Table 1.2 Summary statistics for 17 microsatellite loci for three populations of *Protium subserratum* found on different soil types. 4

Appendix 1.1 Location, soil types, and voucher specimens of three natural populations of *Protium subserratum*. 4

Chapter 2

Table 2.1 Sampling information and summary statistics for all populations sampled. 18

Table 2.2 Analysis of molecular variance for populations of *Protium subserratum* grouped by geography, soil type, and by soil type excluding white-sand 19

Table 2.3 Pairwise F_{ST} values for all population pairs. 19

Table 2.4 Maximum likelihood (ML) migration (M) estimates and their 95% confidence intervals for all population pairs. 20

Figure 2.1 Sample sites and soil types for populations of *P. subserratum* in the region of Loreto, Peru. 21

Figure 2.2 Principal components analysis (PCA) of leaf morphological characters for *P. subserratum* individuals. 22

Figure 2.3 Evolutionary clusters (K=2-4) inferred from STRUCTURE analysis of 201 *P. subserratum* individuals from populations of white-sand, brown-sand and clay soil habitats. 23

Figure 2.4 Average log probabilities for K = 1 – 8. 24

Figure 2.5 Comparison of P_{ST} and F_{ST} values and their 95% confidence intervals. 25

Figure 2.6 Gene flow (as migrants per generation, $M = \theta m/\mu$) estimated in Migrate-N. 26

Chapter 3

Table 3.1 Geographic coordinates, number of individuals, number of genera, and number of species collected at each trapping site as well as total abundance and richness for each habitat type and collection site. 38

Figure 3.1 Sample sites and soil types where stingless bees were sampled in the region of Loreto, Peru. 39

Figure 3.2 Species accumulation curve for stingless bees for all sites sampled. 40

Figure 3.3 Species abundance distribution for all stingless bee species for all sites sampled. 41

Figure 3.4 Results from distance-based redundancy analysis. Gray ellipses correspond to white-sand habitat and black ellipses correspond to non-white-sand habitat. 42

Chapter 4

Table 4.1 Formulas for individual components of reproductive isolation (RI), isolation measure, and the range of values possible for that barrier. 55

Table 4.2 Microsatellite summary statistics for populations sampled. 56

Table 4.3 Total number of pollinator visits observed at white-sand and brown-sand ecotypes of *P. subserratum.* 56

Figure 4.1 The proportion of white-sand and brown-sand individuals of *P. subserratum* in flower across time for January 2006 – December 2009. 57

Figure 4.2 a. Average number of adhered pollen grains per pollination for parental crosses and hybrid hand-crosses. **b.** Average proportion of adhered pollen grains that germinated pollen tubes in parental and hybrid hand-crosses. 58

Figure 4.3 Proportion of hybrid hand-crosses and parental hand-crosses, which developed seeds two weeks after pollination. 59

Figure 4.4 a. Strength of each individual barriers to reproduction (RI). **b.** Relative contribution of each barrier to reproduction to total reproductive isolation (T). **c.** Relative contribution of each barrier to reproduction to total reproductive isolation for individuals found at a contact zone. 60

CHAPTER 1: Microsatellite primers for an Amazonian lowland tropical tree, *Protium subserratum* (Engl.) Engl. (Burseraceae)

ABSTRACT

Premise of the study: The first microsatellite primers were developed for *Protium subserratum* (Engl.) Engl., a widespread Amazonian tree, to investigate genetic differentiation between populations found on clay, brown-sand and white-sand soils.

Methods and Results: Seventeen primer pairs were identified from two individuals of *Protium subserratum,* found on white-sand and brown-sand soil types. Polymorphism was analyzed in 63 individuals from a total of three populations, each found on a different soil type. The primers amplified tetra-, tri- and dinucleotide repeats with three to twenty four alleles per locus. Excluding monomorphic loci, observed and expected heterozygosities ranged from 0 to 0.852 and 0.036 to 0.901 respectively.

Conclusions: These new microsatellite markers will be useful in studies of genetic diversity, population differentiation and gene flow across habitat types in *P. subserratum*.

INTRODUCTION

Protium subserratum (Engl.) Engl. (Burseraceae) is a widespread Neotropical tree found across the lowland Amazon Basin (Daly and Fine, 2011). It is dioecious with flowers that are small, white, nectar-producing and odiferous, indicating generalist pollinator affinities (Daly, 1987). *Protium subserratum* represents one of the few soil generalist species in the genus (Fine et al., 2005), with two morphologically differentiated sub-populations endemic to the widespread and relatively fertile clay and brown-sand forests as well as to the comparatively rare patches of nutrient poor white-sand forest habitats (Daly and Fine, 2011). A recent phylogeographic study comparing measurements of leaf traits in individuals from all three habitat types demonstrated that *P. subserratum* found in white-sand habitats differed morphologically in vegetative traits from populations found on terrace and clay soil types. Nuclear sequence data from the same study also showed that populations of *P. subserratum* from geographically distant clay and terrace soils were more closely related to each other than they were to nearby white sand populations (Fine et al., 2012). While these results are consistent with the idea that adaptations to different soil types may be playing an important role in population divergence, microsatellite markers will provide more powerful tools to examine fine scale population differentiation and gene flow.

METHODS AND RESULTS

Genomic DNA from two individuals of *P. subserratum* from a brown-sand population and a white-sand population (Appendix 1.1) were sent to the Savannah River Ecological Laboratories (SREL) at the University of Georgia for microsatellite marker development and primer design. Microsatellite markers were developed according to the protocol developed by Glenn and Schabel (2005). At SREL genomic DNA was combined, digested, ligated with linkers SimpleXL12_U (5'- AAAGCTGGCGTCGAAGT -3') and SimpleXL12_Lp (5'-pACTTCGACGCCAGC -3'), enriched with biotinylated probes and recovered via polymerase chain reaction (PCR). The enriched library was then sequenced on a 454 using titanium chemistry (454 Life Sciences, a Roche company, Branford CT, USA). A total of 6,123 sequences were obtained. A total of 2,201 reads, identified using MSATCOMMANDER version 0.8.1 (Faircloth, 2008), contained microsatellite repeats suitable for primer design. Primers were

designed using Primer3 (Rozen and Skaletsky, 2000) to utilize a three-primer PCR protocol (Schuelke, 2000). One primer from each pair was modified with the addition of an M13R tag (5'-GGAAACAGCTATGACCAT-3') to enable the use of a third universal primer (identical to the M13R tag), fluorescently labeled for detection. The sequence GTTT was added to primers without the M13R tag to promote adenylation.

One hundred and thirteen primer pairs were tested for amplification and polymorphism using DNA obtained from ten individuals of *P. subserratum*, five from a brown-sand population and five from a white-sand population. PCR amplifications were performed in a total reaction volume of 12.5 µl containing 6.5 µl 2x GoTaq Green Master Mix (400 µM dNTPs, 3 mM $MgCl_2$ and 1 unit of Taq DNA polymerase) (Promega, Madison, Wisconsin, USA) 0.6 µM un-tagged primer, 0.3 µM tag-modified primer, 0.3 µM M13R primer fluorescently labeled with either 6-FAM or Hex, 1 µl of undiluted DNA template and DNase free water. Amplifications for all loci were conducted using a touchdown PCR protocol beginning with an initial denaturation step of 2 min and 30 s at 95°C followed by 30 cycles 95°C for 30 s, annealing at a temperature of 60°C for 30s (decreased by 0.5°C per cycle), 72°C for 1 min; and 30 cycles of 95°C for 30 s, 45°C for 30 s 72°C for 1 min. A final extension was done at 72°C for 10 min. Amplification products were co-loaded on an ABI 3730 DNA analyzer (Applied Biosystems, Foster City, California, USA) with 0.3 µl GS-500 LIZ size standard (Applied Biosystems) to allow allele length sizing. Electrophoretic results were initially scored using Peak Scanner v. 1.0 (Applied Biosystems) followed by visual confirmation. Seventeen of the tested primer pairs amplified high quality PCR product that exhibited polymorphism (Table 1.1).

We assessed the variability of the seventeen polymorphic loci in 63 individuals from three different populations. One voucher specimen per population was deposited at the University of California, Berkeley University Herbarium (Appendix 1.1). Conditions and characteristics for each of the seventeen loci are presented in Tables 1.1 and 1.2. The presence of null alleles was tested using MICRO-CHECKER version 2.2.3 (van Oosterhout et al., 2004). Number of alleles per locus (*A*), observed and expected heterozygosity (H_o and H_e), Shannon's information index (*I*) and deviations from Hardy-Weinberg equilibrium (HWE) were estimated using GenAlEx version 6.4 (Peakall and Smouse, 2006). P-values for tests of deviation from HWE were adjusted using a sequential Bonferroni correction (Rice, 1989).

While each population has at least one monomorphic locus all loci displayed polymorphism when compared across populations with the total number of alleles ranging from three to twenty four alleles per locus. Loci Prot02, Prot08, Prot102, and Prot22 tested positive for the presence of null alleles across all populations. Observed heterozygosity ranged from 0 to 0.852 and expected heterozygosity ranged from 0.036 to 0.901. Locus Prot22 significantly deviated from HWE in the Lagunas and Porvenir 1 populations, Prot02 significantly deviated from HWE in population Porvenir 1 and locus Prot08 showed significant deviation from HWE in population Porvenir 2 ($P< 0.05$). Genetic diversity as measured by Shannon's diversity index ranged from 0.319 to 2.172.

DISCUSSION

These 17 new microsatellite loci described here will provide useful tools in future studies of genetic diversity, gene flow, and genetic differentiation across habitat types in the Amazon tree, *P. subserratum.*

TABLE 1.1 Characteristics of seventeen microsatellite loci developed in *Protium subserratum*.

Locus	Primer sequence (5'- 3')[a]	Repeat motif	T_a (°C)	T_m (°C)	Size range	GenBank accession no.
Prot02	F: *ATAAACCCTCTTACGGTGAG	$ACAT_{(6)}$	60/45	59.2	174-224	JX014415
	R: ◇GGGATTTGTTGACTTTGAAC			59.6		
Prot08	F: ◇TATGTCCCACAATGATCCTC	$ACAT_{(7)}$	60/45	59.9	184-239	JX014416
	R: *TTTATAGGAGCGCTCTGATC			60.4		
Prot13	F: *TGATTTCTTGTCCCAAAGAG	$AAAG_{(7)}$	60/45	60.2	305-313	JX014417
	R: ◇AGCCACATACCGATAAACTC			59.7		
Prot22	F: *TAACCCTTGACAAGCATTTC	$AGAT_{(9)}$	60/45	60.7	336-364	JX014418
	R: ◇AAATTACGGCTTCAGAATTG			60.7		
Prot28	F: ◇CGCAGTTTCAGAAATATCAG	$ACAT_{(8)}$	60/45	58.7	230-251	JX014419
	R: *GCATGATTCGATGTTATAGG			58.4		
Prot29	F: *TGAACTGACCTTTGCATGAC	$AAAG_{(15)}$	60/45	61.1	121-168	JX014420
	R: ◇AAGAGGGTGGTCTGAACTG			59.7		
Prot67	F: ◇TCATGCTGTAATTCCCTGTC	$ACAT_{(7)}$	60/45	60.6	198-234	JX014421
	R: *GAGAAGAGCAAAGATTCGATAG			60.0		
Prot70	F: *CCATTATTAAGCATGCAAAC	$AAG_{(9)}$	60/45	59.4	113-262	JX014422
	R: ◇CAATGGCCTGTTCATATAAAG			60.3		
Prot71	F: *CCATCCTCAGCTCTTACTTTC	$AAG_{(8)}$	60/45	60.7	402-414	JX014423
	R: ◇GATCGGTCACAGATTCAATG			60.3		
Prot78	F: ◇CACACCAGGAAAGACTCAAG	$AAG_{(8)}$	60/45	59.9	137-143	JX014424
	R: *TTGGAAGGAGGATTTATAGG			59.5		
Prot83	F: ◇CGTCTGGATGGAAGATAAAG	$AAC_{(10)}$	60/45	60.0	168-186	JX014425
	R: *TCCTCGTTCTCCACTACAAC			59.6		
Prot97	F: *ATTCCGATTAAACCTCATTC	$AG_{(13)}$	60/45	59.2	150-163	JX014426
	R: ◇GGGTATGAGCTTGAATTAGG			59.7		
Prot99	F: *ATGCTATGATAATCGGTTCC	$AG_{(12)}$	60/45	59.8	159-201	JX014427
	R: ◇GAAATGGTTGCACTTCACTC			60.1		
Prot100	F: *ATCTCTCGTTCCAACTCAAC	$AC_{(10)}$	60/45	58.7	158-184	JX014428
	R: ◇CGTCGAAGTACTCACCACTC			59.9		
Prot101	F: ◇CATTTAGGGACCACGTTTAC	$AG_{(11)}$	60/45	60.3	274-290	JX014429
	R: *ATTGTTCCAGGATCTAGGTG			59.0		
Prot102	F: ◇GTCGACCAAATAATGTCACC	$AG_{(15)}$	60/45	60.0	350-378	JX014430
	R: *ATGGACACACAGGACCTATC			58.7		
Prot104	F: *TAACCGCAATATCAACTCTC	$AC_{(11)}$	60/45	58.2	271-289	JX014431
	R: ◇ACACCACGACTAAAGACTGG			59.7		

Note: T_a = annealing temperatures used for touchdown cycling; T_m = calculated melting temperatures.
[a] * indicates M13R tag (5'- GGAAACAGCTATGACCAT -3'); ◇ indicates GTTT tag.

TABLE 1.2. Statistical analysis of 17 microsatellite loci for three populations of *Protium subserratum* found on different soil types.

	Lagunas (*N*=15)				Porvenir 1 (*N*= 27)				Porvenir 2 (*N*=21)			
Locus	*A*	H_o	H_e	*I*	*A*	H_o	H_e	*I*	*A*	H_o	H_e	*I*
Prot02	7	0.533	0.791	1.717	9	0.250*	0.837	1.965	7	0.636	0.796	1.735
Prot08	10	0.538	0.876	2.172	13	0.565	0.851	2.181	14	0.333*	0.901	2.452
Prot13	1	Monomorphic				Monomorphic			3	0.591	0.483	0.843
Prot22	4	0*	0.625	1.127	3	0.148*	0.427	0.677	5	0.381	0.546	1.093
Prot28	6	0.467	0.709	1.407	2	0.037	0.036	0.092	5	0.810	0.687	1.308
Prot29	6	0.733	0.602	1.280	3	0.519	0.524	0.814	6	0.909	0.731	1.492
Prot67	8	0.733	0.760	1.688	5	0.593	0.676	1.249	6	0.857	0.753	1.563
Prot70	7	0.800	0.698	1.521	4	0.370	0.366	0.690	1	Monomorphic		
Prot71	3	0.375	0.461	0.777		Monomorphic			2	0.176	0.251	0.418
Prot78	3	0.583	0.531	0.829		Monomorphic			2	0.409	0.375	0.562
Prot83	3	0.476	0.438	0.716	4	0.308	0.353	0.677	4	0.286	0.294	0.610
Prot97	5	0.467	0.556	1.112	6	0.778	0.695	1.329	4	0.318	0.350	0.704
Prot99	7	0.467	0.671	1.367	7	0.826	0.750	1.576	8	0.682	0.784	1.738
Prot100	5	0.467	0.673	1.319	5	0.320	0.510	1.054	4	0.333	0.579	1.009
Prot101	4	0.400	0.344	0.703	4	0.444	0.490	0.929	4	0.714	0.695	1.237
Prot102	6	0.533	0.798	1.674	6	0.852	0.694	1.418	8	0.700	0.825	1.866
Prot104	3	0.600	0.518	0.802	2	0.400	0.365	0.551	3	0.545	0.464	0.725

Note: H_e = expected heterozygosity; H_o = observed heterozygosity; *N* = number of samples; *I* = Shannon's Information Index. *Indicates that the observed heterozygosity is significantly departed from the expected heterozygosity under Hardy-Weinberg equilibrium after sequential Bonferonni corrections (P<0.05).

APPENDIX 1.1. Location, soil types, and voucher specimens of three natural populations of *Protium subserratum*. Voucher specimens are deposited in the University Herbarium, University of California, Berkeley.

Populations (locale)	Geographical coordinates (degrees decimal)	Soil type	Voucher specimens
Lagunas (Loreto, Peru)	Lat: -3.8303233202193137, Long: -73.59467610059791	Clay	TM28
Porvenir1 (Loreto, Peru)	Lat: -3.915609635759378, Long: -73.55211354953575	White sand	TM29
Porvenir2 (Loreto, Peru)	Lat: -3.8966818509317767, Long: -73.53723449019981	Brown sand	TM30

WORK CITED

DALY, D. C. 1987. A taxonomic revision of *Protium* (Burseraceae) in eastern Amazonia and the Guianas. Ph.D. dissertation, City University of New York, New York.

DALY, D. C., P. V. A., FINE. 2011. A new Amazonian section of *Protium* (Burseraceae) including both edaphic specialist and generalist taxa. Studies in Neotropical Burseraceae XVI. *Systematic Botany* 36: 939-949.

FAIRCLOTH, B. C. 2008. MSATCOMMANDER: Detection of microsatellite repeat arrays and automated, locus-specific primer design. *Molecular Ecology Resources* 8: 92-94.

FINE, P. V. A., D. C. DALY, G. V. MUNOZ, I. MESONES, and K. M. CAMERON. 2005. The contribution of edaphic heterogeneity to the evolution and diversity of Burseraceae trees in the western Amazon. *Evolution* 59: 1464-1478.

FINE, P. V. A., F. ZAPATA, D. C. DALY, I. MESONES, T. M. MISIEWICZ, H. F. COOPER and C. E. A. BARBOSA. 2012. The importance of environmental heterogeneity and spatial distance in generating phylogeographic structure in edaphic specialist and generalist tree species of *Protium* (Burseraceae) across the Amazon Basin. *Journal of Biogeography* doi: 10.1111/j.1365-2699.2011.02645.x

GLENN, T. C., and N. A. SCHABLE. 2005. Isolating microsatellite DNA loci. *Methods in Enzymology* 395: 202-222.

PEAKALL, R., and P. E. SMOUSE. 2006. GENALEX 6: genetic analysis in Excel. Population software for teaching and research. *Molecular Ecology Notes* 6: 288-295.

RICE, W. R. 1989. Analyzing tables of statistical tests. *Evolution* 43: 223-225.

ROZEN, S. and H. J. SKALETSKY. 2000. Primer3 on the WWW for general users and for biologist programmers. In S. Krawetz and S. Misener [eds.], Bioinformatics methods and portocols: Methods in molecular biology, 365-386. Humana Press, Totowa, New Jersey, USA.

SCHUELKE, M. 2000. An economic method for the fluorescent labeling of PCR fragments. *Nature Biotechnology* 18: 233-234.

VAN OOSTERHOUT, C., W. F. HUTCHINSON, D. P. M. WILLS, and P. SHIPLEY. 2004. MICRO-CHECKER: Software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes* 4: 535-538.

CHAPTER 2: Evidence for ecological divergence across a mosaic of soil types in an Amazonian tropical tree: *Protium subserratum* (Burseraceae)

ABSTRACT

Soil heterogeneity is an important driver of divergent natural selection in plants. Neotropical trees have the highest diversity on earth and frequently soil specialist congeners are distributed parapatrically. While the role of edaphic heterogeneity in the origin and maintenance of tropical tree diversity is unknown it has been posited that natural selection across the patchwork of soils in the Amazon rainforest is important in driving and maintaining tree diversity. We examined genetic and morphological differentiation among populations of the tropical tree *Protium subserratum* growing parapatrically on the mosaic of white-sand, brown-sand and clay soils found throughout western Amazonia. Nuclear microsatellites and leaf morphology were used to (1) Quantify the extent of phenotypic and genetic divergence across habitat types (2) Assess the importance of natural selection vs. drift in population divergence (3) Determine the extent of hybridization and introgression across habitat types (4) Estimate migration rates among populations. We found significant morphological variation correlated with soil type. Higher levels of genetic differentiation and lower migration rates were observed between adjacent populations found on different soil types than between geographically distant populations on the same soil type. P_{ST}-F_{ST} comparisons indicate a role for natural selection in population divergence among soil types. A small number of hybrids were detected suggesting that gene flow among soil specialist populations may occur at low frequencies. Our results suggest that edaphic specialization has occurred multiple times in *P. subserratum* and that divergent natural selection across edaphic boundaries may be a general mechanism promoting and maintaining Amazonian tree diversity.

INTRODUCTION

The role of gene flow and natural selection in the origin and maintenance of species diversity has been a long-standing point of contention in the field of evolutionary biology (Haldane 1948; Ehrlich & Raven 1969; Endler 1977). Traditionally, theory has posited that gene flow among populations will act to homogenize them, limiting local adaptation and maintaining species cohesion (Haldane 1948; Hendry et al. 2001). Therefore, prevention of gene flow through geographic isolation is thought to be the first step in species divergence, with reproductive isolation among populations eventually developing by genetic drift (Mayr 1963). More recently however, there has been an increasing acknowledgement that divergent adaptation across heterogeneous environments can also lead to lineage divergence. Population divergence driven by natural selection can occur in the absence of gene flow if geographic barriers are present or in the face of gene flow if natural selection itself is strong enough to cause significant fitness differences across habitat boundaries. As long as gene flow does not swamp population differentiation, barriers to reproduction can develop, further promoting population divergence and eventually leading to irreversible reproductive isolation (Rundle & Nosil 2005; Nosil 2012).

Genetic differentiation across heterogeneous environments has been well documented in temperate plant populations with some of the best-known examples stemming from local adaption to variable soils (Antonovics & Bradshaw 1970; Rosenthal *et al.* 2005; Anacker *et al.* 2011). This suggests that specialization across edaphically heterogeneous environments could result from selective pressures that are strong enough to promote population level divergence even when plant populations are not geographically isolated. Nevertheless, gene flow can

become limited among populations for many other reasons and a variety of evolutionary processes can result in genetic heterogeneity across populations (Latta 2004). The processes of genetic drift and natural selection however are predicted to leave markedly different genetic signatures across populations. Drift is expected to generate stochastic population genetic structure, which is expected to correlate with geographic distance when it is sufficient to limit migration and gene flow between populations (Latta 2004). Alternatively, if natural selection is important in population divergence then we expect to see a pattern of 'Isolation-by-Adaptation' (IBA) (Nosil *et al*. 2007; Nosil *et al.* 2009). The expected signal for IBA is similar to that of isolation-by-distance except that the level of population differentiation is dictated by environmental similarity as opposed to geographic proximity. Theory predicts that when differential adaptation across ecological gradients leads to a reduction in gene flow, populations found in differing habitats will experience lower effective migration rates and greater genetic differentiation than populations found in ecologically similar habitats. Reproductive isolation, either driven or reinforced by natural selection over a long period of time, will eventually have a genome wide effect on allele frequencies (Grahame *et al.* 2006; Nosil *et al.* 2009; Thibert-Plante & Hendry 2010). Therefore, we can begin to discern the relative importance of these two forces in driving and maintaining differentiation by comparing patterns of genetic divergence relative to variables proximate to genetic drift (Euclidean distance) and natural selection (morphological and environmental variables) (Hendry & Taylor 2004; Nosil *et al.* 2007).

The lowland Amazon rainforest has the highest tree diversity in the world and is also characterized by steep soil gradients yet explanations for the origin of Amazonian species have typically revolved around hypotheses of non-ecologically based divergence by genetic drift (reviewed in Haffer 2008). More recently, an alternative hypothesis of tropical tree speciation that is driven and/or maintained by natural selection across ecological boundaries has also been gaining attention in part due to the observation that congeners across diverse plant families are often found parapatrically on different soil types (Gentry 1988; Tuomisto *et al.* 1995; Fine *et al.* 2005; Fine *et al.* 2010).

Amazonian soils can strongly differ with regards to abiotic factors such as particle size, nutrient content and moisture level and these differences can have a direct effect on trees found in different edaphic environments (Pregitzer *et al.* 2010; Smith *et al.* 2011). Moreover, physical and microclimatic variation among edaphically differentiated forest patches may influence the spatial organization of other organisms (Sääksjärvi *et al.* 2004; Álvarez Alonso *et al.* 2013). As a result, biotic interactions in the form of seed dispersal, herbivory, pollination services, or pathogen prevalence, may interact with abiotic differences to further magnify natural selection across ecological gradients for tropical trees. Yet, while correlations between edaphic shifts and lineage divergence have been demonstrated in a broader phylogenetic context (Fine *et al.* 2005), very few studies have investigated the importance of edaphic heterogeneity in tropical tree diversification at finer taxonomic scales.

Peru's Amazonian lowland rainforest provides an ideal system to examine the role of divergent natural selection in tree diversification because it contains a patchwork of dramatically different soil types each home to distinct tree communities (Fine *et al.* 2005; Fine *et al*. 2010). The genus *Protium* (Burseraceae) is known to occur on a variety of soil types throughout the Neotropics (Daly 1987). It is a diverse clade of ca. 140 species of trees with ca. 100 species found in the lowland Amazon basin, many of which are soil specialists (Fine *et al.* 2005). Edaphic specialization has evolved independently multiple times within the genus, and species adapted to particular soil types exhibit slower growth and higher mortality when grown outside

of their respective habitats (Fine *et al.* 2006). This supports the idea that physiological tradeoffs accompany adaptations to alternative soil types (Fine *et al*. 2004). *Protium subserratum* (Engl.) is one of the few soil generalists in the genus, although the taxon is more accurately described as an incipient species complex with genetically and morphologically differentiated populations endemic to white-sand and non-white-sand soils (Daly & Fine 2011).

A recent phylogeographic study suggested that populations of *P. subserratum* found on white-sand forests that were separated by over 200km were more genetically similar to one another than to populations found on brown-sand and clay soils and that some haplotypes were shared between populations through ongoing gene flow or incomplete lineage sorting (Fine *et al*. 2013a). However, these results were based on DNA sequences from only three nuclear loci with few phylogenetically informative sites and low sample sizes. While these data were valuable for inferring broader evolutionary histories over a large geographic range they were not able to provide insight regarding population dynamics.

Here, we conduct fine-scale analyses using nuclear microsatellites, morphological measurements, and a sampling scheme that includes dense sampling of multiple parapatric population pairs found on the mosaic of white-sand, brown-sand and clay soil types distributed at varying distances across 100 km in the Peruvian Amazon. This experimental design represents a powerful natural experiment to simultaneously test the relative importance of habitat and geographic distance in limiting gene flow and driving divergence among edaphic specialist populations of Amazonian trees. If divergent natural selection across habitat boundaries is strong enough to impede gene flow we expect to find stronger patterns of phenotypic and genetic differentiation in populations found on different habitat types and lower levels of gene flow across habitat boundaries compared to geographically isolated populations found in the same habitat type. Alternatively, if genetic drift has led to barriers to gene flow among populations we would expect to see little to no phenotypic variation among populations, and strong patterns of isolation-by-distance, where populations that are more geographically distant from one another are also more genetically distinct than populations in close geographic proximity to each other regardless of habitat type. Moreover, because neutral forces such as genetic drift are expected to drive variation at putatively neutral loci and natural selection is expected to drive phenotypic variation, comparisons of phenotypic variation and neutral genetic variation across habitat boundaries can provide even further insight into the role of non-neutral evolutionary forces as drivers of divergence (Merilä & Crnokrak 2001; Leinonen *et al.* 2006). If stochastic mechanisms are driving population divergence we expect levels of phenotypic variation among habitats to equal that of neutral genetic variation. If divergent natural selection across habitat boundaries is driving divergence then phenotypic variation is expected to be greater than neutral genetic variation (Merilä & Crnokrak 2001; Leinonen *et al.* 2006).

We specifically addressed four major questions in this study: (1) Are populations phenotypically and genetically differentiated across all three habitat types? (2) Can we detect a signature of natural selection over drift divergence among populations on different soil types? (3) What is the extent of hybridization and introgression across habitat types? (4) What is the relative importance of spatial distance and soil type in structuring *P. subserratum* tree populations and influencing migration rates among populations?

METHODS

Study System

Habitat Types: Non-flooded forests found in the Peruvian Amazon have been classified into three broad categories based on soil type (Fine *et al*. 2005), which can be differentiated largely by their nutrient availability and geologic history (a more thorough discussion can be found in Hoorn 1993; Fine *et al.* 2005 and Frasier *et al*. 2008): (i) White-sand soils, which are extremely nutrient poor, include stunted canopies, and exist today as geographically isolated habitat islands, often covering only a few square hectares (ii) Clay soils which represent the most nutrient rich soils in the western Amazon and have the highest water retention and tallest canopies (iii) Brown-sand soils which have significantly higher nutrient availability than white-sand soils (Fine *et al.* 2005) and the height of their forest canopies represent an intermediate between white-sand and clay soil forests.

Focal plant taxa: *Protium subserratum* is part of the section *Papilloprotium*, which is comprised of four taxa that include both edaphic specialists and generalists (Daly & Fine 2011). It is sister to two white-sand specialist taxa, *P. alvarezianum* and *P. reticulatum* found in white-sand forest patches in the Rio Negro Basin of Venezuela and Brazil (Daly & Fine 2011). *Protium subserratum* is a soil generalist that is common and widespread across the lowland Amazon basin. It has small (~ 4mm length), fragrant white flowers, which are nectiferous and relatively large (~ 1.5cm diameter) red fruits (Misiewicz per obs). Little quantitative data regarding the reproductive biology of *P. subserratum* exists, but a variety of stingless bees have frequently been observed visiting flowers (Misiewicz 2014) and monkeys and large birds are hypothesized to be potential seed dispersers (Daly 1987). While no consistent morphological differences have been observed in flower or fruit characters among populations, significant differences in vegetative characters do exist (Daly & Fine 2011; Fine *et al.* 2013a). Within *P. subserratum* vegetative morphological variation has been noted both across the range as well as within localized populations (Daly & Fine 2011). Daly & Fine (2011) grouped individuals into four distinct morphotypes based primarily on leaf morphology. Two of these morphotypes have very restricted geographic distributions: Morphotype 1 is restricted to non-white-sand forests in French Guiana, and Morphotype 4 is restricted to Colombia's Caquetá. The other two have more widespread distributions: Morphotype 2 is associated with clay and brown-sand soils of the central and western Amazon whereas Morphotype 3 is consistently associated with white-sand soils in the Western Amazon (Daly & Fine 2011). Phylogenetic analysis by Daly & Fine (2011) demonstrated that these morphotypes do not form monophyletic clades and should continue to be considered one taxonomic species.

Further phylogeographic analysis by Fine *et al.* (2013a) included populations found on white-sand and non-white-sand soils sampled throughout the Amazon. They found two well supported clades, one composed of northern Amazonian individuals from non-white-sand soils in Guyana and French Guiana and the other composed of western Amazonian individuals and a single individual from Guyana. Within the western clade, they found groupings composed of Peruvian white-sand individuals, Brazilian non-white-sand individuals, and Peruvian non-white-sand individuals, however none of these had posterior probability support higher than 0.75. These results are consistent with the previous classification of *P. subserratum* as a single species and no evidence of differentiation between individuals collected on brown-sand or clay soil was detected.

While *P. subserratum* morphotypes are not monophyletic, a two-year reciprocal transplant study demonstrated that morphological variation observed between seedlings associated with white-sand and non-white-sand habitats is not completely due to plasticity (Fine

et al. 2013b). Seedlings from three white-sand populations and three non-white-sand Peruvian populations (including both clay and brown-sand soils) were collected and transplanted in white-sand and clay soil habitats. Results indicated that seedlings initially collected from white-sand habitat grew slower and produced fewer leaves in both habitat types than seedlings initially collected from non-white-sand habitats. They also demonstrated that that leaf pubescence on new leaf growth was not influenced by habitat type. Leaflet thickness on the other hand did appear to be plastic. Additionally, there are clear quantitative and qualitative differences in secondary chemical compounds between populations found in white-sand and non-white-sand soil habitats (Fine *et al.* 2013b).

Study Sites and Sampling

Five study sites containing a total of eight populations (n=5-54) of *P. subserratum* growing on white-sand, brown-sand, and clay soil types were established in Loreto, Peru (Fig. 2.1). Individuals found in white-sand habitats corresponded to Morphotype 3 in Daly & Fine (2011) and the Peruvian white-sand morphotype from populations sampled in Fine *et al.* (2013a). Individuals found on brown-sand and clay soil types corresponded to Morphotype 2 described by Daly & Fine (2011) and the Peruvian non-white-sand morphotype from Fine *et al.* (2013a). Furthermore, seedlings from four of the eight populations sampled in this study (WS-A, Clay-A, WS-B, and BS-B (Fig. 2.1)) were included in the reciprocal transplant experiment published by Fine *et al.* (2013b). Adult individuals of *P. subserratum* from eight populations were tagged, mapped and collected in silica for DNA extraction (N=201). Voucher specimens for each population were deposited in the Herbarium Amazonense at the Universidad Nacional de la Amazonía Peruana in Iquitos, Peru (AMAZ) and the University Herbarium at the University of California, Berkeley (UC). To avoid variation in leaf morphology based on age and canopy position, samples used for morphological measurements were only collected from individuals 10m or taller (N=163). Three leaves were collected from each individual, pressed and dried for later processing. All of these individuals were also included in the genetic sampling.

Characterization of phenotypic variation

In order to investigate the extent to which phenotypic differences in leaf morphology were correlated with white-sand, brown-sand and clay soil habitats we characterized variation in eight leaf morphological characters from 163 adult individuals across all eight populations. Measurements were taken three times and averaged. They included leaf length, number of leaflets per leaf, leaflet length, leaflet width, leaflet thickness, number of margin serrations per leaflet, pubescence percent coverage on the abaxial side of the leaflet blade and the leaflet midrib. Pubescence percent coverage of abaxial leaflet blade and leaflet midrib was visually estimated using a dissecting microscope to the nearest 10% within a haphazardly placed three by three millimeter square.

Morphological differentiation among individuals found on white-sand, brown-sand and clay soil was assessed in R (R Development Core Team 2010) using principal components analysis (PCA) and multivariate analysis of variance (MANOVA).

Microsatellite genotyping

Genomic DNA was extracted from the leaf material of all adult individuals and extractions were carried out using a Qiagen DNeasy Plant Mini Kit (Valencia, CA). Genotypes were determined using thirteen nuclear microsatellite markers developed for *P. subserratum* (prot13, prot28, prot29, prot67, prot70, prot78, prot83, prot97, prot99, prot100, prot101, prot102, prot104), following the protocols described in Misiewicz *et al.* (2012).

Genetic variation, Hardy-Weinberg equilibrium, null alleles and linkage disequilibrium

Summary statistics including number of alleles (A), and observed, expected and unbiased expected heterozygosities (Ho He, and UHe) were estimated for each population using GenAlEx v6.4 (Peakall & Smouse 2006). We used all loci equalized to a sample size of five individuals, the number of individuals in our smallest population, to calculate rarefied allelic richness (A_R) and private allele richness (A_P) using the allelic diversity analyzer ADZE (Szpeich *et al.* 2008). Values were calculated as averages across all loci for each population and additional A_P values were calculated for combinations of populations grouped by soil type. The inbreeding coefficient (F_{IS}) for each population was calculated across all loci, as were deviation from Hardy-Weinberg equilibrium (HWE) and linkage disequilibrium (LD), using GENEPOP v4.0 (Raymond & Rousset 1995). Deviation from HWE and LD were calculated using the probability test with 10^4 demorizations, 10^4 batches, and 10^4 iterations per batch. Significance values were adjusted using sequential Bonferroni corrections (Rice 1989). All data were screened for genotyping errors due to stutter, large allele dropout and the presence of null alleles, using the software Micro-Checker (van Oosterhout *et al*. 2003).

Population genetic structure and differentiation across soil types

If natural selection is more important than genetic drift in driving population differentiation then population genetic structure should be more strongly correlated with turnover in habitat type than with geographic distance. Hence, we sought to understand how populations are genetically structured using a combination of descriptive statistics including F-statistics, analysis of molecular variance (AMOVA), and Bayesian clustering analysis. Additionally, we explicitly explored the extent to which genetic variation among populations could be explained by geographic distance and soil type using Mantel and partial Mantel tests. Genetic differentiation among populations was assessed using population level pairwise comparisons of θ, analogous to Wright's F_{ST} (Weir & Cockerham 1984), calculated with and without correction for null alleles with FreeNA (Chapuis & Estoup 2007). Analysis of molecular variance (AMOVA), in which populations are grouped hierarchically to explore how groupings affect the partitioning of genetic variation (Excoffier *et al.* 1992), was calculated among populations grouped according to (i) soil type (white-sand, brown-sand and clay), (ii) brown-sand and clay soils (excluding white-sand populations) and (iii) geographic location (sites 1, 2, 3, 4 and 5) using 10^3 permutations.

Population genetic structuring was determined using Bayesian Markov chain Monte Carlo clustering analysis implemented in STRUCTURE 2.3.3 (Pritchard *et al*. 2000). STRUCTURE was run using the admixture model and assuming uncorrelated allele frequencies. In order to estimate the 'true' K a burn-in period of 500,000 generations was followed by 10^5 Markov chain Monte Carlo (MCMC) generations for each value of K = 1-8. Simulations were repeated twenty times for each value of K. We used Structure Harvester (Earl & vonHoldt 2012) to interpret the output as described by Evanno *et al.* (2005) and Pritchard *et al.* (2000). Admixture proportions were averaged over all runs using CLUMPP (Jakobsson & Rosenberg 2007) and the final matrix was visualized with DISTRUCT 1.1 (Rosenberg 2004).

Mantel and partial Mantel tests were performed in the program IBDWS v. 3.15 (Jensen *et al.* 2005), using 10,000 randomizations. Tests were performed for all populations, brown-sand and clay populations only and populations grouped by soil type. All partial Mantel tests were performed for comparisons of genetic distance and habitat type while controlling for geographic distance.

Selection vs. Drift

We assessed the potential role of divergent natural selection and genetic drift as

mechanisms underlying population differentiation by comparing phenotypic differentiation and neutral genetic differentiation for groupings of individuals across populations based on soil type. While these statistical tests are generally performed using an index of divergence for genes linked to quantitative traits (Q_{ST}) the degree of phenotypic divergence (P_{ST}) can be used as a surrogate when Q_{ST} cannot be estimated (Leinonen *et al.* 2006). If the value of P_{ST} significantly exceeds F_{ST} then a hypothesis of divergent selection is supported, whereas P_{ST} and F_{ST} values that do not differ significantly are consistent with neutral genetic differentiation (Merilä & Crnokrak 2001).

P_{ST} values were calculated as described by Leinonen *et al.* (2006). Analyses of variance (ANOVA) were used to estimate the components of phenotypic variances and calculations were carried out with heritability (h^2) defined as 0.5, where half of phenotypic variation is due to environmental and non-additive effects. Because all morphological traits measured in this study were leaf traits, and therefore non-independent, P_{ST} values for morphology were estimated from our first three principal components (PC). Confidence intervals (CI) were estimated using 1000 bootstrap replicates over individuals in R (http://www.r-project.org/). P_{ST} values and multilocus FST values were calculated and compared for individuals grouped by soil type. Comparisons were considered significantly different when their 95% CI did not overlap.

Hybridization and introgression

We tested for hybridization and introgression between edaphically differentiated population pairs of *P. subserratum* that contained individuals showing significant admixture in our population structure analysis with the model-based clustering method of NewHybrid version 1.1b (Anderson & Thompson 2002). We restricted our analysis to two generations; this resulted in six possible hybrid classes (two parental classes, F_1 and F_2 hybrids and two backcrosses between F_1 hybrids and the parental types). Prior information about parental species from the Structure was included in the analysis. All data was analyzed multiple times from over-dispersed starting values, over 2^6 sweeps after a burn-in of 500,000 sweeps, following the recommendations of Anderson (2002). All simulations performed equally well in the discrimination of individuals based on soil type, and resulted in congruent estimations of posterior probabilities for individual hybrid class assignments. Individuals assigned to a hybrid class with a posterior probability (pp) > 90% were considered to be hybrids with parental contributions from individuals growing in different soil types.

Demographics

Migration rates as low as one migrant per generation can swamp population level differentiation (Wright 1931). Therefore, migration rates were estimated between all population pairs in order to further explore the effect of soil type and geography in limiting migration rates among populations. Migration among populations was assessed using Migrate-N, which samples coalescent genealogies, in order to calculate maximum likelihood estimates of past migration rates (Beerli 1998; Beerli & Felsenstein 1999; Beerli & Felsenstein 2001). Migration rate $(M = \theta m/\mu)$ was estimated for all population pairs. All migration rates were initially calculated between pairs of populations under a model where parameter θ, the ratio between immigration rate and mutation rate per generation, is fixed and the parameter for migration rate M is variable in order to allow for asymmetric bias in migration between populations. A maximum of 20 individuals were randomly sampled from each population for analysis with Migrate-N. All analyses were executed under a Brownian mutation model and relative mutation rates were estimated from the data, with starting values for the M and θ estimated from F_{ST} values. We tested multiple chain lengths, replicates and heating schemes to test for convergence across runs.

Final analyses consisted of 10 short chains and three long chains; the short chains ran for 50,000 generations with a sampling increment of 100 generations, and the long chains ran for 500,000 generations with a sampling increment of 100 generations. A total of 5 million genealogies were visited over the short chain runs and 50 million genealogies were visited over long chain runs. No heating scheme was used and the burn-in was set to 250,000. All analyses were then repeated as described above using the maximum likelihood estimates of parameters θ and M from the previous run as the new parameter start values. Finally, we repeated analyses for a subset of population pairs (WS-A and BS-A; WS-B and BS-B; Clay-A and Clay-C, and WS-A and WS-C) excluding locus Prot100, which tested for null alleles in more than two populations to assess its potential impact on demographic estimates. Gelman's R statistic was used to assess convergence (Beerli & Felsenstein 2001).

RESULTS

Phenotypic variation

The first three principal components accounted for 82% of variation in the data and were strongly associated with soil type (Fig 2.2). As expected from previous studies, pubescence on the leaflet midvein and blade, entire leaf margins, and thicker leaflets differentiated individuals found on white-sand soil from individuals found on brown-sand and clay soils (Fine *et al.* 2013). Furthermore, we detected previously undescribed differences between individuals found on clay and brown-sand soils where the abundance of leaflet margin serrations and the number of leaflets per leaf explained most of the variation. Individuals identified as putative older generational hybrids in hybrid assignment analyses did not reflect intermediate phenotypes with regards to leaf morphology.

Analysis of leaf traits with MANOVA showed that leaflet thickness, number of margin serrations per leaflet, pubescence percent coverage on the abaxial side of the leaflet blade and pubescence percent coverage on the abaxial side of the leaflet midrib varied significantly among individuals found on all three soil types ($p < 0.05$).

Genetic variation, Hardy-Weinberg equilibrium, null alleles and linkage disequilibrium

All summary statistics for microsatellite data are listed in Table 1. Deviations from HWE were observed in 10 of the 104 total within-population comparisons for each locus, and no particular locus or population showed a consistent pattern of deviation from HWE. We found one pair of loci (prot83 and prot100) in linkage disequilibrium (LD) in one population (WS-C); no loci were found to be in LD in any other population.

Microchecker did not detect any genotyping errors due to stutter or large allele dropout. Evidence for the presence of null alleles was found in all populations except Clay-B. No locus consistently showed signs of null alleles across populations. Four populations tested positive for null alleles at one locus (WS-B-prot100; BS-B-prot83; WS-C-prot70; and Clay-C-prot101). Population Clay-A and BS-B tested positive for null alleles at three loci (Prot99, Prot100, Prot102 and Prot83, Prot99, Prot100 respectively) and population WS-B tested positive at four loci (Prot28, Prot78, Prot100 and Prot102).

Genetic structure and ecotypic variation

Pairwise F_{ST} values revealed strong genetic differentiation across soil type and values differed only slightly when calculated with and without null allele corrections (Table 2.3). Accordingly, all subsequent analyses were carried out on the full data set.

AMOVA results confirmed that the majority of genetic variation is explained by soil type and not by geographic locality. Even when white-sand populations were excluded from the

analysis soil type continued to be important in explaining genetic variation among populations found on brown-sand and clay soils (Table 2.2).

Population structure analysis revealed strong patterns of genetic differentiation by soil type. Using the method of Evanno *et al*. (2005) the most likely model contained two clusters, consistent with populations found on white-sand or non-white-sand. Using the method of Pritchard *et al*. (2000) we found K=3 and K=4 to be the best supported models, where LnP(D) begins to asymptote (Fig. 2.3 & Fig. 2.4). The identified clusters for K=3 is consistent with populations found on white-sand, brown-sand and clay soils. Clustering for K=4 remained consistent with clustering by soil type but also showed patterns of isolation-by-distance within soil types with population WS-C distinguished from WS-A and WS-B found 100km away.

Structure analysis also revealed individuals with genetic contributions from multiple soil types. When K=2, four individuals showed admixture between white-sand and non-white-sand clusters in populations WS-A and BS-A at site 2. When K=3, the same four individuals with admixed genomes remained consistent, exhibiting admixture between white-sand and brown-sand populations and an additional individual found in population Clay-B contained admixture from both clay and brown-sand clusters. No additional individuals were identified as having significant admixture in the model where K=4 (Fig 2.3).

The Mantel test comparing all population pairs showed no correlation between genetic distance and geographic distance among populations ($r=0.00$, $P>0.1$) and a strong positive correlation between genetic distance and habitat type ($r=0.72$, $P<0.01$). When controlling for geographic distance in partial Mantel analysis, genetic distance remained positively associated with habitat type ($r=0.72$, $P< 0.01$). When comparing only brown-sand and clay soil populations in the analysis, the Mantel test still showed no significant correlation between genetic distance and geographic distance between populations ($r=0.14$, $P>0.1$) and a positive correlation between genetic distance and habitat type ($r=0.60$, $P<0.01$). When controlling for geographic distance in the partial Mantel test, genetic distance remained positively associated with habitat type ($r=0.65$, $P=0.1$). Mantel tests comparing only populations found on the same soil type showed strong correlations between genetic distance and geographic distance (clay pops. $r=0.84$, $P=0.33$; white-sand pops. $r=0.98$, $P=0.16$). The insignificant p-values are most likely being driven by the small number of comparisons (n=3, df=2 for white-sand and clay pops.).

Selection vs. Drift

With the exception of PC2 in the white-sand and clay comparison all comparisons of phenotypic differentiation and neutral genetic variation (P_{ST}-F_{ST}) among soil types demonstrated much higher P_{ST} values compared to F_{ST} values supporting a hypothesis of divergent natural selection for all comparisons except for PC2 for comparison between white-sand and clay populations (Fig. 2.5).

Hybridization and Introgression

Results from the analysis of hybridization and introgression using NewHybrid corresponded with the Structure results and demonstrated low levels of introgression among populations found on different soil types. In the analysis of populations WS-A and BS-A NewHybrid assigned 52 individuals as pure parental white-sand individuals and 42 individuals as pure parental non-white-sand individuals. One individual could not be classified with confidence as white-sand parental or F_2 hybrid (PP = 0.80 and PP = 0.20 respectively). The four individuals, identified as admixed in Structure, were assigned with confidence as F_2 hybrids (PP > 0.9). No individuals were identified as F_1 hybrids.

Analysis between population pair Clay-B and BS-B also corroborated our structure results where all individuals were identified as parental types with the exception of the one admixed individual. Newhybrid confidently assigned this individual a having a hybrid origin however, was unable to distinguish it as an F1 hybrid or F2 hybrid (PP = 0.75 and PP = 0.13).

Population demographics

Exclusion of locus Prot100 did not substantially change migration estimates thus we have reported results from our analysis using the complete data set. Overall, inferred historical migration rates were consistently higher among populations found on the same soil type than among populations found on different soil types even when those populations were over 100km apart (Table 2.5, Fig. 2.6a, 2.6b). We found that migration rates were higher between brown sand and clay populations than among populations found on either soil type and white-sand populations. Furthermore, we found higher migration rates from white-sand populations into brown sand and clay populations then from either non-white-sand population into white-sand habitat (Table 2.5). Overall migration rates of less than one migrant per generation were observed when M values were averaged among populations found on different soil types and a rate of greater than one migrant per generation was observed when M values were averaged among populations found on the same soil type (Fig. 2.6b).

DISCUSSION

Our analysis of population level genetic and morphological differentiation identified three clearly diverged groups in *P. subserratum*, each associated with white-sand, brown-sand or clay soil habitats. Additionally, our results suggest that *P. subserratum* populations found on brown-sand and clay soils are experiencing a significant, and likely more recent, ecological divergence. We detected morphological differentiation among individuals found on white-sand, brown-sand and clay soils. Populations of *P. subserratum* across all three soil types were found to be more genetically similar to geographically distant populations found on the same soil type than to nearby populations found on different soil types. The average degree of phenotypic variation was much greater than the overall degree of neutral genetic differentiation for all comparisons among soil types suggesting that natural selection may play a more important role than drift in driving divergence among these populations. A small number of hybrid individuals were detected between brown-sand and white-sand groups and between clay and brown-sand groups suggesting that gene flow among populations on different soil types does occur at low frequency. Finally, migration rates were found to be higher between geographically distant populations found on the same soil type than they were between adjacent populations on different soil types. Taken together we observe a signature of fine scale ecological specialization across multiple soil boundaries. While we cannot discern whether initial divergence took place in allopatry or parapatry our results provide evidence consistent with the hypothesis that natural selection plays an important role in the maintaining diversity in paraptric populations within the *P. subserratum* species complex.

Population differentiation across habitat types.

We found high levels of genetic differentiation among populations found on all three soil types, even when those populations were directly adjacent to one another, presenting a clear pattern of 'isolation-by-adaptation'. Similar results across a wide range of taxa including passerines (Smith *et al.* 1997), anolis lizards (Ogden & Thorpe 2002), stickleback fish (Berner *et al.* 2009) and dolphins (Mendez *et al.* 2010) have been used to support the idea that natural selection is an important driver of diversification.

While Fine *et al.* (2013a) did not detect genetic or significant morphological differentiation between individuals found on brown-sand and clay soil, our results were consistent with their overall findings that populations associated with white-sand habitats are distinct from those found in non-white-sand habitats and that soil type as opposed to geographic distance is more important in the structuring of genetic variation between white-sand and non-white-sand populations. However, estimates of genetic diversity within populations strongly differed between the two studies. Fine *et al.* (2013a) reported higher haplotypic and genetic diversity in white-sand populations than clay and brown-sand populations. Here, we found the opposite pattern, with white-sand populations exhibiting lower levels of allelic diversity when compared to clay and brown-sand populations. Our measures of lower genetic diversity in white-sand populations may be expected given the small size and fragmented nature of white-sand habitat islands; higher levels of allelic diversity in brown-sand and clay populations accord well with the wide distribution of these more common soil types.

An explanation for the discrepancy in measurements of genetic diversity between these two studies could be that nuclear DNA sequences and nuclear microsatellite markers represent different temporal depths in the evolutionary history of these diverging groups. Microsatellite markers, with fast rates of evolution, represent more recent evolutionary events while nuclear DNA sequences, exhibiting lower rates of polymorphism, provide insights into evolutionary events that took place deeper in time. As seen in the Hawaiian silversword alliance, when demographic factors change over time this variation may be reflected in the genetic signatures of different molecular markers (Friar *et al.* 2007; Remington & Robichaux 2007 and discussed in Lawton-Rauh *et al.* 2007).

One possible demographic hypothesis that reconciles the results of these two studies and which may also shed light on the divergence histories of these populations relates to the geologic history of the Amazon. Prior to the Andean uplift, white-sand soils in the western Amazon were likely much more widespread than they are today, while brown-sand and clay soil types were less common (Hoorn 1993; Frazier *et al.* 2008). If white-sand specialist populations of *P. subserratum* were previously larger and brown-sand and clay populations were much smaller, the pattern of genetic diversity would correspond to those observed in nuclear sequence data. Microsatellite data, meanwhile, would reflect patterns of genetic diversity under the more recent history of the extent of these soil types. This hypothesis would also make sense with what appears to be a primary diversification from white-sand soils onto brown sand-soils with a secondary and more recent diversification from brown-sand soils onto clay soils, a pattern previously undetected by sequence data by Fine *et al.* (2013).

Hybridization and introgression

While low levels of interspecific gene flow was observed between edaphically differentiated populations, we found no evidence for substantial admixture between any edaphically differentiated population pairs. Low levels of introgression were observed between population pairs of *P. subserratum* found on white-sand and brown-sand soils and brown-sand and clay soils. This pattern is consistent with that of a bimodal hybrid zone which is characterized by parental forms that retain their genetic integrity in spite of gene flow and hybrid zones that are composed of F_2 and backcross hybrids but very few intermediate F_1 hybrids. Bimodal hybrid zones are often attributed to the presence of pre-mating barriers to reproduction in the form of assortative mating (Jiggins & Mallet 2000); they have been particularly well described in *Heliconius* butterflies (Arias *et al.* 2008), ground crickets (Howard *et al.* 1998) and sunflowers (Rieseberg *et al.* 1999). Alternatively, bimodal hybrid zones can also be maintained

by particularly strong natural selection if F_1 hybrids fail to survive and reproduce in alternative habitats (Arnold & Bennett 1993; Ross & Harrison 2002).

Edaphic heterogeneity and reproductive isolation

Edaphic differences play an important role in driving the development of both pre and post-zygotic barriers to reproduction among adaptively differentiated plant populations. Changes in flowering phenology are associated with soil preference in two parapatric species of palms on Lord Howe Island (Savolainen *et al.* 2006). High rates of seed abortion were demonstrated between copper tolerant and intolerant populations of *Mimulus guttatus* (Searcy & Macnair 1990) and strong natural selection against migrants was demonstrated between species of *Lasthenia* growing on and off serpentine soils (Yost *et al.* 2012).

Pre-zygotic barriers to reproduction may be present in soil specialist populations of *P. subserratum*. While flowering times do not differ among populations found on different soil types (PVA Fine unpublished data) edaphic differences may still play an indirect role in limiting pollen movement among populations. Initial pollinator surveys suggest that bee communities found in nearby brown-sand and white-sand forest types significantly differ from one another (Misiewicz 2014). If insect pollinators exhibit habitat preferences they may indirectly decreasing gene flow across habitat boundaries.

As was the case for population differentiation across habitat types, migration rates were higher between distant populations occurring on the same soil type than between adjacent populations found on different soil types. Particularly noteworthy is the result that when migration rates were averaged across pairs of edaphically differentiated populations the values are less than one, the rate by which theory predicts that populations will be strongly differentiated, while migration rates averaged across population pairs found on the same soil type are greater than one, the point where migration is predicted to overcome population level differentiation (Fig. 2.6b; Zhang *et al.* 2011). However, these results should be interpreted with caution. Demographic models implemented in Migrate do not take into account the possibility of retained ancestral polymorphism and assume that all genetic similarity among populations is the result of gene flow. While all migration rates among differing soil types were low we found that rates of migration from white-sand populations into clay and brown sand populations and from brown-sand populations into clay populations were higher than in the reverse direction. Patterns of asymmetric gene flow among edaphically differentiated populations could be a reflection of earlier divergence if ancestral alleles are retained in the more recently diverged populations. Alternatively, the same pattern could be the result of asymmetric strength of reproductive barriers.

Edaphically driven post-zygotic barriers to reproduction may also play a role in reducing successful gene flow among soil specialist populations of *P. subserratum*. White-sand, brown-sand, and clay soil habitats represent a mosaic of microclimatic differences, and the correlated variation in *P. subserratum* vegetative morphology with soil habitat type may reflect adaption to different selective regimes. Moreover, herbivore communities differ markedly between non-white-sand and white-sand habitats, as does the defensive chemistry *P. subserratum* individuals associated with these habitat types (Fine *et al*. 2013b). In the genus *Protium*, tradeoffs in growth rate and defense investment have been found in closely related sister species specialized on different soil types (Fine *et al*. 2006). Physiological tradeoffs between growth and defense in soil specialist populations of *P. subserratum* could lead to strong natural selection against phenotypic intermediates in both soil types, even when the potential for gene flow among habitats is high.

Divergence by drift

Our results suggest that natural selection across heterogeneous soil types plays an important role in driving and maintaining diversity in soil specialist populations of *P. subserratum*. However, high levels of neutral genetic differentiation among habitats and low levels of gene flow are consistent with the late stages of speciation and soil specialist populations could represent cryptic species, which diverged in allopatry and have since come into secondary contact. Furthermore, a hypothesis of initial divergence by genetic drift cannot be excluded. While specialization onto different soil types has clearly occurred multiple times within the species, we cannot disentangle the initial drivers of reproductive isolation without a more complete understanding of the phylogeographic history of the species. Nevertheless, it is worth noting that there are dozens of tree species that are endemic to white-sand forests in Peru with congeners that are associated with parapatric clay and/or brown sand habitats (Fine et al. 2010). Some of these species belong to extremely recently derived groups like *Inga* (Fabaceae) (Richardson et al. 2001), while others are thought to be much older (*Mauritia*, Arecaceae) (Couvreur et al. 2011). Though it cannot be ruled out, a hypothesis of allopatric divergence by drift with subsequent edaphic specialization followed by migration to the current parapatric distributions for so many tropical tree lineages seems less parsimonious. In either case, we believe that edaphic specialization appears to be a general mechanism that promotes and maintains Amazonian tree diversity.

Table 2.1. Populations sampled, region found, soil type, UTM coordinates for each population, Number of individuals sampled (N), observed heterozygotisty (H_O), expected heterozygosity (He), unbiased expected heterozygosity (UHe), average number of alleles (A), rarefied average number of private alleles (A_P), Rarefied average number of alleles (A_R), and inbreeding coefficient (F_{IS}).

Population	Site	Soil Type	Latitude	Longitude	N	Ho	He	UHe	A	A_P	A_R	F_{IS}
Clay-A	1	Clay	-3.830323320	-73.594676100	15	0.05	0.60	0.58	5.3	0.43	3.27	0.08
WS-A	2	White-sand	-3.917752825	-73.551974793	54	0.53	0.56	0.48	5.9	0.14	2.71	0.16
BS-A	2	Brown-sand	-3.909758490	-73.552303691	45	0.44	0.48	0.57	6.0	0.23	3.18	0.01
WS-B	3	White-sand	-3.950526640	-73.408218868	19	0.41	0.43	0.44	3.2	0.13	2.39	0.05
BS-B	3	Brown-sand	-3.976726271	-73.427493203	29	0.51	0.51	0.53	5.2	0.24	2.96	0.05
Clay-B	4	Clay	-4.058777584	-73.432082104	13	0.66	0.65	0.61	5.6	0.46	3.59	0.01
WS-C	5	White-sand	-4.864078785	-73.615967787	21	0.31	0.35	0.36	2.8	0.26	2.03	0.10
Clay-C	5	Clay	-4.887118050	-73.649012166	5	0.61	0.61	0.67	4.5	0.69	4.12	0.11

Table 2.2. Analysis of molecular variance (AMOVA) for populations of *Protium subserratum* grouped by geography (Sites 1,2,3,4, and 5), by soil type (white-sand, brown-sand and clay soil) and by soil type excluding white-sand (brown sand and clay soil).

	Soil Type (white-sand, brown sand, clay)			
AMOVA	d.f.	Sum of squares	Variance of components	% explained
Among groups	2	539.66	1.07*	25.74
Among populations, within groups	5	96.62	0.31*	7.34
Within populations	722	2016.23	2.79*	66.91
	Soil Type (brown-sand and clay)			
Among groups	1	115.13	0.83	19.93
Among populations, within groups	3	53.04	0.32*	7.68
Within populations	377	1138.52	3.02*	72.39
	Geographic Location			
Among groups	4	173.91	-0.36	-9.45
Among populations, within groups	3	462.36	1.39*	36.38
Within populations	722	2016.23	2.79*	73.06

*P<0.01

Table 2.3. Pairwise F_{ST} values for all population pairs. Values below the diagonal are estimated without using corrections for null alleles. Values above the diagonal are estimated using corrections for null alleles.

Pop.	Clay-A	WS-A	BS-A	WS-B	BS-B	Clay-B	WS-C	Clay-C
Clay-A	0	0.315	0.266	0.336	0.282	0.150	0.398	0.159
WS-A	0.334	0	0.340	0.037	0.360	0.295	0.187	0.279
BS-A	0.277	0.358	0	0.367	0.022	0.234	0.411	0.233
WS-B	0.341	0.035	0.377	0	0.389	0.307	0.176	0.298
BS-B	0.295	0.379	0.024	0.400	0	0.249	0.445	0.245
Clay-B	0.155	0.313	0.242	0.317	0.258	0	0.339	0.107
WS-C	0.408	0.186	0.426	0.173	0.460	0.353	0	0.377
Clay-C	0.151	0.290	0.238	0.296	0.250	0.107	0.382	0

Table 2.4. Maximum likelihood (ML) migration (M) estimates and their 95% confidence intervals for all population pairs.

Source Population	Recipient Population	M MLE	M 5%	M 95%	Source Population	Recipient Population	M MLE	M 5%	M 95%
Clay-A	Clay-B	0.58	0.47	0.72	Clay-A	WS-B	0.35	0.26	0.46
Clay-C	Clay-B	1.57	1.32	1.81	Clay-B	WS-B	0.07	0.04	0.13
WS-A	Clay-B	0.22	0.17	0.29	Clay-C	WS-B	0.07	0.04	0.12
WS-B	Clay-B	0.24	0.19	0.31	WS-C	WS-B	3.73	3.31	4.20
WS-C	Clay-B	0.43	0.35	0.52	WS-A	WS-B	5.37	4.71	6.10
BS-A	Clay-B	0.38	0.31	0.46	BS-A	WS-B	0.16	0.11	0.23
BS-B	Clay-B	1.20	1.05	1.36	BS-B	WS-B	0.07	0.04	0.11
Clay-B	Clay-A	1.44	1.22	1.68	Clay-A	WS-C	0.95	0.74	1.19
Clay-C	Clay-A	3.30	2.92	3.71	Clay-B	WS-C	0.40	0.28	0.54
WS-A	Clay-A	0.17	0.12	0.23	Clay-C	WS-C	0.14	0.08	0.23
WS-B	Clay-A	0.80	0.67	0.95	WS-A	WS-C	1.90	1.63	2.21
WS-C	Clay-A	0.38	0.31	0.46	WS-B	WS-C	1.81	1.44	2.28
BS-A	Clay-A	1.7	1.49	1.94	BS-A	WS-C	0.06	0.03	0.11
BS-B	Clay-A	0.74	0.62	0.88	BS-B	WS-C	0.08	0.04	0.15
Clay-A	Clay-C	1.36	1.80	1.57	Clay-A	BS-A	1.27	1.09	1.45
Clay-B	Clay-C	2.30	2.03	2.57	Clay-B	BS-A	0.68	0.56	0.84
WS-A	Clay-C	0.26	0.20	0.34	Clay-C	BS-A	1.97	1.74	2.25
WS-B	Clay-C	0.29	0.23	0.36	WS-A	BS-A	0.73	0.61	0.88
WS-C	Clay-C	0.24	0.16	0.31	WS-B	BS-A	0.55	0.45	0.67
BS-A	Clay-C	0.68	0.59	0.79	WS-C	BS-A	0.26	0.19	0.34
BS-B	Clay-C	0.43	0.35	0.53	BS-B	BS-A	5.12	4.74	5.51
Clay-A	WS-A	0.12	0.08	0.17	Clay-A	BS-B	0.42	0.33	0.53
Clay-B	WS-A	0.06	0.04	0.10	Clay-B	BS-B	1.20	1.03	1.40
Clay-C	WS-A	0.57	0.43	0.74	Clay-C	BS-B	0.99	0.80	1.21
WS-B	WS-A	3.98	3.63	4.38	WS-A	BS-B	0.46	0.37	0.57
WS-C	WS-A	3.19	2.92	3.48	WS-B	BS-B	0.47	0.39	0.57
BS-A	WS-A	0.13	0.09	0.18	WS-C	BS-B	0.28	0.19	0.39
BS-B	WS-A	0.18	0.12	0.25	BS-A	BS-B	3.65	3.30	4.00

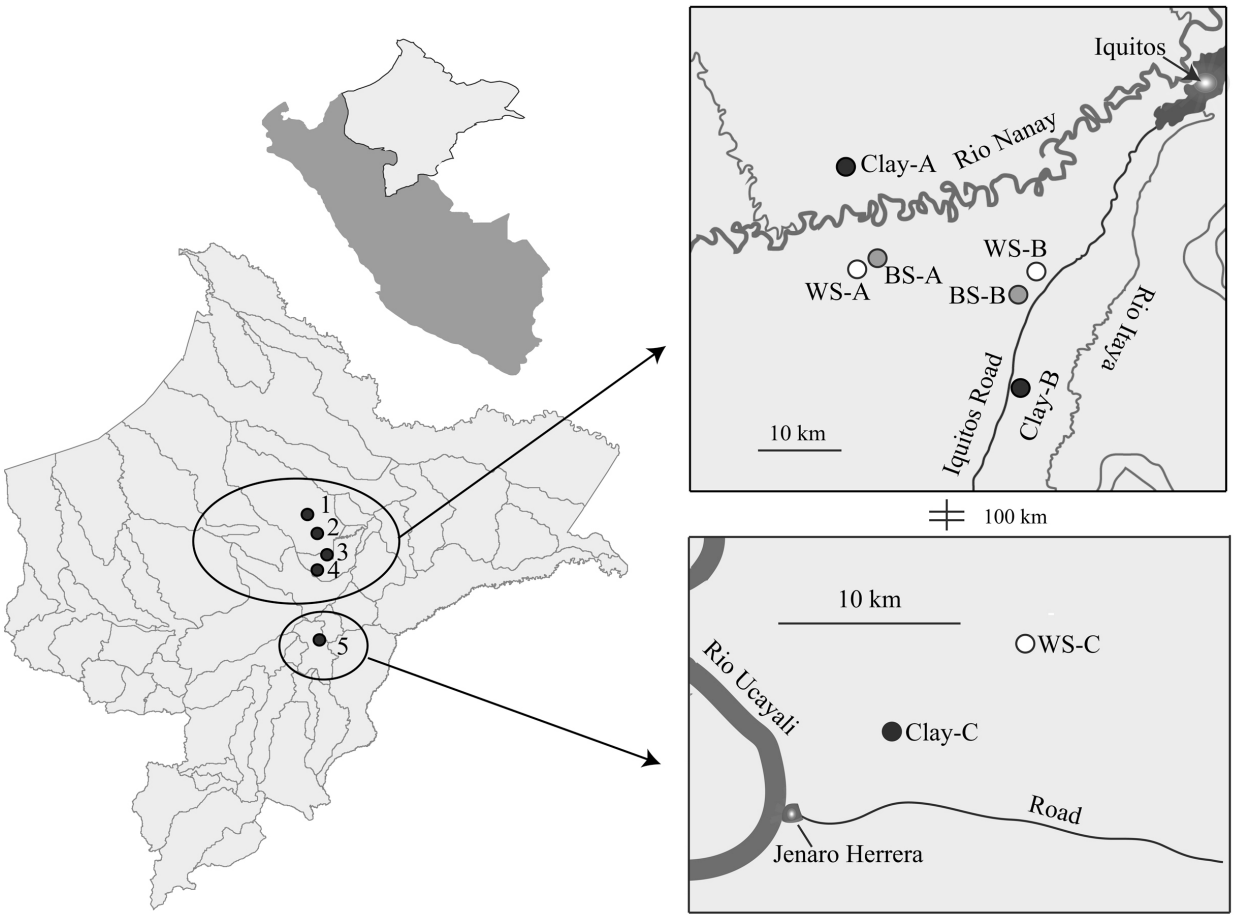

Figure 2.1. Sample sites and soil types for populations of *P. subserratum* in the region of Loreto, Peru. Numbered points represent the five sites where populations were found. Each individual population is displayed in the inset. Black circles represent populations found on clay soil, grey circles represent populations found on brown-sand soil and white circles represent populations found on white-sand soils.

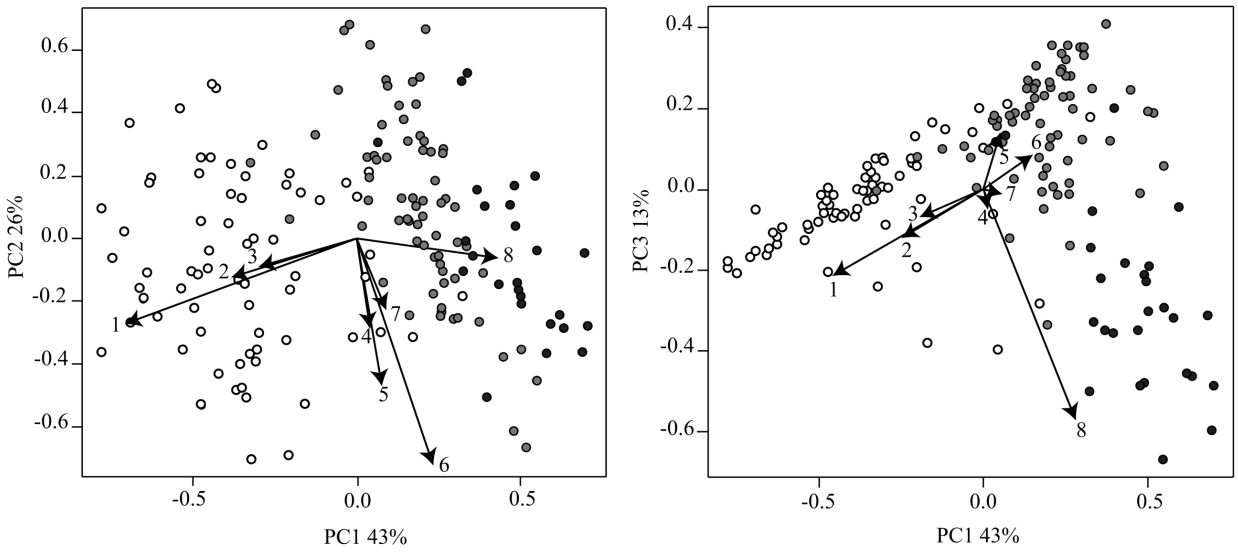

Figure 2.2. Principal components analysis (PCA) of leaf morphological characters for *P. subserratum* individuals. White, grey and black circles represent individuals found in white-sand, brown-sand and clay soil habitats respectively. Axis one through eight respectively represent the following characters; percent pubescence coverage on leaflet midrib, percent pubescence coverage on abaxial side of leaflet blade, leaflet thickness, leaflet width, number of leaflets per leaf, leaf length, leaflet length, and number margin serrations per leaflet.

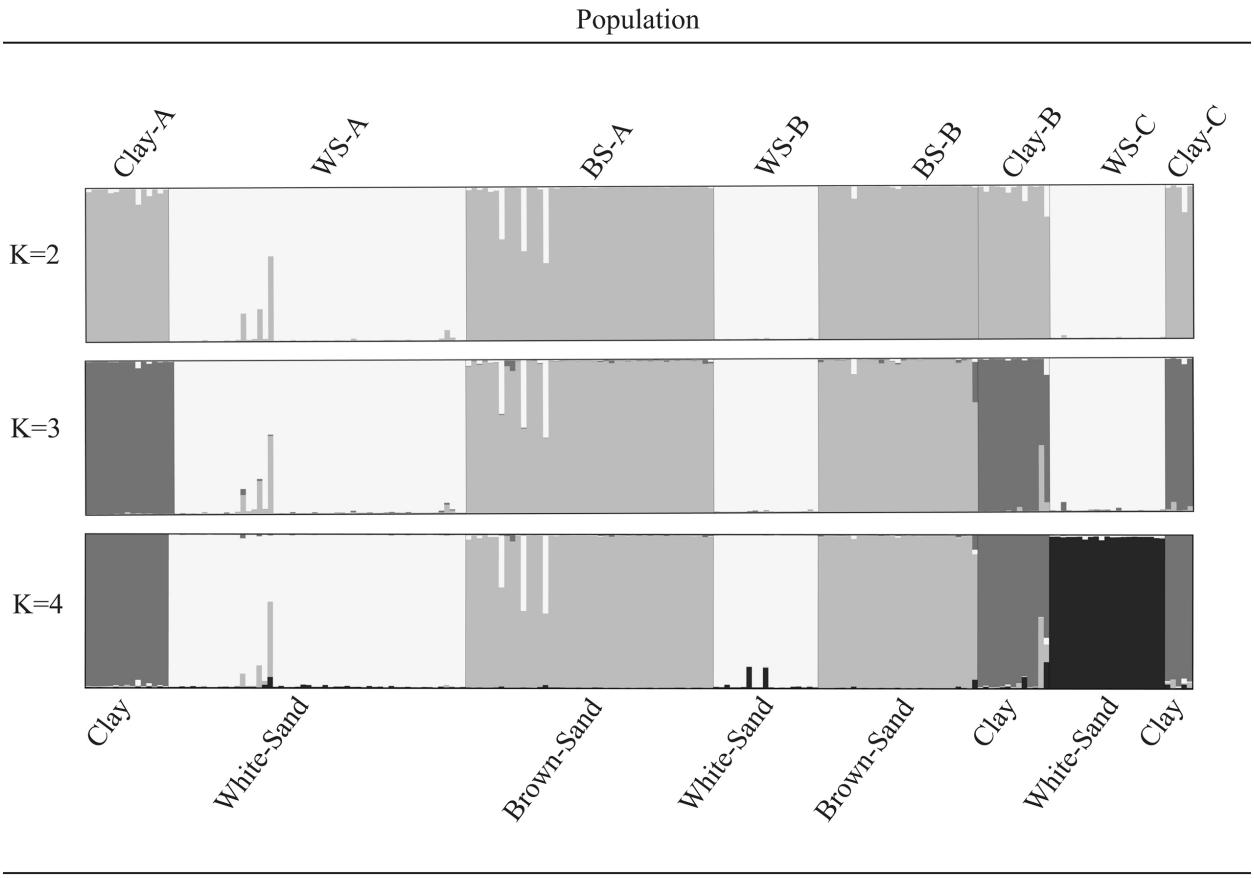

Figure 2.3. Evolutionary clusters (K=2-4) inferred from STRUCTURE analysis of 201 *P. subserratum* individuals from populations of white-sand, brown-sand and clay soil habitats. Each color represents an inferred character and each individual is represented by a vertical line shaded according to its probability of assignment to a given population.

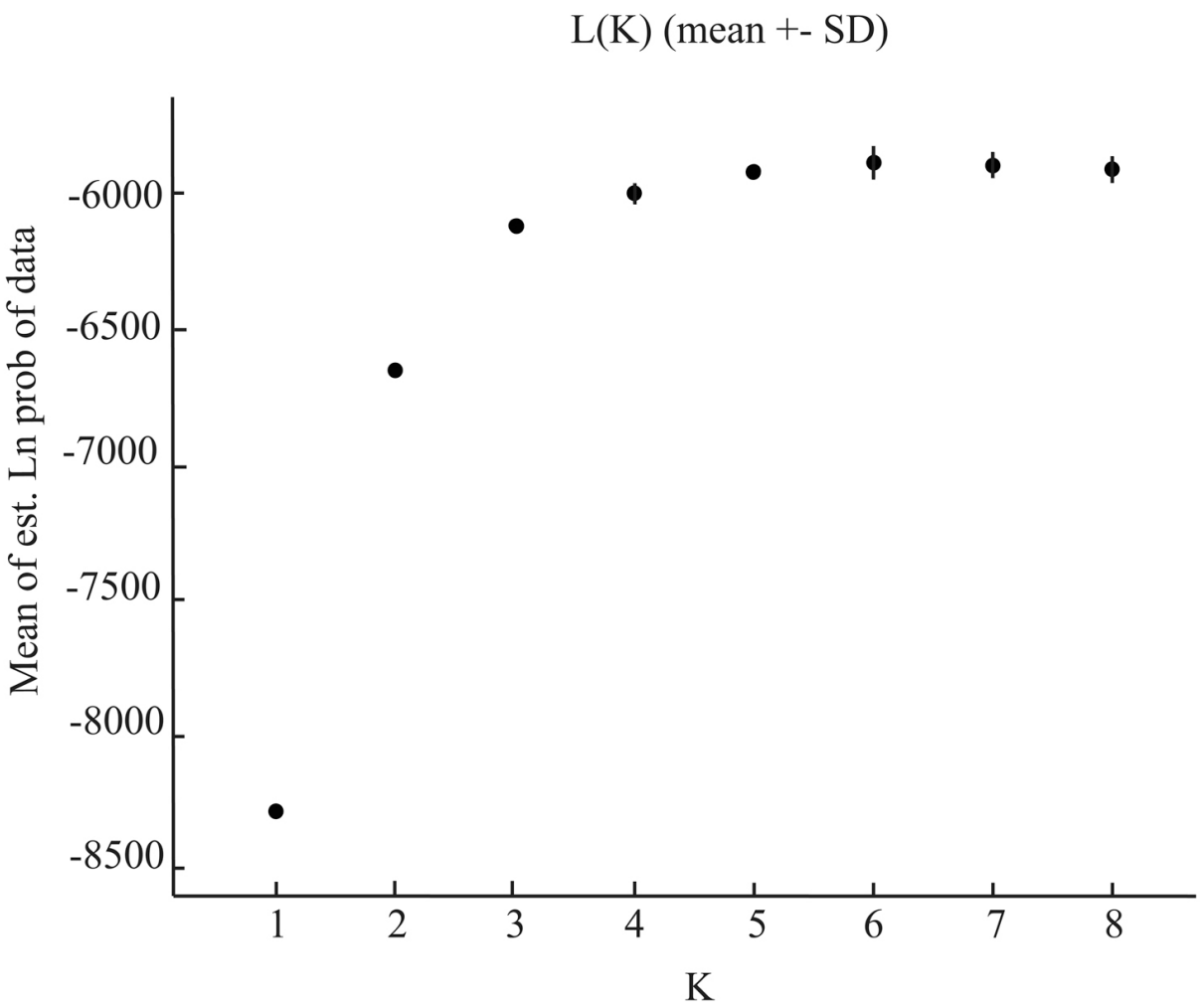

Figure 2.4. Average log probabilities for K = 1 – 8.

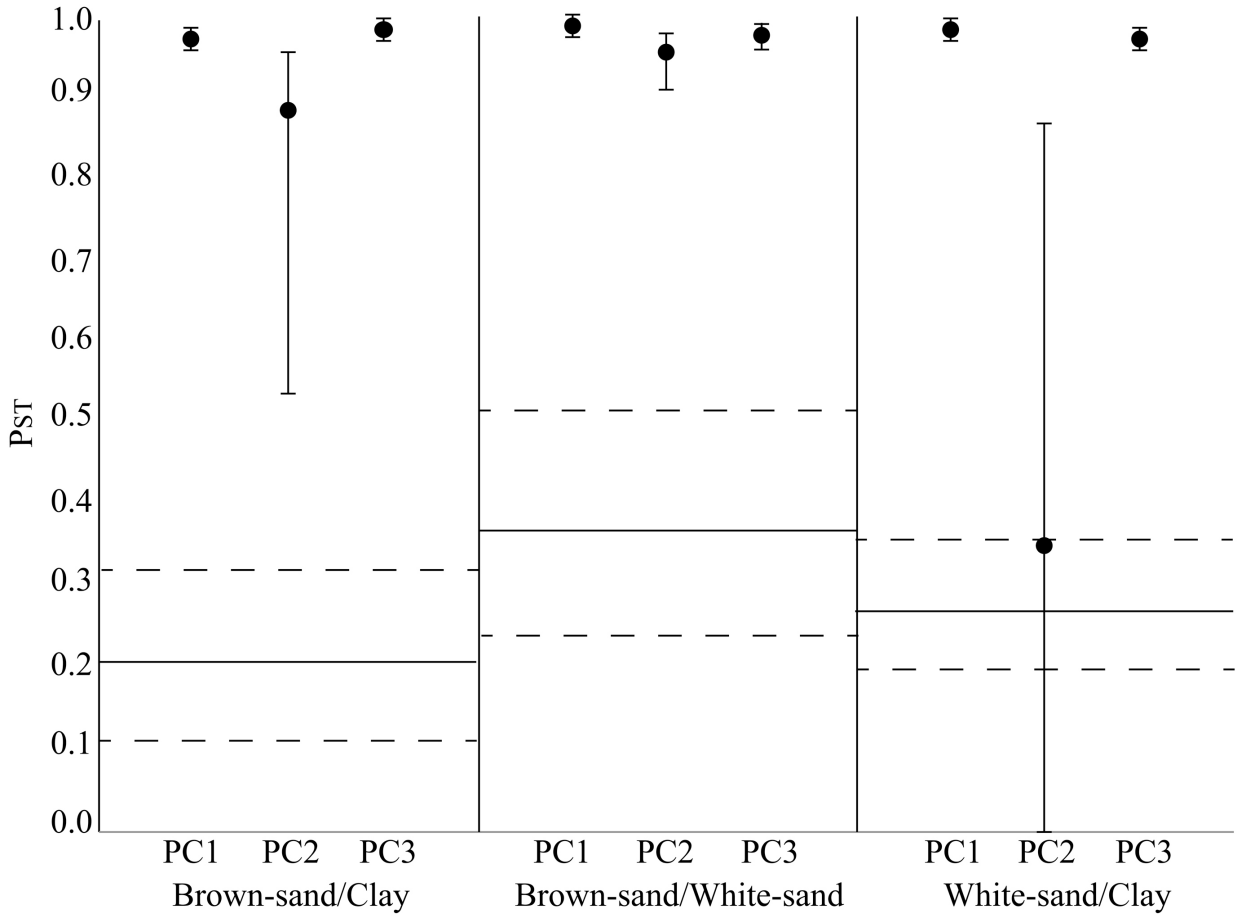

Figure 2.5. Comparison of P_{ST} and F_{ST} values and their 95% confidence intervals. P_{ST} values were calculated from principal components for individuals found on clay, white-sand and brown-sand soil types. Solid horizontal lines represent the multilocus F_{ST} value and dotted lines represent the 95% CI of the F_{ST} value.

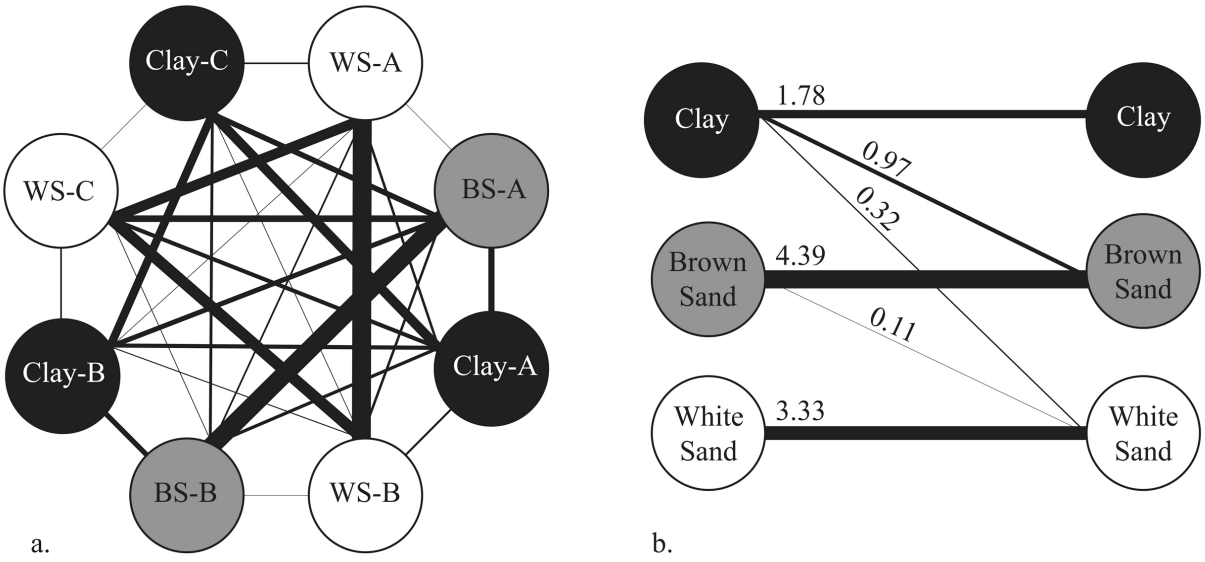

Figure 2.6. Gene flow (as migrants per generation, $M = \theta m / \mu$) estimated in Migrate-N. **a**. ML Migration rates between population pairs were averaged for one overall estimate to which lines between population pairs are proportional. **b**. ML estimates of migration across all soil types were averaged. Lines are proportional to the average migration rate between populations found on each soil type.

WORK CITED

Álvarez Alonso J, Metz MR, Fine PVA. (2013) Habitat specialization by birds in western Amazonian white-sand forests. *Biotropica,* **45**, 365-372.

Anderson AB (1981) White-sand vegetation of Brazilian Amazonia. *Biotropica*, **13**, 199-210.

Anderson EC, Thompson EA (2002) A model-based method for identifying species hybrids using multilocus genetic data. *Genetics*, **160**, 1217-1229.

Anderson EC (2002) User's guide to the program NewHybrids version 1.0.

Anacker B, Whittal J, Goldberg E, Harrison S (2011) Origins and consequences of serpentine endemism in the California flora. *Evolution,* **65**, 365-376.

Antonovics J, Bradshaw AD (1970) Evolution in closely adjacent plant populations VIII. Clinal patterns at a mine boundary. *Heredity*, **25**, 349-362.

Arias CF, Munoz AG, Jiggins CD, Mavarez J, Bermingham E, Linares M (2008) A hybrid zone provides evidence for incipient ecological speciation in *Heliconius* butterflies. *Molecular Ecology*, **17**, 4699-4712.

Arnold ML, Bennett BD (1993) Natural hybridization in Louisiana irises: genetic variation and ecological determinants. Pp. 115-139 in R.G. Harrison, ed. Hybrid zones and the evolutionary process. Oxford University Press, New York.

Beerli P (1998) Estimation of migration rates and population sizes in geographically structured populations. Pp. 39-53 in G. Carvalho, ed. in Advances in molecular ecology. NATO-ASI workshop series. IOS Press, Amsterdam.

Beerli P, Felsenstein J (1999) Maximum-likelihood estimation of migration rates and effective population numbers in two populations using a coalescent approach. *Genetics*, **152**, 763-773.

Beerli P, Felsenstein J (2001) Maximum-likelihood estimation of migration matrix and effective population sizes in n subpopulations by using a coalescent approach. *Proceedings of the National Academy of Sciences*, **98**, 4563-4568.

Berner DAC, Grandchamp, Hendry AP (2009) Variable progress toward ecological speciation in parapatry:stickleback across eight lake-stream transitions. *Evolution*. **63**, 1740-1753.

Chapuis MP, Estoup A (2007) Microsatellite null alleles and estimation of population differentiation. *Molecular Biology and Evolution*, **24**, 621-631.

Couvreur TLP, Forest F, Baker WJ (2011) Origin and global diversification patterns of tropical forests: inferences from a complete genus-level phylogeny of palms. *BMC Biology*, **9**:44.

Daly DC (1987) A taxonomic revision of *Protium* (Burseraceae) in eastern Amazonia and the Guianas. Ph.D. diss., City University of New York, New York.

Daly DC, Fine PVA (2011) An ecologically significant new section of *Protium* (Burseraceae) from Amazonia. Studies in neotropical Burseraceae XVI. *Systematic Botany*, **36**, 939-949.

Earl DA, vonHoldt BM (2012) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources*, **4**, 359-361.

Endler JA (1977) Geographic variation, clines, and speciation. Princeton University Press, Princeton, N.J.

Ehrlich PR, Raven PH (1969) Differentiation of populations. *Science*, **165**, 1228-1232.

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology,* **14**, 2611-2620.

Excoffier L, Smouse PE, Quattro, JM (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics,* **131**, 479-491.

Fine PVA, Mesones I, Coley PD (2004) Herbivores promote habitat specialization by trees in Amazonian forests. *Science*, **305**, 663-665.

Fine PVA, Daly DC, Munoz GV, Mesones,I, Cameron KM (2005) The contribution of edaphic heterogeneity to the evolution and diversity of Burseraceae trees in the western Amazon. *Evolution* **59**, 1464-1478.

Fine PVA, Miller ZJ, Mesones I, Irazuzta S, Appel HM, Stevens HH, Saaksjarvi I, Schultz JC, Coley PD (2006) The growth-defense trade-off and habitat specialization in plants in Amazonian forests. *Ecology*, **87**, S150-S162.

Fine PVA, Garcia-Villacorta R, Pitman NCA, Mesones I, Kembel SW (2010). A floristic study of the white-sand forests of Peru. *Annals of the Missouri Botanical Garden*, **97**, 283-305.

Fine PVA, Zapata F, Daly DC, Mesones I, Misiewicz TM, Cooper HF (2013a) Phylogeography of edaphic specialist and generalist species of *Protium* (Burseraceae): the relative importance of geographic distance and environmental heterogeneity in generating phylogeographic structure across the Amazon basin. *Journal of Biogeography*, **40**, 646-661.

Fine PVA, Metz MR, Lokvam J, Mesones I, Ayarza Zuniga JM, Lamarre M, Vasquez Pilco, Baraloto C (2013b) Insect herbivores, chemical innovation and the evolution of habitat specialization in Amazonian trees. *Ecology*, **94**, 1764-1775.

Frasier CL, Albert VA, Struwe L (2008) Amazonian lowland, white-sand areas as ancestral regions for South American biodiversity: Biogeographic and phylogenetic patterns in *Potalia* (Angiospermae: Gentianaceae). *Organisms Diversity & Evolution*, **8**, 44-57.

Friar EA, Cruse-Sanders JM, and McGlaughlin ME (2007) Gene flow in *Dubautia arborea* and *D. cilolata*: the roles of ecology and isolation by distance in maintaining species boundaries despite ongoing gene hybridization. *Molecular Ecology*, **16**, 4028-4038.

Gentry AH (1988) Patterns of plant community diversity and floristic composition on environmental and geographical gradients. *Annals of the Missouri Botanical Garden*, **75**, 1-34.

Goodman SJ, Barton NH, Swanson G, Abernathy K, Pemberton JM (1999) Introgression through rare hybridization: a genetic study of a hybrid zone between red and sika deer in Argyll, Scotland. *Genetics*, **152**, 355-371.

Grahame JW, Wilding CS, Butlin RK (2006) Adaptation to a steep environmental gradient and an associated barrier to gene exchange in *Littorina saxatilis*. *Evolution*, **60**, 268-278.

Haffer J (2008) Hypotheses to explain the origin of species in Amazonia. *Brazilian Journal of Biology*, **68**, 917-947.

Haldane JBS (1948) The theory of a cline. *Journal of Genetics*, **48**, 277-284

Hendry AP, Day T, Taylor EB (2001) Population mixing and the adaptive divergence of quantitative traits in discrete populations: a theoretical framework for empirical tests. *Evolution*, **55**, 459-466.

Hendry AP, Taylor EB, McPhail JD (2002) Adaptive divergence and the balance between selection and gene flow: lake and stream stickleback in the Misty system. *Evolution*, **55**, 1199-1216.

Hoorn C (1993) Marine incursions and the influence of Andean tectonics on the Miocene depositional history of northwestern Amazonia: results of a palynostratigraphic study. *Paleogeography Paleoclimatology Paleoecology*, **112**,187-238.

Hoorn C (1994) An environmental reconstruction of the paleo-Amazon River system (Middle-Late Miocene, NW Amazonia). *Paleogeography, Paleoclimatology, Paleoecology*, **112**, 187-238.

Howard DJ, Gregory PG, Chu J, Cain ML (1998) Conspecific sperm precedence is an effective barrier to hybridization between closely related species. *Evolution*, **52**, 511-516.

Jakobsson M, Rosenberg NA (2007) CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics*, **23**, 1801-1806.

Jensen JL, Bohonak AJ, Kelley ST (2005) Isolation by distance, web service. *BMC Genetics*, **6**, 6-13.

Jiggins CD, Mallet J (2000) Bimodal hybrid zones and speciation. *Trends in Ecology and Evolution*, **15**, 250-255.

Latta RG (2004) Gene flow, adaptive population divergence and comparative population structure across loci. *New Phytologist*, **161**, 51-58.

Lawton-Rauh A, Friar EA, Remington DL (2007) Coordinated temporal patterns of differentiation among molecular markers in recently-derived species of the Hawaiian silversword alliance adaptive radiation (Heliantheae, Asteraceae). *Molecular Ecology*, **16**, 3993-3994.

Leinonen T, Cano JM, Makinen H, Merilä J (2006) Contrasting patterns of body shape and neutral genetic divergence in marine and lake populations of threespine sticklebacks. *Journal of Evolutionary Biology*, **19**, 1803-1812.

Mendez M, Rosenbaum HC, Subramaniam A, Yackulic C, Bordino P (2010) Isolation by environmental distance in mobile marine species: molecular ecology of franciscana dolphins at their southern range. *Molecular Ecology*, **19**, 2212-2228.

Merilä J, Crnokrak P (2001) Comparison of genetic differentiation at marker loci and quantitative traits. *Journal of Evolutionary Biology*, **14**, 892-903.

Mayr E (1963) Animal species and evolution. Cambridge, MA: Harvard University Press.

Medina E, Cuevas E (1989) Patterns of nutrient accumulation and release in Amazonian forests of the upper Rio Negro basin. Pp. 217-240 in J. Proctor, ed. Mineral nutrients in tropical forest and savanna ecosystems. Blackwell Scientific, Oxford, U.K.

Misiewicz TM, Barbosa CEA and Fine PVA (2012) Microsatellite Primers for an Amazonian lowland tropical tree, *Protium subserratum* (Burseraceae). *AJB Primer Notes and Protocols in the Plant Sciences*. E465-E467.

Misiewicz TM (2014) Ecological divergence and reproductive isolation in an Amazonian tropical tree: *Protium subserratum* (Burseraceae). Ph.D. diss., University of California Berkeley, California.

Nosil P, Egan SP, Funk DJ (2007) Heterogeneous genomic differentiation between walking-stick ecotypes: "Isolation by adaptation" and multiple roles for divergent selection. *Evolution*, **62**, 316-336.

Nosil P, Funk DJ, Ortiz-Barrientos D (2009) Divergent selection and heterogeneous genomic divergence. *Molecular Ecology.* **18**, 375-402.

Nosil P (2012) Ecological speciation. Oxford University Press, Oxford.

Ogden R, Thorpe RS (2002) Molecular evidence for ecological speciation in tropical habitats. *Proceedings of the National Academy of the Sciences*, **99**, 13612-13615.

Peakall R, Smouse PE (2006) GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes*, **6**, 288-295.

Pregitzer CC, Bailey JK, Hart SC, Schweitzer JA (2010) Soils as agents of selection: feedbacks between plants and soils alter seedling survival and performance. *Evolutionary Ecology*, **24**, 1045-1059.

Pritchard JK, Stephens P, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics*, **155**, 945-959.

Raymond M, Rousset F (1995) GENEPOP (version 1.2): Population genetics software for exact tests and Ecumenicism. *Journal of Heredity*, **86**, 248-249.

Remington DL, Robichaux RH (2007) Influences of gene flow on adaptive speciation in the *Dubautia arborea--D. ciliolata* complex. *Molecular Ecology*, **16**, 4014-4027.

Rice WR (1989) Analyzing tables of statistical tests. *Evolution*, **43**, 223-225.

Richardson JE, Pennington TR, Pennington TD, Hollingsworth PM (2001) Rapid diversification of a species-rich genus of neotropical rainforest trees. *Science*, **293**, 2242-2245.

Rieseberg LH, Whitton J, Gardner K (1999) Hybrid zones and the genetic architecture of a barrier to gene flow between two sunflower species. *Genetics*, **152**, 713-727.

Rosenberg NA (2004) DISTRUCT: a program for the graphical display of population structure. *Molecular Ecology Notes*, **4**, 137-138.

Rosenthal DM, Ludwig F, Donovan LA (2005) Plant responses to an edaphic gradient across an active sand dune/desert boundary in the great basin desert. *International Journal of Plant Sciences*, **166**, 247-255.

Rundle HD, Nosil P (2005) Ecological speciation. *Ecology Letters*, **8**, 336-352.

Sääksjärvi IE, Haataja S, Neuvonen S, Gauld ID, Jussila R, Salo J, Burgos AM (2004) High local species richness of parasitic wasps (Hymenoptera: Ichneumonidae;Pimplinae and Rhyssinae) from the lowland rainforests of Peruvian Amazonia. *Ecological Entomology*, **29**, 735-743.

Savolainen V, Anstett MC, Lexer C, Hutton I, Clarkson JJ, Norup MV, Powell MP, Springate D, Salamin N, Baker WJ (2006) Sympatric speciation in palms on an oceanic island. *Science*, **441**, 210-213.

Searcy KB, Macnair MR (1990) Differential seed production in *Mimulus gattatus* in response to increasing concentrations of copper in the pistil by pollen from copper tolerant and sensitive sources. *Evolution*, **44**, 1424-1435.

Smith TB, Wayne RK, Giman DJ, Bruford MW (1997) A role for ecotones in generating rainforest biodiversity. *Science*, **276**, 1855-1857.

Smith SS, Schweitzer JA, Turk P, Bailey JK, Hart SC, Shuster SM, Whitham TG (2011) Soil-mediated local adaptation alters seedling survival and performance. *Plant and Soil*, **352**, 243-251.

Szpiech ZA, Jakobsson M, Rosenberg NA (2008) ADZE: a rarefaction approach for counting alleles private to combinations of populations. *Bioinformatics*, **24**, 2498-2504.

Thibert-Plante X, Hendry AP (2010) When can ecological speciation be detected with neutral loci? *Molecular Ecology*, **19**, 2301-2314.

Tuomisto H, Ruokolainen K, Linna A, Danjoy W, Rodriguez Z (1995) Dissecting Amazonian biodiversity. *Science*, **299**, 241-244.

van Oosterhout C, Hutchinson WF, Wills DPM, Shipley P (2004) Micro-checker: Software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes*, **4**, 535-538.

Weir BS, Cockerham CC (1984) Estimating F-statistics for the analysis of population structure. *Evolution*, **38**, 1358-1370.

Wright S (1931) Evolution in Mendelian populations. *Genetics*, **16**, 97-259.

Yost JM, Barry T, Kay KM, Rajakaruna N (2012). Edaphic adaptation maintains the coexistence of two cryptic species on serpentine soils. *American Journal of Botany*, **99**, 890-897.

Zhang C, Zhang D, Zhu T, Yang Z (2011) Evaluation of a bayesian coalescent method of species delimitation. *Systematic Biology*, **60**, 747-761.

CHAPTER 3: Distance and Habitat Drive Stingless Bee Community Turnover Across Naturally Heterogeneous Forests in the Western Amazon

ABSTRACT

High tree species richness in the western Amazon has been attributed to heterogeneous soils, which harbor edaphic specialist trees. While rapid transitions in tree communities are well documented across these variable soils few studies have investigated the role of habitat heterogeneity in structuring animal communities. Stingless bees are taxonomically diverse and important natural pollinators in Neotropical forests. However, little is known about their community structuring at local scales in naturally heterogeneous environments. We systematically sampled stingless bee communities found across three paired sites that included adjacent patches of white-sand and non-white-sand forest in the lowland Amazonian region of Loreto, Peru. We sought to understand: (1) How stingless bee species richness and abundance differs among white-sand and non-white-sand habitats and (2) The relative influence of fine scale geographic distance and habitat type in structuring stingless bee communities. We found that species richness did not differ between habitats and that species abundances were highest in white-sand habitats. Community analyses for sampling sites pooled across all months demonstrated that location and soil type played a significant role in structuring bee communities and that community turnover may be more strongly influenced by distance in white-sand habitats than non-white sand habitats. Our results suggest that distance and habitat play an important role in driving stingless bee community turnover at fine scales and that the interaction between habitat and geographic distance may promote higher stingless bee community turnover in white-sand habitats than non-white sand habitats.

INTRODUCTION

The lowland amazon basin is notable for its exceptionally high levels of species diversity across a wide range of taxa (Erwin 1988, Kress *et al.* 1998, Lamas *et al* 1991, Pitman *et al.* 2001). High levels of species richness have been attributed to extreme habitat heterogeneity with particular emphasis on the mosaic of soil types found across the Western Amazon in Peru (Gentry 1988, Terborgh 1985, Whitney & Alvarez 1989). These variable soils create a patchwork of forest types differentiated by plant communities that have strong edaphic associations with nutrient poor white-sand soil patches and the more fertile brown-sand and clay soils which surround them (Tuomisto & Ruokolainen 1994, Rukolainen *et al.* 1997, Fine *et al.* 2010). White-sand forests differ from surrounding forests in that they harbor significantly less species richness, are shorter in stature, and experience higher temperatures on average below the canopy due to increased light penetration than forests found on surrounding soils (Fine *et al*. 2010, Anderson 1981, Medina & Cuevas 1989). In the western Amazon white-sand forests exist as small habitat islands, usually no larger than a few square hectares (Fine *et al.* 2005).

While the majority of research examining the role of edaphic variation in species turnover in the Amazon has been centered on tree communities (Gentry 1988, Tuomisto & Ruokolainen 1994, Rukolainen *et al.* 1997, Fine *et al.* 2010 but see Alvarez Alonso *et al*. 2013) similar habitat specialization may also be present in animals. Because forests constitute the primary habitat and food source for many forest dwelling animals we expect that abrupt changes in floristic composition across habitats may in turn drive turnover in animal communities.

Animals that provide pollination services play a particularly important role in tropical ecosystems where the majority of trees are reliant on animal interactions for pollen transfer

(Bawa 1990). While turnover in bee communities across forest fragments and agriculturally modified landscapes has been well studied in relation to crop production (Perfecto *et al.* 1996, Tylianakis *et al.* 2005, Jha & Vandameer 2010) the role of naturally heterogeneous habitats in driving species turnover in the lowland Amazon has largely been neglected (but see Abrahamczyk *et al.* 2010).

Stingless bees (Hymenoptera: Apidae: Meliponini) are an important taxon for studies of biodiversity and species turnover in the lowland Amazon because they are highly diverse (*ca* 500 species) with their center of diversity found in the Neotropics (*ca* 400 species) (Michener 2013). Additionally, they are the most important native providers of pollination services in the Amazon, making them essential for ecosystem functioning (Roubik 1995, Engel & Dingemans-Bakels 1980).

Most stingless bee species are considered to be generalist pollinators and they exhibit a wide range of variation in nesting habits across species. Nests are usually arboreal or subterranean and are constructed using diverse construction materials including mud, wood pulp, feces, and plant exudates (Schwarz 1948, Roubik 1989). Foraging distance away from nest site is dependent on the size of the bee with distances ranging from less than 500 meters to 2 km (Kerr 1959, Araujo *et al.* 2004, Kuhn-Neto *et al.* 2009). Given the variation in nest site preferences between lineages relatively little attention has been paid to the fine scale distribution and ecology of Neotropical stingless bees and no studies have investigated species turnover across naturally occurring environmental gradients in undisturbed forest sites.

Furthermore, because the movement of animal pollinators directly influences the distance, direction and degree of pollen dispersal, they ultimately determine the spatial pattern of gene movement within and among plant populations (Dick *et al.* 2003, Garcia *et al.* 2007). If pollinators are restricted in their foraging area due to habitat preference (Dieckmann 2004) then the question of ecological specialization in bee communities may be of particular interest to plant ecologists as well.

In this study we simultaneously examined the effect of habitat and distance in structuring stingless bee communities at a local scale. We systematically sampled native bee communities found across three paired sites that included adjacent patches of white-sand and non-white-sand forest across more than 100km in the lowland Amazon in the region of Loreto, Peru in order to answer two questions (1) How does stingless bee species richness and abundance differ among white-sand and non-white-sand habitats? (2) What is the relative influence of fine scale geographic distance and habitat type in structuring stingless bee communities?

If stingless bees are generalist pollinators with relatively large foraging ranges, we expect that geographic distance will play a greater role in structuring bee communities than habitat type providing that trees exhibit similar flowering phenology across habitat types. Alternatively, If bees prefer floral resources provided by soil specialist trees, nesting sites that are more common in one particular habitat (*i.e.* large vs. small stems or clay vs. sandy soil in the case of subterranean nesters) or environmental differences such as temperature or predation risk then we may find that habitat type plays a stronger role than distance in structuring stingless bee communities.

METHODS

STUDY SITES.–Three primary study sites, each containing adjacent white-sand and non-white-sand forest patches, were established in the region of Loreto, Peru (Fig. 3.1; Table 3.1). Site one and two are located within the Allpahuayo Mishana National Reserve in the Nanay

River watershed and site three is located approximately 100km to the south in the Ucayali River watershed.

SAMPLING DESIGN.–All trapping was conducted using bee pan traps. These traps are easily standardized and avoid collector bias (Westphal *et al.* 2008). Traps were created using 12oz clear plastic soup bowls painted fluorescent blue, fluorescent yellow, or white in order to account for variation in color preference among bee species. Four trapping stations consisting of six bee pan traps were established in each white-sand and non-white-sand forest at each site for a total of 24 trapping stations across six forest patches (Fig. 3.1). Within forest patches each trapping station was established 200-250m distant from any other trapping station.

Each trapping station contained two sets of yellow, blue and white traps. One set was suspended one meter above the ground with each individual bowl spaced at a distance of five meters to avoid bowl competition (Droege 2010). The second set of bowls was suspended at a height of 15–20m in the canopy directly above the ground traps. Bowls were filled with six ounces of soapy water solution (one tsp blue Dawn brand soap per two liters of water).

Pan traps were set out at each site once per month for 24h March-July 2010. While Loreto, Peru exhibits very little seasonality our sampling period extended from the high water season, when rivers rise substantially, through the low water season. All trapped specimens were collected in the field and transferred to 96% ethanol. Specimens were separated, pinned and identified and have been deposited at the Essig Museum of Entomology at University of California, Berkeley. Identification of all specimens was done by C.R. by direct comparison with a large synoptic collection of Peruvian stingless bees previously identified by J.M.F. Camargo

GENERAL DIVERSITY.–We assessed the effectiveness of our sampling method using species accumulation curves and the Chao estimator (Chao 1987). Rarefaction curves were calculated using the individual-based species matrix and the species accumulation curve and Chao total richness estimates were calculated using the 'specacum' and 'specpool' functions respectively within the package 'vegan' (Oksanen *et al.* 2013). We assessed the dominance structure within our dataset by ranking the relative abundance of each species using a regular base plot (Magurran 2004). Collection data for all trap clusters in each sampling region were pooled to assess the importance of temporal heterogeneity, spatial distance and habitat type on species richness and abundance using multiple linear models to select the best combination of explanatory variables. Models were fitted using step-wise selection with AIC criterion, using function "stepAIC" in package MASS.

FINE SCALE COMMUNITY STRUCTURE.–We investigated the influence of trap height, trap color, month collected, soil type and trap site on species richness and abundance using an Analysis of Variance (ANOVA) followed by stepwise regression to determine which model best explains the data.

BROAD SCALE COMMUNITY STRUCTURE.–In order to assess the broad scale affects of soil type and location, we combined all monthly collection data for each of the four trap stations within each of the six sites for a total of 24 sample units. We transformed the data by first applying square root to the abundance data followed by standardization using the Wisconsin method (Bray & Curtis 1957), which reduces the effect of overly dominant species in the data set and controls for sampling effort in each trap cluster (Legendre & Gallagher 2001). We first quantified relative contributions of soil type and location to the community structure by performing variance partitioning with the function 'varpart'. The analysis partitions the explained variation in community structure into different components based on the studied environmental factors (Borcard *et al.* 1992). Then, in order to visualize results and specifically test the significance of

effects of soil and location, as well as their interaction in driving community structure, we used a distance-based Redundancy Analysis (db-RDA) with soil and location as constraints. This method allowed us to carry out constrained ordinations using non-Euclidean distance measures (Gower 1966, Gower 1985, Legendre & Anderson 1999, Legendre & Legendre 2012). Our distance matrix was created using Bray-Curtis distance, which only accounts for shared presences between two sites (Anderson *et al.* 2011), and the redundancy analysis was carried out using the 'capscale' function (Anderson & Willis 2003). The significance of constrained ordination was assessed using a permutation test for Constrained Correspondence Analysis (Legendre *et al.* 2011, Legendre & Legendre 2012) using the function 'anova.cca.' The *P*-value is calculated by comparing the observed *F*-value with the values from 999 permutations of community data.

We also used a Permutational Multivariate Analysis of Variance (PERMANOVA) to further test the effects of soil, location, and their interaction using the function 'adonis.' This analysis is analogous to parametric Multivariate Analysis of Variance (MANOVA), but has been shown to be more robust for community data, as the P-value is derived from permutation, as opposed to the comparison against a known distribution (Anderson 2001).

Finally, using the function 'mantel' we implemented Mantel tests (Legendre & Legendre 1998) individually on white-sand and non-white-sand populations to determine if geographic distance was more important in structuring bee communities in one habitat or the other. All functions for community analysis are available the R package 'vegan.' All statistical analyses were carried out in R 3.0.3 (R Development Core Team 2013).

RESULTS

GENERAL DIVERSITY.–We trapped a total of 1109 bees representing three families, 17 genera and 39 species. All but three taxa were Apidae. (Appendix S3.1). Thirty-one species (79%) were identified to the species level and the remaining eight were sorted to morphospecies. Eight of the collected specimens representing six species were identified as solitary bees and were therefore discarded from the dataset for further analysis. All other specimens for the remaining 33 species were stingless bees and were included in all analyses (Table 3.1).

The species accumulation curve approached, but did not reach an asymptote (Fig. 3.2) suggesting that our sampling was adequate but not exhaustive. The estimated Chao total for total species richness across all habitats and sites was 38.6 species (Standard Error=±3.85) meaning we captured approximately 77 percent to 95 percent of the estimated total number of stingless bee species. Overall, our richness estimates are slightly lower than was found in a study by Rasmussen and Gonzalez (2009) that detected 51 stingless bee species in a medium elevation tropical rainforest in Peru after intensive targeted sampling. *Plebia minima* and *Plebia* sp. A were particularly abundant in our data set (N=584, N=235) while 6 of the 33 species were represented by singletons. The remaining 29 species were represented by medium to low abundances (Fig. 3). A total of 19 species (N=1,048) were found in both white-sand and non-white-sand habitats, five species (N=13) were found only in non-white-sand habitats and nine species (N=40) were found only in white-sand habitats. Total abundance (A) and species richess (S) for each habitat type and sampling location are reported in Table 3.1.

FINE SCALE COMMUNITY STRUCTURE.–Soil type (F=10.62, P=0.001) and location (F=7.10, P<0.001) had a significant effect on richness, while it did not vary significantly with month collected, trap height and pan color (P=0.1, P=0.45, P=0.76, respectively). When testing for effects on abundance, soil type had the largest effect (F=5.2, P=0.02) followed by the month in

which collections were made (F=4.3, P=0.04). Location, trap height and pan color were insignificant (P=0.27, P=0.09, P=0.8). The multiple linear model that best explained species richness included soil type, location, and month collected (AIC=-86.64). The model that best explained species abundance included soil type, month collected, and height (AIC=905.66).
BROAD SCALE COMMUNITY STRUCTURE.–The db-RDA indicated that both soil type and geographic distance play a role in structuring bee communities (Fig. 3.4). The permutation test for this ordination yielded significant results with both location and soil type as constraints (F=2.35, P<0.001) with soil type contributing to 1.6 percent of the variation and location contributing to 11.15 percent for a total contribution of 13.47 percent of the variation. Results from the PERMANOVA corroborated these results with soil type (F=2.00, P=0.01) and location (F=3.32, P=0.001), as well as the interaction of the two (F=1.65, P=0.01) having a significant impact on the observed community structure. Mantel tests confirmed that geographic distance is correlated with species turnover in both habitat types and suggests that distance may play a stronger role in structuring white-sand populations than it does in non-white-sand populations (white-sand, r=0.4, P<0.001); non-white-sand, r=0.3, P=0.02).

DISCUSSION

To our knowledge this is the first study to examine the role of geographic distance and forest type in structuring stingless bee communities at local scales across a naturally heterogeneous landscape in the Western Amazon. We found that both location and habitat were important in structuring stingless bee communities even over extremely small spatial scales. Our results were consistent with other tropical bee studies, which demonstrate changes in bee communities across a variety of spatial scales and environmental gradients (Tylianakis *et al.* 2005, Abrahamczyk *et al.* 2011, Batista Matos *et al.* 2013).

Bee communities sampled in geographically distant white-sand patches were more dissimilar to one another than bee communities from non-white-sand patches were to each other at the same distances. White-sand forests exist as small patches or habitat 'islands' surrounded by a matrix of non-white-sand forest. Accordingly, factors such as migration, colonization and local extinction may play a stronger role in structuring white-sand bee communities than non-white-sand bee communities. If non-white-sand habitat is less favorable for bee species found in white-sand forests then these communities may experience higher levels of isolation due to the compounded effect of habitat and distance. In this case metacommunity dynamics could play an important role in increasing turnover among white-sand forests effectively amplifying the effect of geographic distance and habitat alone.

While our results demonstrated that stingless bee community structure is strongly influenced by location, habitat type also plays significant role particularly at very fine geographic scales. We found that variation across sampling sites was driven in part by soil type however, species specific to one forest type tended to be rare in our collections, making it difficult to discern between true habitat specificity and insufficient sampling. Species abundances were much higher in white-sand-forests than non-white-sand forests suggesting that while many stingless bee species utilize both forest types habitat preferences may dictate where they are more commonly found. Floral and nesting resources are both important in structuring bee communities across habitats (Tependino & Stanton 1981, Petanidou & Ellis 1996, Wuellner 1999). Fierro *et al.* (2012) found that stingless bee species show preferences for particular tree taxa, which commonly provide ideal nesting sites, as well as species-specific foraging behavior suggesting that turnover in tree diversity likely drives changes in stingless bee distributions.

While we did not quantify differences in floral or nesting resource availability between habitat types in this study marked differences in floristic composition, forest structure, microclimate and abiotic resources are likely driving differences in these neighboring bee communities.

Habitat based differences in stingless bee communities may also reinforce tropical tree specialization across habitat boundaries. Many tree species that are endemic to white-sand forest patches in Peru have congeners associated with parapatric non-white-sand forests (Fine *et al.* 2010) and divergent natural selection across adjacent white-sand and non-white-sand habitats has been shown to play an important role in maintaining boundaries between ecologically divergent tree populations (Misiewicz & Fine in press). If pollinators forage less frequently outside of their preferred habitat type they may indirectly limit pollen flow between ecologically divergent plant populations increasing reproductive isolation.

This study indicates that geographic distance, forest type and the interaction between the two are important in structuring stingless bee communities supporting the hypothesis that dispersal processes such as migration and colonization interact with niche specialization in determining local patterns of community composition.

TABLE 3.1. Geographic coordinates, number of individuals, number of genera, and number of species collected at each trapping site as well as total abundance (A) and richness (S) for each habitat type and collection site.

White-Sand						Non-White-Sand						Totals per site	
Site	Lat.	Long.	Ind. (*N*)	Gen. (*N*)	spp. (*N*)	Site	Lat.	Long.	Ind. (*N*)	Gen. (*N*)	spp. (*N*)	*A*	*S*
1-WS	-3.91	-73.55	215	8	16	1-NWS	-3.90	-73.55	92	9	15	307	20
2-WS	-3.95	-73.40	112	5	6	2-NWS	-3.97	-73.42	236	9	12	348	14
3-WS	-4.86	-73.61	415	19	16	3-NWS	-4.88	-73.64	31	6	9	446	20
A			742			*A*			359				
S			26			*S*			24				

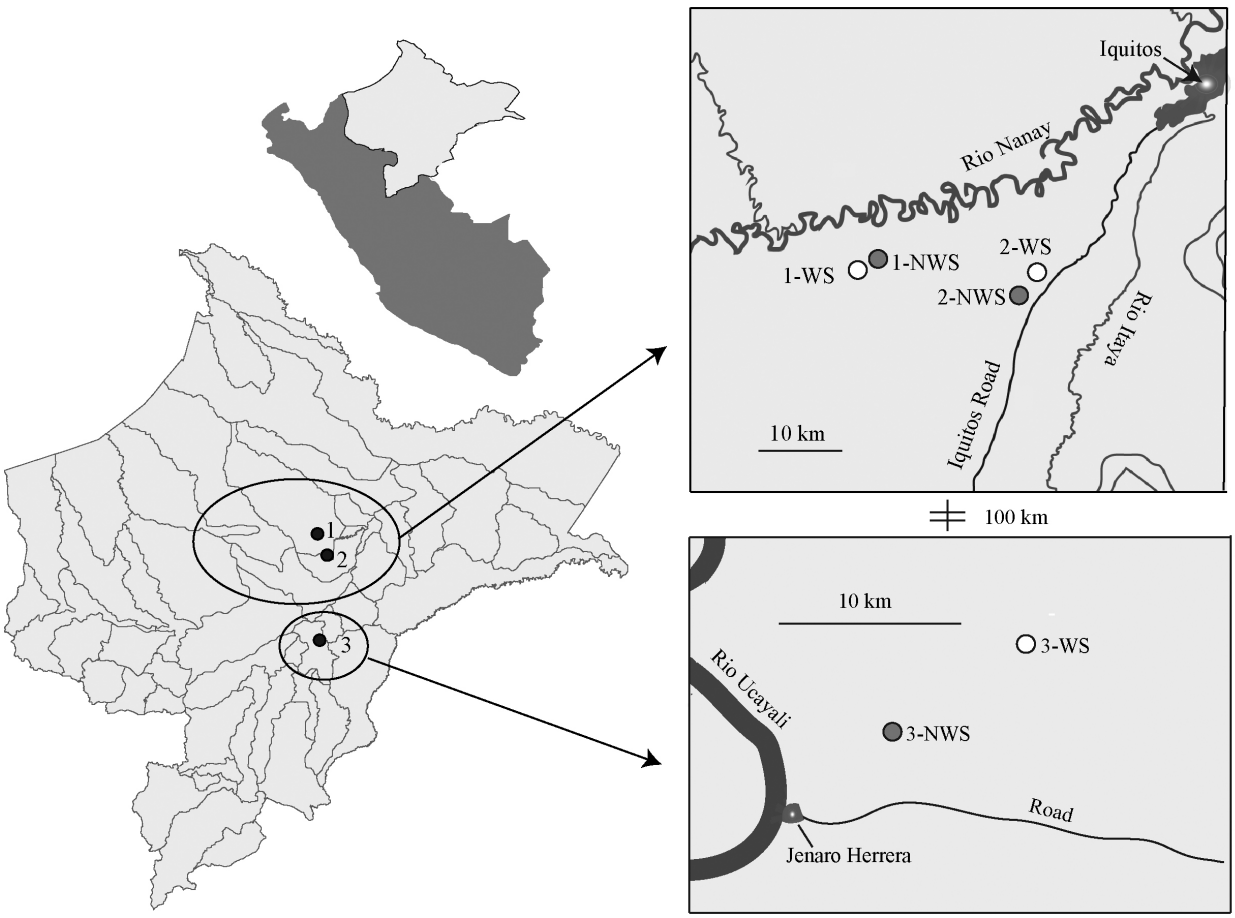

FIGURE 3.1. Sample sites and soil types where stingless bees were sampled in the region of Loreto, Peru. Numbered points represent the three areas where adjacent white-sand and non-white-sand forests were found. Each individual sampling site is displayed in the inset. Grey circles represent non-white-sand forests and white circles represent white-sand forests.

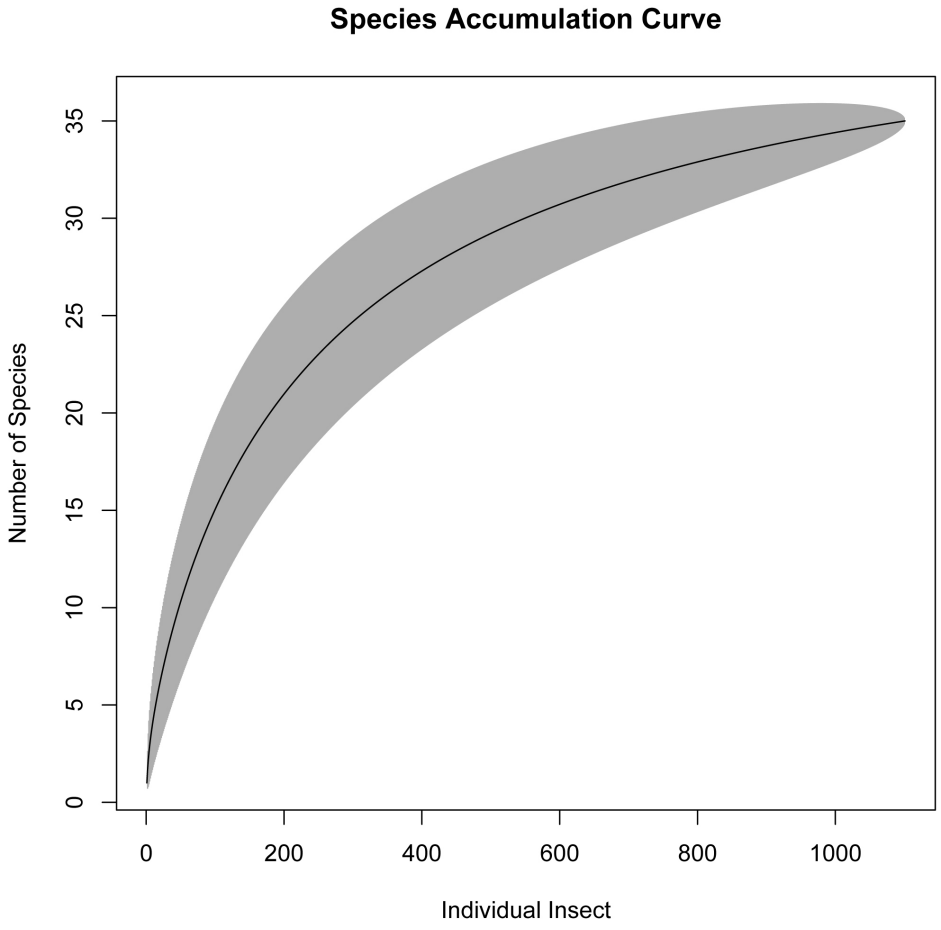

FIGURE 3.2. Species accumulation curve for stingless bees for all sites sampled.

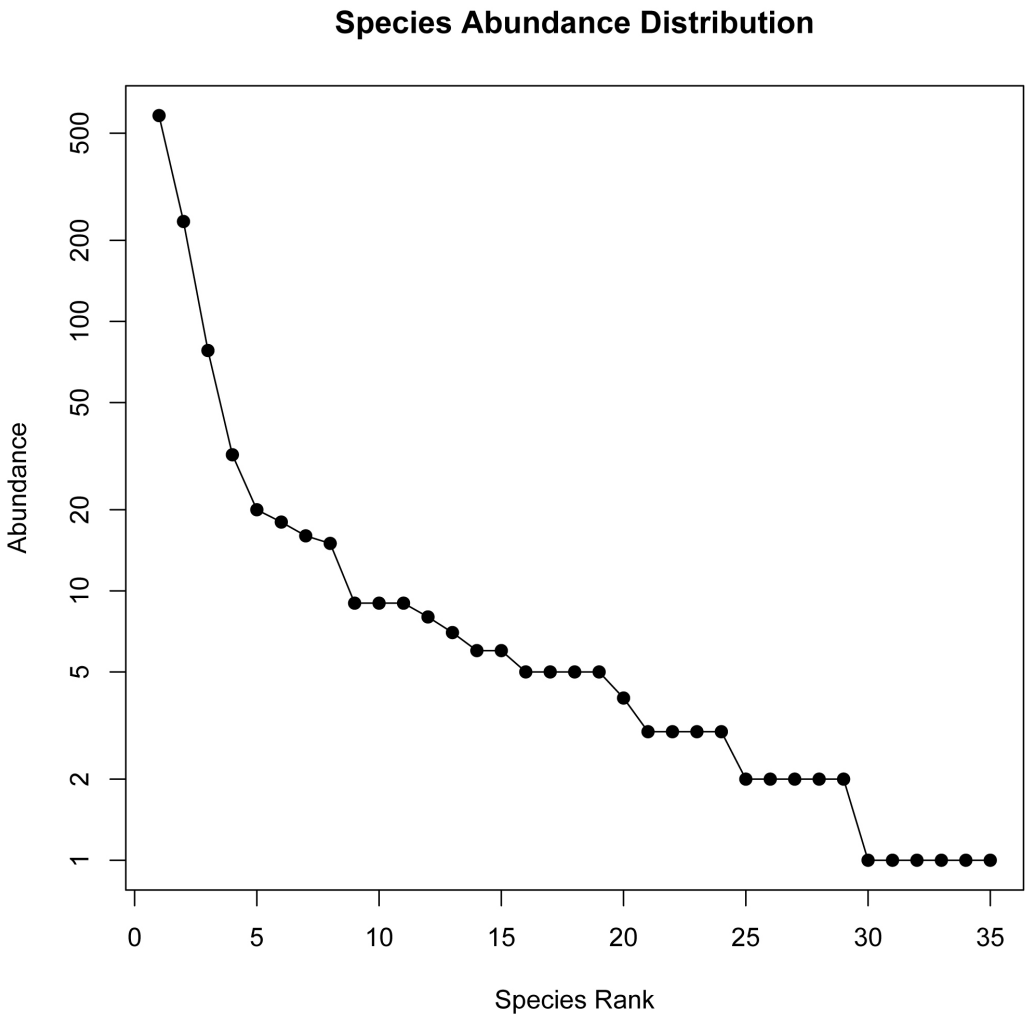

FIGURE 3.3. Species abundance distribution for all stingless bee species for all sites sampled.

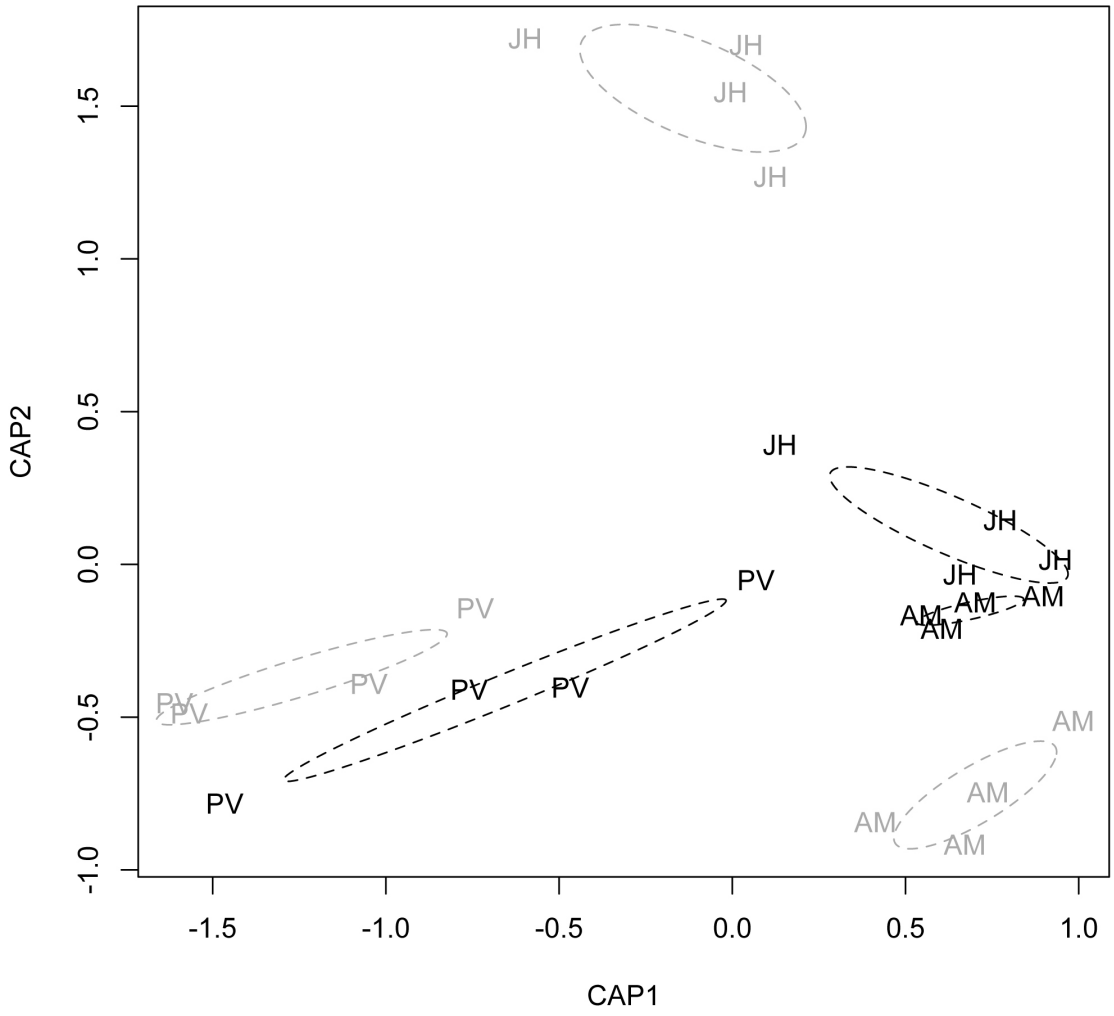

FIGURE 3.4. Results from distance-based redundancy analysis. Gray ellipses correspond to white-sand habitat and black ellipses correspond to non-white-sand habitat.

WORK CITED

ABRAHAMCZYK S., B. STEUDEL, AND M. KESSLER. 2010. Sampling Hymenoptera along a precipitation gradient in tropical forests: The effectiveness of different coloured pan traps. Entomol Exp Appl 137: 262–268.

ABRAHAMCZYK, S., J. KLUGE, Y. GARECA, S. REICHLE, AND M. KESSLER. 2011. The influence of climatic seasonality on the diversity of different pollinator groups. PLOS ONE. DOI: 10.1371/journal.pone.0027115.

ALVAREZ ALONSO, J., M. R. METZ, AND P.V.A. FINE. 2013. Habitat specialization by birds in Western Amazonian White-sand forests. Biotropica 45: 365-372.

ANDERSON, A. B. 1981. White-sand vegetation of Brazilian Amazonia. Biotropica 13: 199-210.

ANDERSON, M. J. 2001. A new method for non-parametric multivariate analysis of variance. Austral Ecol. 26: 32–46.

ANDERSON, M. J., AND T. J. WILLIS. 2003. Canonical analysis of principal coordinates: a useful method of constrained ordination for ecology. Ecology 84: 511–525.

ANDERSON, M. J., T. O. CRIST, J. M. CHASE, M. VELLEND, B. D. INOUYE, A. L. FREESTONE, N. J. SANDERS, H. V. CORNELL, L. S. COMITA, K. F. DAVIES, S. P. HARRISON, N. J. B. KRAFT, J. C. STEGEN, AND N. G. SWENSON. 2011. Navigating the multiple meanings of β diversity: a roadmap for the practicing ecologist. Ecol. Lett. 14: 19–28.

BATISTA MATOS, M. C., L. SOUSA-SOUTO, R. S. ALMEIDA, AND A. V. TEODORO. 2013. Contrasting patterns of species richness and composition of solitary wasps and bees (Insecta: Hymenoptera) according to land-use. Biotropica 45: 73-79.

BAWA, K. S. 1990. Plant-pollinator interactions in tropical rain forests. Ann. Rev. Ecol. Syst. 21: 299–422.

BORCARD, D., P. LEGENDRE, AND P. DRAPEAU. 1992. Partialing out the spatial component of ecological variation. Ecology 73: 1045–1055.

BRAY, J. R. AND J. T. CURTIS. 1957. An ordination of the upland forest communities of southern Wisconsin. Ecol. Monogr. 27:325-349.

CHAO, A. 1987. Estimating the population size for capture-recapture data with unequal catchability. Biometrics 43: 783–791.

CORBET, S., M. FUSSELL, R. AKE, A. FRASER, C. GUNSON, A. SAVAGE, AND K. SMITH. 1993. Temperature and the pollinating activity of social bees. Ecol. Entomol. 18: 17-30.

DICK, C. W., G. ETCHELECU, AND F. AUSTERLITZ. 2003. Pollen dispersal of tropical trees (*Dinizia excelsa*: Fabaceae) by native insects and African honeybees in pristine and fragmented Amazonian rainforest. Mol. Ecol. 12: 753-764.

DIECKMANN, U., AND M. DOEBELI. 2004. Adaptive dynamics of speciation: Sexual populations. *In* Adaptive Speciation. Dieckmann, U., Metz, J. A. J, Doebeli, M. and Tautz, D. (Eds.). pp 76-111. Cambridge University Press, Cambridge, UK.

DROEGE, S., V. J. TEPEDINO, G. LEBUHN, W. LINK, R. L. MINCKLEY, Q. CHEN, AND C. CONRAD, 2010. Spatial patterns of bee captures in North American bowl trapping surveys. Insect Conser. Diver. 3: 15-23.

ENGEL, M. S., AND S. F. DINGEMANS-BAKELS. 1980. Nectar and pollen resources for stingless bees (Meliponinae, Hymenoptera) in Surinam (South America). Apidologie 11: 341-350.

ERWIN, T. L. 1988. The tropical forest canopy: The heart of biotic diversity. *In* Wilson, E. O. (Ed.). Biodiversity. pp 123–129. National Academy Press, Washington DC, USA.

FIERRO, M. M., L. CRUZ-LÓPEZ, D. SÁNCHEZ, R. VILLANUEVA-GUTIÉRREZ AND R. VANDAME. 2012. Effect of biotic factors on the spatial distribution of stingless bees (Hymenoptera: Apidae, Meliponini) in fragmented neotropical habitats. Neotrop. Entomol. 41: 91-104.

FINE, P. V. A., R. GARCIA-VILLACORTA, N. C. A. PITMAN, I. MESONES, AND S. W. KEMBEL. 2010. A floristic study of the white-sand forests of Peru. Ann. Missouri Bot. Gard. 97: 283-305.

GARCIA, C., P. JORDANO, AND J. A. GODOY. 2007. Contemporary pollen and seed dispersal in a *Prunus mahaleb* population: patterns in distance and direction. Mol. Ecol. 16: 1947-1955.

GENTRY, A. H. 1988. Changes in plant community diversity and floristic composition on environmental and geographical gradients. Ann. Missouri Bot. Gard. 75: 1-34.

GOWER, J. C. 1966. Some distance properties of latent root and vector methods used in multivariate analysis. Biometrika 53: 325-338.

GOWER, J. C. 1985. Properties of Euclidean and non-Euclidean distance matrices. Linear Algebra Appl. 67: 81–97.

JHA, S., AND J. H. VANDERMEER. 2010 Impacts of coffee agroforestry management on tropical bee communities. Biol. Cons. 143: 1423-1431.

KERR, W. E. 1959. Bionomy of meliponids VI, *In* Symposium on food gathering behavior of Hymenoptera. pp. 24–31. Ithaca, NY, USA.

KRESS, J., W. R. HEYER, P. ACEVEDO, J. A. CODDINGTON, D. COLE, T. L. ERWIN, B. J. MEGGERS, M. G. POGUE, R. W. THORINGTON, R. P. VARI, M. J. WEITZMAN, AND S. H. WEITZMAN. 1998. Amazon biodiversity: Assessing conservation priorities with taxonomic data. Biodivers. Conserv. 7: 1577–1587.

KUHN-NETO, B., F. A. L. CONTRERA, M. S. CASTRO, AND J. C. NIEH. 2009. Long distance foraging and recruitment by a stingless bee, *Melipona mandacaia*. Apidologie 40: 472-480.

LAMAS, G., R. K. ROBBINS, AND D. J. HARVEY. 1991. A preliminary survey of the butterfly fauna of Pakitza, Parque Nacional del Manu, Perú, with an estimate of its species richness. Publicaciones del Museo Historia Natural, Universidad Nacional Mayor de San Marcos 40: 1–19.

LEGENDRE, P., AND L. LEGENDRE. 1998. Numerical Ecology. 2nd English Edition. Elsevier Science, Amsterdam, The Netherlands.

LEGENDRE, P., AND M. J. ANDERSON. 1999. Distance-based redundancy analysis: testing multispecies responses in multifactorial ecological experiments. Ecol. Monogr. 69: 1–24.

LEGENDRE, P., AND E. D. GALLAGHER. 2001. Ecologically meaningful transformations for ordination of species data. Oecologia 129: 271–280.

LEGENDRE, P., J. OKSANEN, AND C. J. F. TER BRAAK. 2011. Testing the significance of canonical axes in redundancy analysis. Methods Ecol. Evol. 2: 269–277.

LEGENDRE, P., AND L. LEGENDRE. 2012. Numerical Ecology. 3rd English Edition. Elsevier Science, Amsterdam, The Netherlands.

MISIEWICZ, T. M., AND P. V. A. FINE. In Press. Evidence for ecological divergence across a mosaic of soil types in an Amazonian tropical tree: *Protium subserratum* (Burseraceae). Mol. Ecol.

MAGURRAN, A. E. 2004. Measuring biological diversity. Blackwell Publications, Oxford, UK.

MEDINA, E., AND E. CUEVAS. 1989. Patterns of nutrient accumulation and release in Amazonian forests of the upper Rio Negro basin. *In* Proctor, J. (Ed.). Mineral nutrients in tropical forest and savanna ecosystems. pp. 217-240. Blackwell Scientific, Oxford, UK.

MICHENER, C. D. 2013. The Meliponini. *In* Pot-honey: a legacy of stingless bees. Vit, P., Pedro, S. R. M., and Roubik, D. (Eds.). pp. 3-18. Springer, New York, USA.
OKSANEN, J., F. G. BLANCHET, R. KINDT, P. LEGENDRE, P. R. MINCHIN, R. B. O'HARA, G. L. SIMPSON, P. SOLYMOS, M. HENRY, H. STEVENS, AND H. WAGNER. 2013. Vegan: Community Ecology Package. R package version 2.0.9. Http://CRAN.R-project.org/package=vegan.
PERFECTO, I., R. A. RICE, R. GREENBERG, AND M. E. VAN DER VOORT. 1996. Shade coffee: a disappearing refuge for biodiversity. Bioscience 46: 598–608.
PETANIDOU, T., AND W. N. ELLIS. 1996. Interdependance of native bee faunas and floras in changing Mediterranean communities. *In* Matheson, A., Buchmann, S. L., Toole, C. O., Westrich, P., and Williams, I. H. (Eds). The conservation of bees. Linnean Society Symposium Series 18. pp. 201-226. Academic Press, London, UK.
PITMAN, N. C. A., J. TERBORGH, M. R. SILMAN, P. NÚÑEZ, D. A. NEILL, C. E. CERÓN, W. PALACIOS, AND M. AULESTIA. 2001. Dominance and distribution of tree species in upper Amazonian terra firme forests. Ecology 82: 2101–2117.
ROUBIK, D. W. 1989. Ecology and natural history of tropical bees. Cambridge University Press, Cambridge, UK.
ROUBIK, D. W. 1995. Pollination of Cultivated Plants in the Tropics, Food and Agriculture Organization of the United Nations, Bulletin 118, Rome, Italy.
RUOKOLAINEN, K., H. TUOMISTO, M. J. MACIA, M. A. HIGGINS, AND M. YLI-HALLA. 2007. Are floristic and edaphic patterns in Amazonian rain forests congruent for trees, pteridophytes and Melastomataceae? J. Trop. Ecol. 23: 13–25.
SCHWARZ, H. F. 1948, Stingless bees (Meliponinae) of the western hemisphere. Bull. Am. Mus. Nat. Hist. 90: 1-546.
TEPENDINO, V. J., AND N. L. STANTON. 1981. Diversity and competition in bee-plant communities on short-grass prairie. Oikos 36: 35-44.
TERBORGH, J. 1985. Habitat selection in Amazonian birds. *In* M. L. Cody (Ed.). Habitat selection in birds. pp 311-338. Academic Press, New York, USA.
TUOMISTO, H., AND K. RUOKOLAINEN. 1994. Distribution of Pteridophyta and Melastomataceae along an edaphic gradient in Amazonian rain forest. J. Veg. Sci. 5: 25–34.
TYLIANAKIS, J. M., A. M. KLEIN, AND T. TSCHARNTKE. 2005. Spatiotemporal variation in the diversity of hymenoptera across a tropical habitat gradient. Ecology 86: 3296-3302.
WESTPHAL, C., R. BOMMARCO, G. CARRE, E. LAMBORN, N. MORISON, T. PETANIDOU, S. G. POTTS, S. P. M. ROBERTS, H. SZENTGYORGYI, T. TSCHEULIN, E. VAISSIERE, M. WOYCIECHOWSKI, J. C. BIESMEIJER, W. E. KUNIN, J. SETTELE, AND I. STEFFAN-DEWENTER. 2008. Measuring bee diversity in different European habitats and biogeographical regions. Ecol. Monogr. 78: 653-671.
WHITNEY, B. M., AND A. J. ALVAREZ. 1998. A new *Herpsilochmus* antwren (Aves: Thamnophilidae) from northern Amazonian Peru and adjacent Ecuador: the role of edaphic heterogeneity of tierra firme forest. Auk 115: 559–576.
WUELLNER, C. T. 1999. Nest preference and success in a gregarious, ground-nesting bee *Dieunomia triangulifera.* Ecol. Entomol. 24: 471-479.

CHAPTER 4: The contribution of multiple barriers to reproduction between edaphically divergent lineages in the Amazonian tree *Protium subserratum* (Burseraceae).

ABSTRACT

Disentangling the strength and importance of various barriers to reproduction that arise between diverging lineages is central to our understanding of species origin and maintenance. Evidence suggests that divergent adaptation can drive the development of ecologically and genetically based barriers to reproduction yet, relatively few studies have investigated which reproductive isolating mechanisms are most important in naturally occurring ecologically divergent lineages and most have focused on animal lineages. To date, only a handful of studies have comprehensively quantified the strength of individual reproductive isolating mechanisms among closely related plant species. Even fewer studies have focused on tropical plants and those have focused on herbaceous lineages. In this study we systematically examine multiple barriers to reproduction between diverging lineages of soil specialist ecotypes of *Protium subserratum* (Burseraceae), an Amazonian tropical tree. Specifically we aimed to (1) Quantify the contributions of pre-zygotic and post-zygotic barriers to isolation including ecogeographic isolation, flowering phenology, pollinator assemblage, pollen adhesion, pollen germination, pollen tube growth, fertilization/initial seed development and hybrid fitness and (2) Calculate the total amount of reproductive isolation as well as the relative contribution of each barrier to total reproductive isolation. We were able to identify four active barriers to reproduction including ecogeographic isolation, differential pollen adhesion, differences in pollinator assemblages, and low levels of hybrid seed development. We demonstrate that a combination of pre-zygotic and post-zygotic barriers to reproduction act to maintain near complete reproductive isolation between edaphically divergent populations of the tropical tree, *P. subserratum*.

INTRODUCTION

Disentangling the strength and relative importance of the different barriers to reproduction that arise between diverging lineages is central to our understanding of new species formation and the maintenance of species diversity. Numerous different reproductive isolating mechanisms have been characterized across a wide range of taxa (Coyne & Orr 2004) however; complete reproductive isolation is rarely the consequence of any one isolating mechanism. More commonly, a number of barriers to reproduction will accumulate over time, additively contributing to the total level of reproductive isolation between lineages. Consequently, assessing the relative importance of many different barriers to reproduction between closely related lineages or species pairs is essential to our understanding of speciation (Dobzhansky 1951; Coyne 1992; Schluter 2001; Coyne & Orr 2004).

Barriers to reproduction act sequentially to limit gene flow and hence are categorized by the point in an organism's life in which they act. While individual barriers to reproduction may be equally strong in limiting gene flow, early-acting barriers will have proportionately large effect on the level of total reproductive isolation (Coyne & Orr 2004). In angiosperms, reproductive barriers are temporally classified in four ways, those that act prior to pollination and after pollination (Grant 1971), those that act prior to the fusion of parental gametes (pre-zygotic barriers), and those that act afterwards (post-zygotic barriers). Pre-pollination, and therefore also pre-zygotic, barriers to reproduction include geographic and ecogeographic isolation, temporal isolation, and mechanical floral isolation (Grant 1949; Rieseberg & Willis 2007; Lowry *et al.* 2008; Widmer *et al.* 2008; Schiestl & Schluter 2009). Post-pollination, pre-zygotic barriers

include competition between conspecific and heterospecific pollen, pollen-pistil incompatibilities and gametic incompatibilities. Final barriers to reproduction are incurred through post-pollination, post-zygotic isolating mechanisms and include embryo abortion, ecologically based low hybrid fitness, and hybrid sterility (Dobzhansky 1937; Mayr 1942; Coyne & Orr 2004; Rieseberg & Willis 2007).

While it is generally thought that ecologically based barriers to reproduction are most likely to develop as a result of divergent natural selection because they rely on a phenotype environment interaction, genetically based reproductive isolation may also readily arise as a result of differential adaptation (Gavrilets 2004). Theory posits that when different alleles are favored by divergent selection in different habitats combinations of habitat specific alleles within each environment will work well together. Alternatively, alleles from different environments that are not generally coupled may be incompatible when brought together in the genome of a hybrid (Gavrilets 2004). Evidence that divergent adaptation can play a role in driving the development of genetically based barriers has been demonstrated in numerous studies (Bolnick *et al.* 2006, Funk *et al.* 2006, Dettman *et al.* 2007 Sambatti *et al.* 2008). However, given that divergent selection can drive the evolution of a large variety of barriers to reproduction few studies have investigated which reproductive isolating mechanisms are most important in naturally occurring ecologically divergent lineages and most have focused on animal lineages (reviewed in Nosil *et al.* 2005). To date, only a handful of studies have comprehensively quantified the strength of individual reproductive isolating mechanisms among closely related plant species, the majority of which are herbaceous and found in temperate zones (Morrison *et al.* 1994; Chari & Wilson 2001; Ramsey *et al.* 2003; Martin & Willis 2007; Husband & Sabara 2004; Hoffmann *et al.* 2008; Dell'Olivo *et al.* 2011). Even fewer studies have focused on tropical plants and those that have focused on herbaceous lineages (Kay 2006; Scopece *et al.* 2013). Yet, Neotropical forests harbor the highest plant diversity on Earth, and some authors have suggested that this may be due to elevated speciation rates (Mittelbach et al. 2007). Furthermore, the majority of Neotropical plant species richness is comprised of trees (Gentry & Dodson 1987). To our knowledge no study has thoroughly examined reproductive isolation across a suite of pre-zygotic and post-zygotic barriers in trees, let alone tropical trees.

In this study we provide a critical missing link in speciation studies by systematically examining multiple barriers to reproduction between diverging lineages of habitat specialist ecotypes of *Protium subserratum* (Burseraceae), an Amazonian tropical tree. Specifically, we aimed to (1) Quantify the contributions of pre-mating, pre-zygotic barriers to reproduction (ecogeographic isolation, flowering phenology, pollinator assemblage), post-mating, prezygotic, barriers to reproduction (pollen adhesion, pollen germination and pollen tube growth), and post-zygotic barriers to reproduction (fertilization/initial seed development and hybrid fitness). (2) Calculate the total amount of reproductive isolation as well as the relative contribution of each barrier to total reproductive isolation As per Coyne & Orr (1997) and Ramsey *et al.* (2003).

METHODS

Study System

The genus *Protium* (Burseraceae) comprise small to large canopy trees with the center of diversity existing in the Neotropics where ca. 100 of the ca. 140 species are found (Daly 1987, Fine *et al.* in press). Specialization onto different soils is common within the genus and has been particularly well documented on the mosaic of different soil types found across the Western lowland Amazon basin. Edaphic specialization onto nutrient poor white-sand habitat islands as

well as common and more fertile clay and brown sand soils found in the Peruvian Amazon have occurred independently multiple times within the genus (Fine *et al.* 2005).

Protium subserratum, a soil generalist tree found across the Amazon, contains both genetically and morphologically differentiated sub-populations endemic to brown-sand and white-sand soils often found in parapatry (Misiewicz & Fine *in press*). While population genetic analysis suggests the presence of low levels of gene flow across habitat boundaries, populations of both ecotypes are clearly maintaining their genetic and morphological integrity, suggesting that barriers to reproduction are present (Misiewicz & Fine *in press*).

Soil specialist populations of *P. subserratum* differ from one another in leaf morphology, with individuals found on white-sand soils exhibiting thick, pubescent leaflets with entire margins, whereas brown-sand individuals have thinner, glabrous leaflets with some marginal teeth (Misiewicz & Fine *in press*). *P. subserratum* is dioecious and male and female flowers closely resemble each other. Flowers within and between ecotypes are relatively uniform with the exception that the adaxial side of the petals of the white-sand ecotype is pubescent (Misiewicz per obs), a character that has been overlooked in previous studies (Daly & Fine 2011). Flowers are fragrant and nectiferous with small white petals ~ 5mm in length. Pollen grains are approximately 6.4μm in length, invisible to the naked eye. Once a tree begins to flower it does so abundantly. While the lifespan of an individual flower is about 48 hours, a tree will generally remain at flowering peak for one to two weeks and can produce flowers over one to two months (Misiewicz per obs.).

All components of reproductive isolation were studied at a contact zone between a parapatric white-sand and brown-sand population in the Allpahuayo Mishana Reserve in the Amazonian region of Loreto, Peru.

Sampling and Genotyping

Each population of white-sand and brown-sand *P. subserratum* individuals covered approximately one square km. All adult individuals of *P. subserratum* found in each population were tagged and mapped (white-sand, N=14; brown-sand, N=14). Leaves were collected and dried in silica for DNA extraction. A voucher specimen for each population was deposited in the Herbarium Amazonense at the Universidad Nacional de la Amazonía Peruana in Iquitos, Peru (AMAZ) and the University Herbarium at the University of California, Berkeley (UC). Seedlings (N=178) were also collected from the seed shadows of white-sand (N=3) and brown-sand (N=3) maternal trees at the contact zone where male and female trees of both ecotypes are found within 10m of each other to test for a signal of natural selection against hybrids (see section *Hybrid fitness*).

All individuals were genotyped using thirteen nuclear microsatellite markers previously developed for *P. subserratum* (Misiewicz *et al.* 2012) and shown to be effective in population level differentiation and the identification of hybrids between soil ecotypes (Misiewicz & Fine *in press*). DNA extraction and genotyping protocols followed those described in Misiewicz *et al.* (2012).

Genetic variation, Hardy-Weinberg equilibrium, null alleles and linkage disequilibrium

Number of alleles (A), observed and expected heterozygosities (Ho and He), the inbreeding coefficient (F_{IS}), deviation from Hardy-Weinberg equilibrium (HWE) and linkage disequilibrium (LD) for each population was calculated across all loci as described in Misiewicz & Fine (*in press*).

Ecotypic differentiation and hybrid assignment

Population genetic structure was assessed using STRUCTURE 2.3.3 (Pritchard *et al.* 2000) as described in Misiewicz & Fine (*in press*). The model-based clustering method of NewHybrid version 1.1b (Anderson & Thompson 2002) was then used to assign individuals to one of six hybrid classes, white-sand parental, brown-sand parental, F1 hybrid, F2 hybrid, white-sand backcross hybrid, or brown-sand backcross hybrid using the methods described in Misiewicz & Fine (*in press*). Individuals assigned to a hybrid class with a posterior probability (pp) > 90% were considered to be hybrids between the two ecotypes.

Pre-pollination Isolating Barriers

Ecogeographic isolation

Adaptation to different habitats can directly influence the geographic distribution of species, potentially limiting reproductive potential between populations (Sobel *et al.* 2010). Because *P. subserratum* ecotypes are known to be associated with different soils, which have their own distinct geographic distributions we quantified the contribution of geographic isolation due to habitat specialization, ecogeographic isolation, as a barrier to reproduction. Ecogeographic isolation was estimated by mapping the distribution of all known adult brown-sand and white-sand *P. subserratum* ecotypes in the Allpahuayo-Mishana National Reserve. Overall we included 79 brown-sand individuals and 94 white-sand individuals. Spatial isolation was determined *in silico* using custom scripts in R (3.0.3) by placing a grid of quadrats over the entire area within which individuals were mapped. We then sampled every quadrat that included at least one tree. Quadrats (N=1798) of 300m X 300m were placed over an area of approximately 162.5km2. We chose $300m^2$ quadrats as the distance that we thought would best represent the distance of pollen flow among trees because stingless bees, with average foraging ranges of up to 500m from their nest site (citation), are thought to be the effective pollinators for both ecotypes. We compared the number of heterospecific quadrats in which only individuals of the same ecotype were sampled to the number of conspecific quadrats, in which individuals of both ecotypes were sampled. The index for ecogeographic isolation ($RI_{ecogeography}$; Table 4.1) ranges from zero, where ecotypes are found in complete sympatry to one, where ecotypes are found in complete allopatry.

Flowering phenology

We quantified the overlap in flowering times by censusing brown-sand (N= 39) and white-sand trees (N= 12) bi-weekly from January 2006 – December 2008. Using binoculars each tree's canopy was surveyed for the presence or absence of flowers. The proportion of individuals flowering in each population was plotted over time.

Pollinator assemblage

Although only relatively minor differences in floral morphology have been observed between *P. subserratum* ecotypes, pollinator preferences may still exist. In order to determine the extent to which pollinator specialization may limit gene flow between ecotypes we compared pollinator assemblages on brown-sand and white-sand ecotype trees. Floral visitors to each ecotype were observed using digital video cameras placed in the canopies of six trees: white-sand and brown-sand male and female trees (white-sand male, N=2; white-sand female, N=2, brown-sand male, N=1; brown-sand female, N=1). Video recordings of infloresences were taken between 0900h and 1700h between June 1 and July 30 2012. Each individual recording lasted a duration of 15-30 min. A total of 6 hours of video footage (BS=192min; WS=169min) covering a total of 572 individual flowers was collected (BS, N=313; WS, N=169). All visits were noted and each insect species recorded was assigned a morphospecies identification. A single visit was

defined from the time that an individual entered the recording frame to the time it exited. If antagonistic behavior such as nectar robbing was observed the episode was counted separately and omitted from the final analysis. We compared pollinator assemblages across ecotypes by comparing the visitation rates at each ecotype for every insect visitor (visits per insect morphospecies/ total recording time for the respective ecotype). Values were tested using Wilcoxon rank sum tests. The index for pollinator assemblage ($RI_{pollinator\ overlap}$; Table 4.1) ranges from zero, where pollinator visitors to both ecotypes completely overlap to one, where there is complete turnover of pollinator visitors among ecotypes.

Post-pollination Isolating Barriers

Hand pollinations

In order to prevent pollination contamination female inflorescences were bagged while flowers were still in their buds using a tied bag made from synthetic woven interfacing. Male and female flowers were checked daily. Upon anthesis of female flowers, we collected newly opened male flowers from neighboring trees to use male stigmas for hand pollination experiments. Floral bags were removed from female inflorescences and flowers were hand pollinated by gently bringing anthers into contact with the receptive stigma. Hand pollinated flowers were marked on their pedicel and the inflorescence was re-bagged. No more than 30% of flowers on any given inflorescence were hand-pollinated and the remaining flowers were left as negative controls. Inflorescences of flowering female white-sand trees were inaccessible and as a result all hand pollinations were made using brown-sand maternal trees. However, because population genetic structure results suggest that hybrids exist in both white sand and brown sand populations (Misiewicz and Fine in press), we have no reason to believe that reproductive isolation is unidirectional. Only one type of pollination treatment (hybrid or parental) was made per inflorescence to avoid selective abortion or preferential resource allocation to some flowers over others.

Pollen adherence, germination and pollen tube growth

Parental (N=33) and hybrid (N=35) hand pollinations were made using one maternal brown-sand individual. Hand pollinated flowers were collected 48 hours after pollinations were made and fixed for 24 hours at room temperature in 4% paraformaldehyde solution (w/v) in phosphate buffered saline (PBS) solution (0.01M phosphate buffer, 0.0027M potassium chloride, and 0.137 M sodium chloride). Flowers were then transferred to PBS solution and stored at 20°C. Prior to tissue preparation, petals were removed from the preserved flowers. Floral tissue was cleared and softened in 4M NaOH for 6 hours, rinsed and stained with 0.1% de-colorized aniline blue (Martin 1959) for an additional 6 hours and then squashed and mounted in a drop of the aniline blue staining solution for pollen tube visualization. Slides were observed under fluorescence microscopy (Zeiss Axioimager), with UV (400nm) excitation and photographed (QIClick digital CCD camera). All adhered pollen grains and germinated pollen grains were counted. While initial pollen tube germination was easily observed in all treatments pollen tubes were not observed growing down the style in either treatment. Whether or not this was due to a lack of fluorescence in the pollen tubes or because 48 hrs was not sufficient for pollen tube growth to extend into the style is equivocal.

The number of adhered pollen grains and proportion of adhered pollen grains showing germination and initial pollen tube growth for each treatment was compared using a Mann-Whitney U test. The index for pollen adherence ($RI_{pollen\ adherance}$; Table 4.1) ranges from zero to one, where zero indicates no difference between conspecific adhesion to the stigma and

heterospecific pollen adhesion to the stigma and one indicates that heterospecific pollen does not adhere to the stigma. The index for pollen germination ($RI_{pollen\ germination}$; Table 4.1) ranges from zero to one, with a zero meaning that there is no difference in the proportion of conspecific pollen grain germination and a value of one meaning that there is no heterospecific pollen grain germination on the stigma.

Fertilization and seed set

A total of 178 hand crosses were made using female flowers from brown-sand maternal trees (N=4) out of which 140 hybrid hand crosses were made with white-sand males (N=4) and 38 parental crosses were made using brown-sand males (N=2). Floral bags were removed two weeks after pollination once all negative controls were no longer receptive and fruit formation was observable in the hand-pollinated flowers. An additional 46 inflorescences with a total of 1,166 flowers from the same maternal trees were monitored as positive controls in order to determine the natural pollination rate. Seed set among parental and hybrid crosses were pooled among treatments and compared using a Fisher exact test for 2X2 contingency table. The index for fertilization and initial seed development ($RI_{seed\ development}$; Table 4.1) ranges from zero to one. A value of zero indicates that all cross-fertilizations were equally successful and a value of one indicates that fertilizations with heterospecific pollen were never successful.

Hybrid fitness

In order to test the prediction that hybrid individuals are less fit than parental type offspring we compared genotypes of first year seedlings, juveniles and adult trees. We predicted that if hybrids were less fit than parental types we would observe a higher relative abundance of F1 hybrids in first year seedlings and that they would decrease in frequency as age class increased due to environmental filtering. First year seedlings (N=73 white-sand; N=43 brown-sand) were identified by their persistent seed coat, no more than two leaves, and a height under 10cm. Juvenile seedlings (N=19 white-sand; N=14 brown-sand) had two or more leaves present with heights ranging from 15-30cm. All seedlings were collected within a 15m radius of maternal white-sand (N=3) and maternal brown-sand (N=4) trees found at the contact zone between the two ecotypes. All known adult trees (N=14 white-sand; N=14 brown-sand) from the two populations were also sampled and genotyped. All individuals were genotyped and assigned to hybrid classes as described in Misiewicz & Fine (*in press*).

Total isolation

Cumulative reproductive isolation between white-sand and brown-sand ecotypes of *P. subserratum* was calculated following the methods of Coyne and Orr (1989, 1997) and Ramsey *et al*. (2003) where total reproductive isolation (T) is determined from individual components of reproductive isolation (RI) at successive stages in the life history. We first estimated the strength of RI for each mechanism followed by the absolute contribution of that mechanism. The absolute contribution (AC) is the reduction in gene flow after reductions from previous stages of reproductive isolation have been taken into account. Comparisons across isolating mechanisms were made using the relative contribution (RC) of each component, calculated as the AC of a particular component divided by T.

RESULTS

Genetic variation, Hardy-Weinberg equilibrium, null alleles and linkage disequilibrium

We found 7-20 alleles per locus (summed across both populations) and an average of 4.9-6.6 alleles per population (averaged across all loci). Observed and expected heterozygosity varied from 0.45-0.49 and 0.44-0.52 across both populations. Inbreeding coefficients were low

for both populations (F_{IS} = -0.02; 0.07) (Table 1). Deviation from HWE was observed in 2 loci (prot13 in the white-sand population and locus prot104 in the brown-sand population). LD was observed between loci prot70 and prot83, prot29and100, prot70 and prot100, prot83 and prot100, and prot104 and prot100 in the white-sand population and between loci prot29 and prot67, and prot70 and prot78 in the brown-sand population.

Ecotypic differentiation and hybrid assignment

Population structure analysis revealed strong patterns of genetic differentiation across soil types. K=2 was the best supported model using evaluation methods of both Evanno *et al.* (2005) and Pritchard *et al.* (2000) where the two genetic clusters are clearly segregated by soil ecotype. Two individuals found in brown sand-soils appeared to have low genetic contributions from the white-sand soil ecotype however further hybrid assignment analysis confidently assigned all individuals as pure parental white-sand or pure parental brown-sand (pp>0.95). No individuals were identified as F_1, F_2 or backcross hybrids.

Pre-pollination Isolating Barriers

Ecogeographic isolation

Twenty-nine of the sampled quadrats contained *P. subserratum* individuals. Eleven quadrats contained only brown-sand individuals, 12 contained only white-sand individuals and six contained at least one white-sand and brown-sand individual. $RI_{ecogeography}$ was calculated as 0.80.

Flowering phenology

Flowering times for brown-sand and white-sand individuals overlapped across all four years (Figure 4.1). As overlap in flowering phenology is high it is unlikely to contribute to reproductive isolation between the two ecotypes and therefore was not used in the overall calculations of reproductive isolation.

Pollinator assemblage

A total of 14 different insect visitors were observed visiting *P. subserratum* ecotypes, 4 of which were observed visiting both ecotypes (Table 4.2). Pollinator assemblages visiting brown-sand and white-sand ecotypes differed significantly from one another (U=49, Z=-2.61, p<0.01). $RI_{pollinator\ overlap}$ was set to 0.23.

Post-pollination Isolating Barriers

Pollen adherence, germination and pollen tube growth

Significantly more pollen grains were adhered to stigmas from the parental cross treatment (mean = 159) than the hybrid cross treatment (mean = 112) (U=371, Z=2.36, p<0.05) however the proportion of adhered pollen grains which germinated did not significantly differ between treatments (U=538, Z=-0.27, p>0.1). None of the negative controls set seed. $RI_{pollen\ adhesion}$ was 0.3 and $RI_{pollen\ germination}$ was 0.04 (Figure 4.2a).

Fertilization and seed set

Seed set per pollination was significantly lower in the hybrid pollination treatment (16.4%) compared to the parental pollination treatment (39%) (p<0.01). $RI_{seed\ development}$ was 0.57 (Figure 4.2b).

Hybrid fitness

All first year seedlings, juveniles and adults genotyped were confidently assigned as pure parental individuals (PP>0.95). No hybrid individuals were sampled in any age class. Because no hybrid individuals were detected at any age class, ecologically based low hybrid fitness is

unlikely to be and important barrier to reproduction and therefore this measure was not used in the overall calculations for barriers to reproduction.

Total isolation

The strength of each individual barrier to reproduction, the relative contribution of each barrier to total isolation as well as the relative contribution of each barrier to isolation at the contact zone (assuming no geographic isolation) are summarized in Figure 4.5. White-sand and brown-sand ecotypes were found to be in near complete reproductive isolation (T=0.96) with the majority of the isolation occurring prior to pollination. Post-pollination reproductive barriers appeared to be genetically based as opposed to ecologically based.

DISCUSSION

Our study showed that reproductive isolation between parapatric populations of edaphic specialist white-sand and brown-sand ecotypes of *P. subserratum* is almost complete. We quantified seven potential isolating barriers and showed that ecogeographic isolation accounts for the majority of isolation between ecotypes. In sympatry we found that genetically based post-pollination isolating mechanisms including pollen adhesion and failed fertilization/low hybrid seed development were most important in isolating the two ecotypes. Pollinator assemblages also appeared to differ between ecotypes potentially limiting pre-pollination gene flow to some extent. We found high overlap in flowering times and no evidence of hybrids at a contact zone suggesting that genetic barriers to gene flow are much stronger than ecological barriers. Results from this study are congruent with our previous study, which demonstrated very low levels of gene flow and hybridization between three adjacent pairs of *P.subserratum* soil specialist ecotypes.

Ecogeographic Isolation

Ecogeographic isolation is often used to describe differences in geographic ranges that are dictated in part by habitat specialization. Essentially, the geographic ranges of multiple species will not overlap any more than the geographic overlap of their associated habitats (Ramsey *et al*. 2003; Sobel *et al*. 2010). While a variety of studies have demonstrated the clear importance of ecogeographic isolation as a strong early acting barrier to reproduction (Ramsey *et al*. 2003; Glennon *et al*. 2012) it has been largely ignored in studies of reproductive isolating barriers (reviewed in Sobel *et al*. 2010). We found very high levels of ecogeographic isolation between *P. subserratum* brown-sand and white-sand ecotypes. Previous reciprocal transplant studies have demonstrated that differences in morphological vegetative characters between white-sand and non-white sand ecotypes are genetically based (Fine et al. 2013b). Additionally, the fact that populations of white-sand and brown-sand ecotypes are commonly found in parapatry (Fine *et al.* 2013a, Misiewicz & Fine *in press*) suggests that seeds may frequently be dispersed across habitat boundaries but offspring may fail to establish due to low fitness in either habitat type. The strong edaphic associations seen in closely related ecotypes of *P. subserratum* is a pattern shared by numerous tropical tree congeners found across multiple plant families (Fine *et al.* 2010) suggesting that ecogeographic isolation as a result of adaptation to differing soil types likely plays an important role in diversification and maintenance of diversity in many Amazonian trees.

Pollinator Assemblages

Many studies of reproductive isolation have focused on closely related lineages that show high levels of floral differentiation. As expected, differences in floral architecture and reward leads to results in differences in both visiting pollinator communities as well as pollinator

efficiency sometimes resulting in near complete pre-pollination isolation when lineages were found in sympatry (Ramsey 2003, Kay 2006, Scopece 2013; Whitehead & Peakall 2014). These studies suggest that in some cases animal pollinators may play an important role in maintaining reproductive isolation particularly between lineages with divergent floral morphologies. While white-sand and brown-sand ecotypes of *P. subserratum* exhibit a generalist floral morphology and display only slight floral variation compared to lineages studied by Ramsey (2003), Scopece (2013), and Kay (2006), we still detected significant differences in bee communities visiting each ecotype. While observed floral morphologies do not differ substantially between the two ecotypes, pollinator communities may be indirectly influenced by biological differences in ecotype or habitat differences. For instance, white-sand and brown-sand ecotypes of *P. subserratum* show distinct quantitative and qualitative differences in their leaf chemistry (Fine *et al.* 2013b), while these chemical differences have been attributed to herbivore defense they may also be expressed to some extent in floral tissue. Chemical volatiles have been shown to play an important role in plant herbivore defense across many plant groups and some studies have demonstrated that differences in anti-herbivore defense chemistry may also indirectly influence bee preference (Kessler *et al.* 2011; Adler *et al.* 2012). Alternatively, Misiewicz *et al.* (*in review*) demonstrated high turnover in stingless bee communities across brown-sand and white-sand soils and over very small areas. If pollinators are foraging over small distances then turnover in pollinator communities visiting white-sand and brown-sand trees may be the indirect result of pollinator behavior.

Pollen adhesion and germination

Intraspecific pollen–pistil recognition at the stigma surface is common in plants as a mechanism to avoid conspecific pollen germination and growth that would deplete female tissue resources (Heizmann *et al*. 2000). Prezygotic genetic incompatibilities may evolve over time due to genetic drift or as a direct result of natural selection through the process of reinforcement (Dobzhansky 1940; Grant 1965). If pollen transfer between populations that are locally adapted to different habitats results in less fit hybrid offspring then natural selection should favor individuals that do not waste resources on inferior offspring. We found significantly lower levels of white-sand pollen adhered to brown-sand flowers than we did brown-sand pollen adhered to brown-sand flowers but no difference in the proportion of adhered pollen germination between treatments. Our results suggest that pollen-pistil incompatibilities limiting pollen adhesion could be an early acting post-pollination barrier to reproduction. However, we cannot exclude alternative explanations of these results. Lower levels of pollen adherence in hybrid hand pollinations could be explained if white-sand flowers produce lower over-all quantities of pollen than brown-sand ecotype flowers. While dehiscent anthers of both white-sand and brown-sand ecotypes were observed using a dissecting microscope and flowers of both ecotypes appeared to release equally large quantities of pollen we did not quantify pollen production. Regardless, pollen adhesion in hybrid crosses, while lower than in parental crosses, was still high (averaging 112 pollen grains per stigma) and the fact that the proportion of adhered grains that germinated did not differ between treatments suggests that if initial pollen-pistil incompatibilities were present they are far from complete.

Fertilization/Seed development

We detected additional genetically-based barriers to reproduction in hand crossing experiments. However, because we were unable to visualize pollen-tube growth in the style for any treatment we could not discern whether the difference in seed development two weeks after pollination was due to a lack of fertilization through genetically based pre-zygotic barriers such

as pollen tube/style or gametic incompatibilities or if it was due to genetically based post-zygotic hybrid mortality. In any case it is clear that genetic post-pollination isolating mechanisms are in place.

Asymmetrical barriers to reproduction

One limitation of this study was that we were not able to compare the effectiveness of reproductive isolating barriers in both directions due to the inability to safely access flowers on female white-sand trees using our climbing techniques. While we couldn't test for asymmetry of individual barriers to reproduction the presence of a small number of putative hybrid individuals were detected on both soil types by Misiewicz & Fine (*in press*) suggests that barriers to reproduction are not asymmetrical. Additionally, we failed to detect any hybrid seedlings in seed shadows of white-sand and brown-sand maternal trees at contact zone suggesting that total RI is complete or near complete in both directions.

We were able to identify five active barriers to reproduction between soil specialist ecotypes of the Amazonian tree, *P. subserratum*. While studies have explored the relative strength of pre-pollination versus post-pollination and prezygotic versus postzygotic barriers to reproduction in plants most of these studies focused on temperate systems and none have comprehensively explored barriers in trees. White-sand and brown-sand ecotypes of *P. subserratum* are strongly isolated from one another. We show that hybridization between ecotypes is extremely rare as a result of both ecologically based pre-pollination barriers, genetically based post-pollination pre-zygotic barriers and likely post-zygotic barriers to reproduction. Our results show the clear importance of multiple barriers occurring before and after pollination and likely before and after fertilization in maintaining reproductive isolation between edaphically divergent populations of the tropical tree, *P. subserratum*.

Table 4.1. Formulas for individual components of reproductive isolation (RI), isolation measure, and the range of values possible for that barrier.

Barrier/ formula	Isolation measure	Range
Premating		
$RI_{ecogeography}$= 1 – (no. heterospecific quadrats/(heterospecific quadrats + conspecific quadrats)	Ecogeographic isolation	0-1
$RI_{pollinator\ overlap}$ = 1 – (visitation rate of pollinator species visiting both ecotypes/ total visitation rate)	Pollinator species assemblage	0-1
Postmating- intrinsic		
$RI_{pollen\ adhesion}$ = 1- (mean no. pollen grains adhered to stigma per interspecific pollination/ mean no. pollen grains adhered to the stigma per intraspecific pollination)	Pollen adhesion	0-1
$RI_{pollen\ germination}$= 1- (proportion of adhered pollen grains germinated per interspecific pollination/ proportion of adhered pollen grains germinated per intraspecific pollination)	Pollen germination	0-1
Postmating		
$RI_{seed\ development}$= 1 – (no. seeds per interspecific pollinations/ no. seeds per intraspecific pollinations)	Fertilization/ seed development	0-1

Table 4.2. Populations sampled, number of individuals sampled (N), observed heterozygosity (Ho), expected heterozygosity (He), average number of alleles (A), and inbreeding coefficient (F_{IS}).

Population	N	Ho	He	A	F_{IS}
White-Sand	121	0.45	0.44	4.92	-0.02
Brown-Sand	88	0.49	0.52	6.62	0.07

Table 4.3. Total number of pollinator visits observed at white-sand and brown-sand ecotypes of *P. subserratum.*

	# Visits	
Insect Visitor	White-sand	Brown-sand
Bee Morph A	22	4
Bee Morph B	6	4
Bee Morph C	9	84
Bee Morph D	0	2
Bee Morph E	0	9
Bee Morph F	0	10
Bee Morph G	0	7
Bee Morph H	0	1
Bee Morph I	2	34
Bee Morph J	0	10
Bee Morph K	0	2
Green fly	0	8
Brown wasp	0	1
Black wasp	0	1
Lepidoptera	1	0

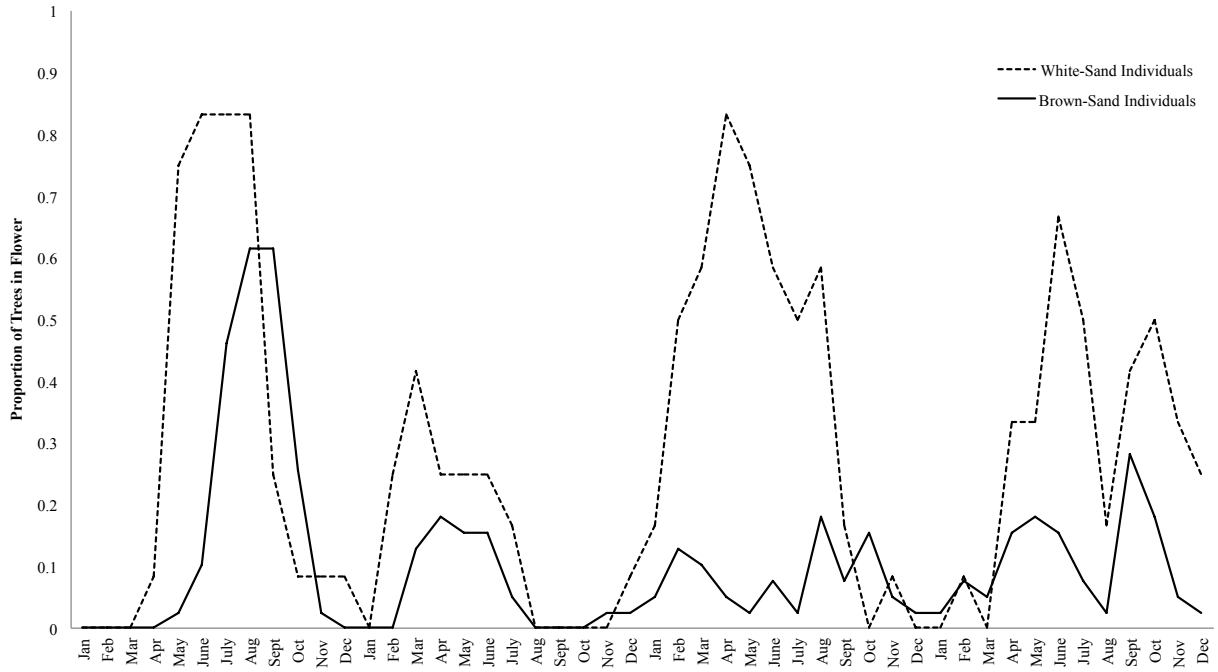

Figure 4.1. The proportion of white-sand and brown-sand individuals of *P. subserratum* in flower across time for January 2006 – December 2009.

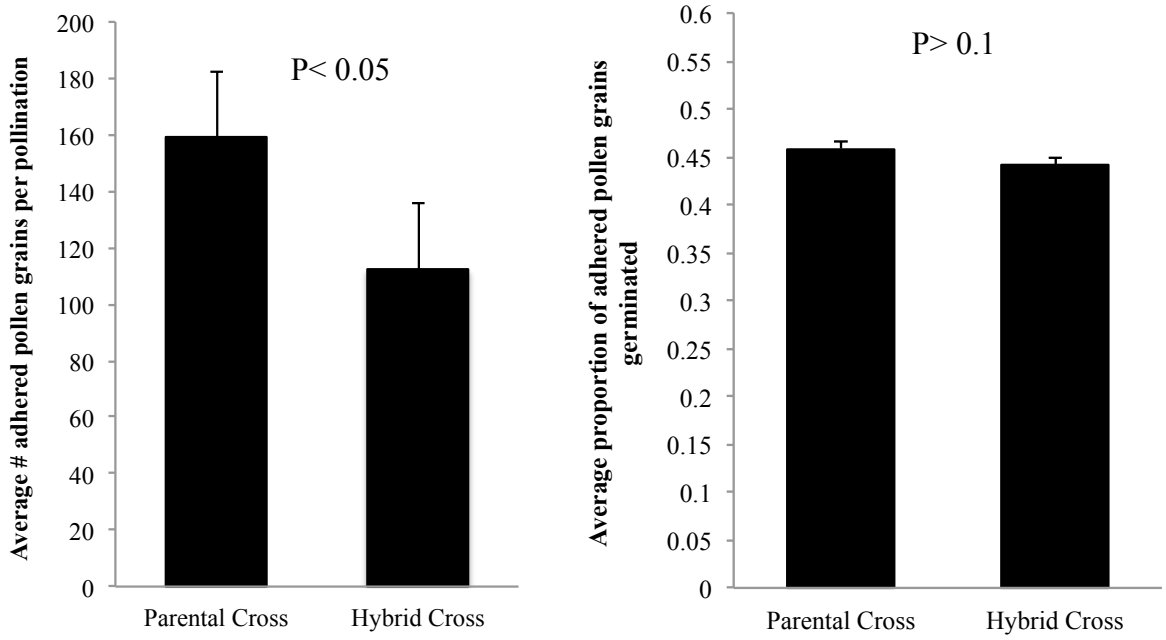

Figure 4.2. a. Average number of adhered pollen grains per pollination for parental crosses and hybrid hand-crosses. **b.** Average proportion of adhered pollen grains that germinated pollen tubes in parental and hybrid hand-crosses.

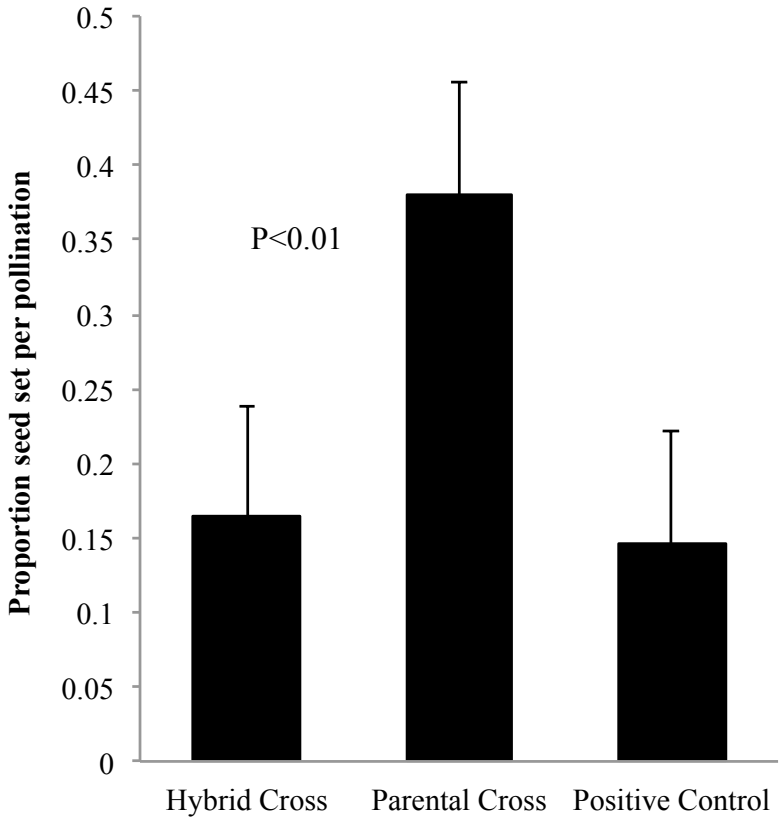

Figure 4.3. Proportion of hybrid hand-crosses and parental hand-crosses, which developed seeds two weeks after pollination.

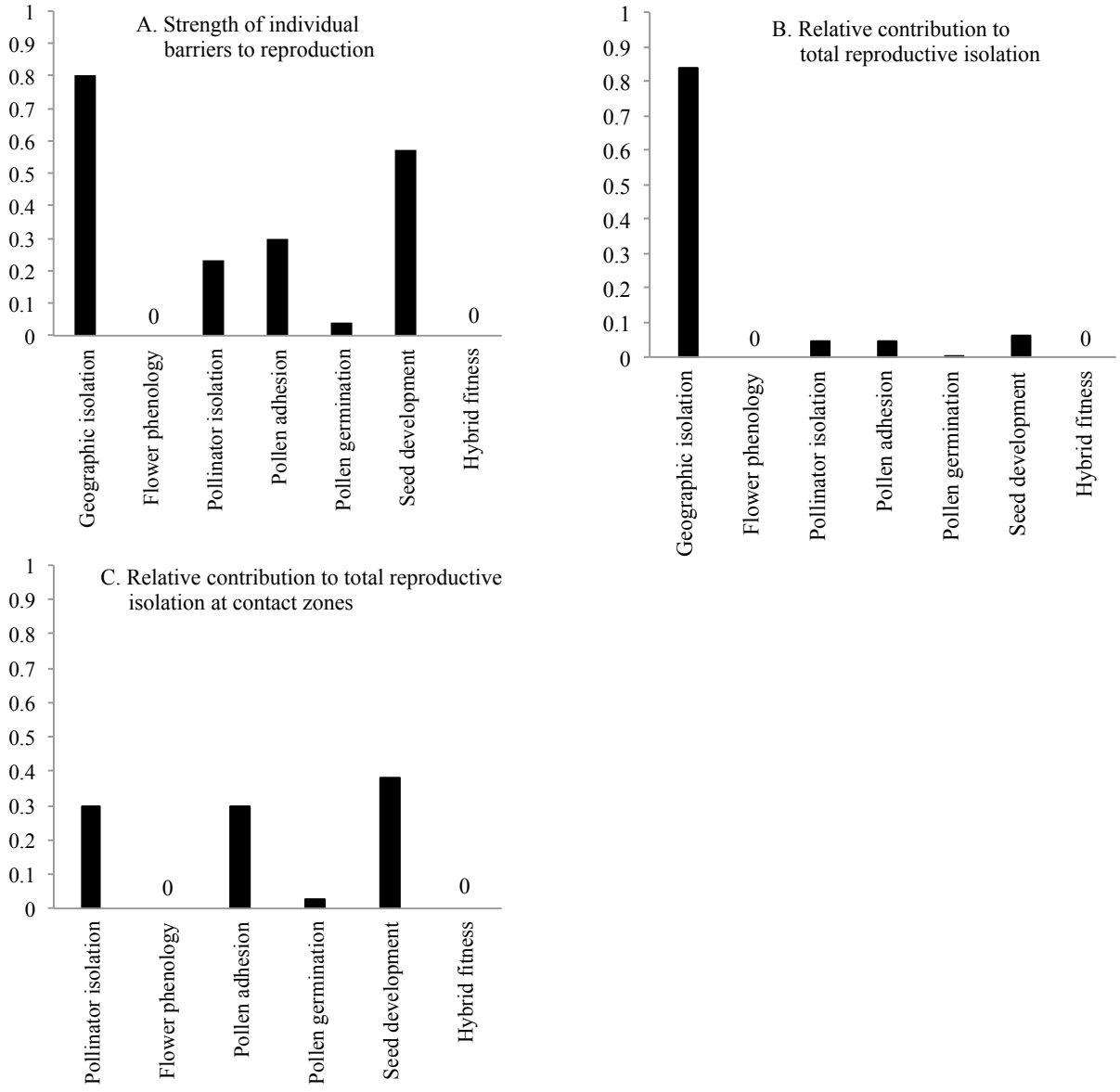

Figure 4.4. a. Strength of each individual barriers to reproduction (RI). **b.** Relative contribution of each barrier to reproduction to total reproductive isolation (T). **c.** Relative contribution of each barrier to reproduction to total reproductive isolation for individuals found at a contact zone.

WORK CITED

Adler, L. S, Seifert, M. G., Wink, M. and Morse, G. E. 2012. Reliance on pollinators predicts defensive chemistry across tobacco species. *Ecology Letters*. **15**:1140-1148.

Anderson, E. C., Thompson, E. A. 2002. A model-based method for identifying species hybrids using multilocus genetic data. *Genetics*, **160**: 1217-1229.

Bolnick, D. I. and Fitzpatrick, B. M. 2006. Sympatric Speciation: Models and empirical evidence. *Annu. Rev. Ecol. Evol. Syst.* **38**: 459-487.

Chari, J. and Wilson, P. 2001. Factors limiting hybridization between *Penstemon spectabilis* and *Penstemon centranthifolius*. *Canadian Journal of Botany*. **79**: 1439-1448.

Coyne, J. A. and Orr, H. A. 1989. Patterns of speciation in Drosophila. *Evolution* **43**: 362-381.

Coyne, J.A. 1992. Genetics and speciation. *Nature*. **355**: 511-515.

Coyne, J. A. and H. A. Orr 1997 "Patterns of speciation in Drosophila" Revisited. *Evolution*. **51**: 295-303.

Coyne, J. A. & Orr, H. A. 2004. Speciation. Sunderland, MA; Sinauer Associated, Inc.

Daly, D. C. 1987. A taxonomic revision of *Protium* (Burseraceae) in eastern Amazonia and the Guianas. Ph.D. diss., City University of New York, New York.

Daly, D. C., Fine, P. V. A. 2011. An ecologically significant new section of *Protium* (Burseraceae) from Amazonia. Studies in neotropical Burseraceae XVI. *Systematic Botany*, **36**: 939-949.

Dell'Olivo, A., Hoballah, M. E., Gubitz, T., and Kuhlemeier, C. 2011. Isolation barriers between *Petunia axillaris* and *Petunia integrifolia* (Solanaceae). *Evolution* **65**: 1979-1991.

Dettman, J. R., Sirjusingh, C., Kohn, L. M., and Anderson, J. B. 2007. Incipient speciation by divergent adaptation and antagonistic epistasis in yeast. *Nature*. **447**: 585-586.

Dobzhansky, T. 1937 Genetics and the origin of species. New York, NY Colombia University Press.

Dobzhansky, T., 1940. Speciation as a stage in evolutionary divergence. *Am. Nat.* **74**: 312-321.

Dobzhansky, T. 1951. Genetics and the Origin of Species, 3rd edition, Columbia University Press, New York, New York.

Evanno, G., Regnaut, S., Goudet, J. 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology,* **14**: 2611-2620.

Fine, P. V. A., Daly D. C., Munoz G. V., Mesones, I, Cameron. K. M. 2005. The contribution of edaphic heterogeneity to the evolution and diversity of Burseraceae trees in the western Amazon. *Evolution* **59**: 1464-1478.

Fine, P. V. A., Garcia-Villacorta, R., Pitman N. C. A., Mesones, I., Kembel, S. W. 2010. A floristic study of the white-sand forests of Peru. *Annals of the Missouri Botanical Garden*, **97**: 283-305.

Fine, P. V. A., Zapata F., Daly, D. C., Mesones, I., Misiewicz, T. M., Cooper, H. F. 2013a Phylogeography of edaphic specialist and generalist species of *Protium* (Burseraceae): the relative importance of geographic distance and environmental heterogeneity in generating phylogeographic structure across the Amazon basin. *Journal of Biogeography*, **40**, 646-661.

Fine, P. V. A., Metz, M. R., Lokvam, J., Mesones, I., Ayarza Zuniga, J. M., Lamarre, M., Vasquez Pilco, Baraloto, C. 2013b. Insect herbivores, chemical innovation and the evolution of habitat specialization in Amazonian trees. *Ecology*, **94**, 1764-1775.

Fine, P. V. A., Zapata, F., and Daly, D. C. In press. Investigating processes of Neotropical rain forest tree diversification by examining the evolution and historical biogeography of the Protieae (Burseraceae). *Evolution*.

Funk, D. J., Nosil, P and Etges, W. J. 2006. Ecological divergence exhibits consistently positive assoications with reproeuctive isolation across disparate taxa. *Proc. Natl. Acad. Sci. USA* **103**: 3209-3213.

Gavrilets, S. 2004. Fitness Landscapes and the Origin of Species. Princeton, New Jersey, Princeton University Press.

Glennon, K. L., Rissler, L. J., Church, S. A. 2012. Ecogeographic isolation: reproductive barrier between species and between cytotypes in *Houstonia* (Rubiaceae). *Evolutionary Ecology* **26**: 909-926.

Grant, V. 1949. Pollination systems as isolating mechanisms in angiosperms. *Evolution* **3**: 82-97.

Grant V. 1965. Evidence for the selective origin of incompatibility barriers in the leafy-stemmed *gilias*. *Proc Natl Acad Sci USA*. **54**:1567–1571.

Gentry, A. H. and Dodson, C. 1987. Contribution of nontrees to species richenss of a tropical rainforest. *Biotropica* **19**: 149-156.

Heizmann, P., Luu, D. T., and Dumas, C. 2000. Pollen-Stigma Adhesion in the Brassicaceae. *Annals of Botany* **85**: 23-27.

Hoffmann, V., Verboom, G. A., and Ellis, A. G. 2008. Components of reproductive isolation and their role in speciation between *Jamesbrittenia racemosa* and *J. pedunculosa* (Manuleae). *South African J. Bot.* **74**: 368-368.

Husband, B. C. and Sabara, H. A. 2004. Reproductive isolation between autotetraploids and their diploid progenitors in fireweed, *Chamerion angustifolium* (Onagraceae). *New Phytologist*. **161**: 703-713.

Kay, K. M. 2006. Reproductive isolation between two closely related hummingbird-pollinated neotropical gingers. *Evolution* **60**: 538-552.

Kessler, A., Halitschke, R., and Poveda, K. 2011. Herbivory-mediated pollinator limitation: negative impacts of induced volatiles on plant-pollinator interactions. *Ecology*. **92**: 1769-1780.

Lowry, D. B., Rockwood, R. C. and Willis J. H. 2008. Ecological reproductive isolation of coast and inland races of Mimulus gattatus. *Evolution*.

Mayr, E. 1942. Systematics and the origin of species. New York, NY, Columbia University Press.

Misiewicz, T. M., Barbosa, C. E. A., and Fine, P. V. A. 2012. Microsatellite primers for an Amazonian lowland tropical tree, *Protium subserratum* (Burseraceae). *American Journal of Botany Primer Notes and Protocols*. **99**:e465-e467.

Misiewicz, T. M. and P. V. A. Fine. In press. Evidence for ecological divergence across a mosaic of soil types in an Amazonian tropical tree: *Protium subserratum* (Burseraceae). *Molecular Ecology*.

Mittelbach, G. G., Schemske, D. W., Cornell, H. V., Allen, A. P., Brown, J. M., Bush, M. B., Harrison, S. P., Hurlbert, A. H., Knowlton, N., Lessios, H. A., McCain, C. M., McCune, A. R. McDade, L. A., McPeek, M. A., Near, T. J., Price, T. D., Ricklefs, R. E., Roy, K., Sax, D. F., Schluter, D., Sobel, J. M., and Turelli, M. 2007. Evolution and the latitudinal diversity gradient: speciation, extinction and biogeography. *Ecology Letters*. **10**: 315-331.

Morrison, D. A., McDonald, M., Bankoff, P., Quirico, P., and MacKay, D. 1994. Reproductive isolation mechanisms among 4 closely related species of *Conospermum* (Proteaceae). *Bot. J. Linn. Soc.* **116**: 13-31.

Nosil, P., Vines, T. H. and Funk, D. J. 2005. Perspective: reproductive isolation caused by natural selection against immigrants from divergent habitats. *Evolution* **59**, 705-719.

Pritchard, J. K., Stephens, P., Donnelly, P. 2000 Inference of population structure using multilocus genotype data. *Genetics*, **155**: 945-959.

Ramsey, J., Bradshaw Jr, H. D., and Schemske, D. W. 2003. Components of reproductive isolation between the monkeyflowers *Mimulus lewisii* and *M. cardinalis* (Phrymaceae). *Evolution* **57**: 1520-1534.

Rieseberg, L. H. and Willis J. H. 2007. Plant speciation. *Science* **317**: 910-914.

Sambatti, J. B. M., Ortiz-Barrientos, D., Baack, E. J., and Rieseberg, L. H. 2008. Ecological selection maintains cytonuclear incompatibilities in hybridizing sunflowers. *Ecology Letters*. **11**: 1082-1091.

Schiestl, F. P. and Schluter P. M. 2009. Floral isolation, specialized pollination, and pollinator behavior in orchids. *Annual Review of Entomology* **54**:425-446.

Schluter, D. 2001. Ecology and the origin of species. *Trends Ecol. Evol.* **16**: 372-380.

Scopece, G., Croce, A., Lexer, C., and Cozzolino, S. 2013. Components of reproductive isolation between *Orchis mascula* and *Orchis pauciflora*. *Evolution* **67**: 2083-2093.

Sobel, J. M., Chen, G. F., Watt, L. R., and Schemske, L. R. 2010. The biology of speciation. *Evolution* **64**: 295-315.

Whitehead, M. R. and Peakall, R. in press. Pollinator specificity drives strong prepollination reproductive isolation in sympatric sexually deceptive orchids. *Evolution*.

Widmer, A., Lexer, C. and Cozzolino, C. 2008. Evolution of reproductive isolation in plants. *Heredity* **102**: 31-38.