DIVERSIFICATION ON CONTINENTS AND ISLANDS: A HERPETELOGICAL PERSPECTIVE

by

Kyle Anthony O'Connell

DISSERTATION
Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Biology at
The University of Texas at Arlington
December, 2017

Arlington, Texas

Supervising Committee:

Matthew K. Fujita, Supervising Professor
Eric N. Smith
Todd A. Castoe
Esther Betrán
Matthew R. Walsh

ABSTRACT

DIVERSIFICATION ON CONTINENTS AND ISLANDS: A HERPETELOGICAL PERSPECTIVE

Kyle Anthony O'Connell, Ph.D

The University of Texas at Arlington, 2017

Supervising Professor: Matthew K. Fujita

This dissertation seeks to understand the geological and climatological processes that have promoted diversification on continental and island systems. Using molecular genetic data generated using Sanger sequencing and Next Generation Sequencing platforms, I conduct phylogenetic and biogeographic analyses, estimate gene flow, and conduct species delimitation. Using these analyses, I explore diversification processes on continents and islands using reptile and amphibian systems. In Chapter 2 I evaluate the role of geographical features to whipsnake diversification. Chapter 3 resolves the taxonomy of several poorly understood whipsnakes species and tests the effect of missing data on species delimitation. Chapter 4 investigates the biogeographical processes acting on parachuting frog diversification on the Sunda Shelf, specifically by quantifying the roles of within and between island diversification. Finally, Chapter 5 focuses on the processes that promoted *in situ* diversification on the island of Sumatra. I found that allopatric diversification is the predominant mode of diversification in whipsnakes and parachuting frogs, and that parachuting frogs diversified via *in situ* diversification on the islands of Sumatra and Borneo.

ACKNOWLEDGEMENTS

This dissertation work would not have been possible without the support hard work of all those who participated in specimen collection in Indonesia, including the Eric Smith lab, Michael B. Harvey, and all Indonesian participants from LIPI and Universitas Brawijaya. I am also thankful to all the museums who loaned me tissues for the *Masticophis* project, and to Ed Myers for the use of his sequences of western whipsnakes. I am thankful for the mentoring in both the lab and with writing given by Jeffrey Streicher, and especially for the mentoring in all things related to science given by Dan Portik, who make a big difference in my academic career. I am thankful for all the members past and present of the Fujita lab for their discussion and troubleshooting related to research questions and analyses, and to Jose Maldonado for always helping me in the lab. I could never have finished my research in Indonesia without the ideas given me from the Eric Smith lab, especially from Utpal Smart and also Elijah Wostl, who's realism tempered my enthusiasm, and who's persistence in insistence on using primary sources has strengthened my own research. I am thankful to Kathleen Currie for generating data as an undergraduate. This work was generously funded by the University of Texas at Arlington, the National Science Foundation Doctoral Dissertation Improvement Grant, the UTA Phi Sigma Society, the UTA Biology Department, and the American Museum of Natural History Theodore Roosevelt Grant. Finally, I am thankful to Eric Smith for his generosity with funding, field work, tissues, and project ideas, and to my supervising professor Matthew Fujita for his willingness to help, whether in the lab, writing letters, or reviewing my writing.

Dedication

This work is dedicated to my wife Katie who encourages me after every failure and never ceases to remind me of the big picture. I love you

TABLE OF CONTENTS

LIST OF FIGURES

Figure Page

LIST OF TABLES

Tables Page

CHAPTER 1

DIVERSIFICATION PRIMARILY OCCURS IN ALLOPATRY ACROSS GEOGRAPHICAL FEATURES

The process of diversification has intrigued and inspired biologists for over two centuries (Darwin, 1858). The most common mechanism that initiates diversification is that of allopatric distributions (Dobzhansky, 1940; Mayr, 1942; Coyne & Orr, 2004). In this case, speciation is preceded by the separation of populations by a geographical feature (Avise et al*.,* 1987; Coyne & Orr, 2004). This separation can occur via two processes: vicariance or dispersal (Kirkpatrick & Barton, 1997; Diamond, 1977). The difference between the two processes lies chiefly in the order of events, whether the barrier preceded the ancestral species, or whether the ancestral species preceeded the barrier. In the vicariant scenario, the formation of a barrier such as a river, mountain, or marine incursion divides an existing ancestor species into multiple populations (Zink et al*.,* 2000). In the dispersal scenario, a subsection of the ancestral species disperses across an existing barrier, creating multiple geographically isolated populations. Over time, drift, selection, or both forces will fix differences (genetic, phenotypic, or behavioral) that may lead to reproductive incompatibility, and as a result, new species (Futuyma & Mayr, 1980; Coyne & Orr, 2004).

However, the tempo of this processes can be strongly affected by the level of gene flow between the geographically separated populations. Gene flow is dependent on both biotic and abiotic factors (Futuyma & Mayr, 1980; Steeves et al*.,* 2005). The dispersal potential of the species, including niche specificity, strongly influences the level of gene flow (Zink et al*.,* 2001; Bell et al*.,* 2017). However, the age and permeability of the barrier also influences the level of

gene flow (Pyron & Burbrink, 2010). For example, most mountain chains are very long lasting, and will continue to isolate allopatric lineages well after their formation. In contrast, many marine incursions are ephemeral in nature, rising and falling relatively rapidly with changing climate. Thus, a barrier such as a marine incursion may have isolated lineages in the past, but today allows for secondary contact that may erase the signal of divergence (Jordan, 1905; Coyne & Orr, 2004).

The field of species delimitation attempts to identify independently evolving lineages to more accurately quantify biodiversity (De Queiroz, 2007; Fujita et al*.,* 2012; Petit & Excoffier, 2009). Researchers can delimiting species within an integrative framework by utilizing diverse data types, including phenotype, genotype, ecology, and geographic distributions (Camargo et al., 2012; Fujita et al., 2012; Yeates et al., 2011). In fact, methods have recently been developed to integrate multiple data types into a single analysis, theoretically reducing the bias introduced by relying on a single datatype (Pyron et al., 2016; Solís Lemus et al., 2015). Yet, as is well documented, an integrative approach that incorporates few loci with morphological or ecological data can present many challenges, and at times, lead to incorrect inferences (Bickford et al., 2007; Herrera & Shank, 2016). When different lines of evidence support different species models, it can be difficult to decide which analysis to prioritize. This commonly occurs in morphologically cryptic lineages, or young lineages (Hey, 2009; Knowles & Carstens, 2007). One possible solution to these challenges is to use of model-based methods that utilize species trees rather then gene trees, thus reducing investigator bias and accounting for incomplete lineage sorting. Bayes Factor Delimitation (BFD), first proposed by Grummer et al., (2013), and expanded to incorporate genomic data by Leaché et al., (2014; BFD*), does not require an a priori guide tree and estimates a species tree directly from biallelic markers. The method

estimates the marginal likelihood of each model (MLE) using path sampling, which can be compared between models to calculate a Bayes Factor (BF; Kass & Raftery, 1995). This allows investigators to objectively rank models of species relationships.

Building upon this theoretical background, this dissertation investigates the role of allopatric diversification and conducts species delimitation of taxa at two geographic scales: the continent of North America, and the island system of the Sunda Shelf (Sundaland). Within these two geographic contexts, I leverage two herpetological study systems, North American whipsnakes of the genus *Masticophis*, and Southeast Asian parachuting frogs of the genus *Rhacophorus*. Chapter 2 conducts comparative phylogeography of whipsnakes and asks questions about the role of geographical features to diversification. In addition, Chapter 2 quantifies gene flow to assess the isolating potential of each geographical feature. Chapter 3 is a natural extention of Chapter 2, but focuses long-standing taxonomic issues within whipsnakes. Using integrative species delimitation, I evaluate the species status of several subspecies within the genus, and test the effect of missing data on coalescent species delimitation. Chapters 4 and 5 shift their focus to the second study system: Sundaland *Rhacophorus*. Chapter 3 focuses on the biogeographical patterns exhibited by the whole genus across Asia, and quantifies the roles of within and between island diversification. I also probe the relationship between within-island diversification and species richness and endemism. Finally, Chapter 5 focuses on the processes that promoted extensive within-island diversification on the island of Sumatra. I find that several highland species show congruent phylogeographic structure across the island, but investigate if these divergence events are synchronous, suggesting common causality, or asynchronous, suggesting pseudocongruence.

Each chapter includes several co-authors who contributed to these works. In Chapter 2, Jeffrey Streicher helped with laboratory work, and with writing and editing the paper. Eric Smith contributed ideas and tissues to the project, and Matthew Fujita helped to write and edit the paper, and guided much of the paper's direction. In Chapter 3, Eric Smith aided with morphological work and made taxonomic recommendations. In Chapter 4, Utpal Smart conducted the biogeographic analysis, guided the divergence dating, and helped to edit the manuscript. Eric Smith contributed ideas to the project and collected tissues. Amir Hamidy and Nia Kurniawan aided with field work. Matthew Fujita helped to write the paper and contributed to the discussion about island biogeography. In Chapter 5, Amir Hamidy and Nia Kurniawan aided with field work and permitting. Eric Smith contributed tissues and ideas regarding the geographical barriers on Sumatra. Matthew Fujita helped edit the paper.

REFERENCES

Avise JC, Arnold J, Ball RM, Bermingham E, Lamb T, Neigel JE, Reeb CA, Saunders NC (1987) Intraspecific phylogeography: the mitochondrial DNA bridge between population genetics and systematics. *Annual Review of Ecology and Systematics*, 18, 489-522.

Bickford, D., Lohman, D. J., Sodhi, N. S., Ng, P. K., Meier, R., Winker, K., Ingram, K. K., Das I (2007) Cryptic species as a window on diversity and conservation. *Trends in Ecology & Evolution*, 22, 148–155.

Bell RC, Parra JL, Badjedjea G, Barej MF, Blackburn DC, Burger M, ... & Kielgast J. (2017) Idiosyncratic responses to climate-driven forest fragmentation and marine incursions in reed frogs from Central Africa and the Gulf of Guinea Islands. *Molecular Ecology*, 26, 5223–5244.

Camargo A, Morando M, Avila LJ, Sites JW (2012) Species delimitation with ABC and other coalescent-based methods: a test of accuracy with simulations and an empirical example with lizards of the *Liolaemus darwinii* complex (Squamata: Liolaemidae). *Evolution*, 66, 2834–2849.

Coyne JA, Orr HA (2004) Speciation. Sinauer Associates, Sunderland.

Darwin C (1858). On the Origin of Species by Means of Natural Selection, Or, the Preservation of Favoured Races in the Struggle for Life. London: J. Murray.

De Queiroz K (2007). Species concepts and species delimitation. *Systematic Biology*, 56, 879–886.

Diamond JM (1977) Continental and insular speciation in pacific island birds. *Systematic Zoology*, 26, 263-268.

Dobzhansky T (1940) Speciation as a stage in evolutionary divergence. *American Naturalist,* 74, 302-321.

Fujita MK, Leaché AD, Burbrink FT, McGuire JA, Moritz C (2012). Coalescent-based species delimitation in an integrative taxonomy. *Trends in Ecology & Evolution*, 27, 480–488.

Futuyma DJ, Mayer GC (1980) Non-allopatric speciation in animals. *Systematic Biology*, 29, 254-271.

Grummer JA, Bryson RW, Reeder TW (2014) Species delimitation using Bayes factors: simulations and application to the Sceloporus scalaris species group (Squamata: Phrynosomatidae). *Systematic Biology*, 63, 119–133.

Herrera S, Shank TM (2016) RAD sequencing enables unprecedented phylogenetic resolution and objective species delimitation in recalcitrant divergent taxa. *Molecular Phylogenetics and Evolution*, 100, 70–79.

Hey J (2009) On the arbitrary identification of real species. In: Butlin RK, Bridle JR, Schulter D, eds. *Speciation and Patterns of Diversity.* Cambridge: Cambridge University Press, 15–28.

Hotaling S, Foley ME, Lawrence NM, Bocanegra J, Blanco MB, Rasoloarison R, Kappeler PM, Barret MA, Yoder AD, Weisrock DW (2016) Species discovery and validation in a cryptic radiation of endangered primates: coalescent-based species delimitation in Madagascar's mouse lemurs. *Molecular Ecology*, 25, 2029-2045.

Jordan DS (1905) The origin of species through isolation. *Science*, 22, 545-562.

Kass RE, Raftery AE (1995) Bayes factors. *Journal of the American Statistical Association,* 90, 773–795.

Kirkpatrick M, Barton NH (1997) Evolution of a species' range. *The American Naturalist*, 150, 1–23.

Knowles, LL, Carstens, BC (2007) Delimiting species without monophyletic gene trees. *Systematic Biology*, 56, 887–895.

Leaché A, Fujita M, Minin V, Bouckaert R (2013) Species Delimitation using Genome-Wide SNP Data. *Systematic Biology*, 63, 534-542.

Mayr E (1942) Systematics and the Origin of Species. Columbia University Press, New York.

Petit RJ, Excoffier L (2009) Gene flow and species delimitation. *Trends in Ecology & Evolution*, 24, 386–393.

Pyron RA, Burbrink FT (2010). Hard and soft allopatry: physically and ecologically mediated modes of geographic speciation. *Journal of Biogeography*, 37, 2005-2015.

Pyron RA, Hsieh FW, Lemmon AR, Lemmon EM, Hendry CR (2016) Integrating phylogenomic and morphological data to assess candidate species-delimitation models in brown and red-bellied snakes (Storeria). *Zoological Journal of the Linnean Society*, 177, 937-949.

Solís Lemus C, Knowles LL, Ané C (2015) Bayesian species delimitation combining multiple genes and traits in a unified framework. *Evolution*, 69, 492–507.

Steeves TE, Anderson DJ, Friesen VL (2005) The Isthmus of Panama: a major physical barrier to gene flow in a highly mobile pantropical seabird. *Journal of Evolutionary Biology*, 18, 1000–1008.

Yeates DK, Seago A, Nelson L, Cameron SL, Joseph L, Trueman JW (2011) Integrative taxonomy, or iterative taxonomy? *Systematic Entomology*, 36, 209–217.

Zink RM, Blackwell-Rago RC, Ronquist F (2000) The shifting roles of dispersal and vicariance in biogeography. *Proceedings of the Royal Society of London B: Biological Sciences*, 267, 497-503.

Zink RM, Kessen AE, Line TV, Blackwell-Rago RC (2001) Comparative phylogeography of some aridland bird species. *Condor*, 103, 1-10.

GEOGRAPHICAL FEATURES ARE THE PREDOMINANT DRIVER OF MOLECULAR DIVERSIFICATION IN WIDELY DISTRIBUTED NORTH AMERICAN WHIPSNAKES[1]

Kyle A. O'Connell[1,2], Jeffrey W. Streicher[3], Eric N. Smith[1,2], Matthew K. Fujita[1,2]

[1]*Department of Biology, The University of Texas at Arlington, Arlington, Texas 76019, USA.*
[2]*The Amphibian and Reptile Diversity Research Center; University of Texas at Arlington, Arlington, Texas 76010, USA.*
[3]*Department of Life Sciences; Natural History Museum, London, UK*

Citation: O'Connell, KA, Streicher, JW, Smith, EN, Fujita, MK (2017). Geographical features are the predominant driver of molecular diversification in widely distributed North American whipsnakes. *Molecular Ecology,* 26: 5729–5751.

[1] Used with permission of the publisher, 2017

CHAPTER 2

GEOGRAPHICAL FEATURES ARE THE PREDOMINANT DRIVER OF MOLECULAR DIVERSIFICATION IN WIDELY DISTRIBUTED NORTH AMERICAN WHIPSNAKES

ABSTRACT

Allopatric divergence following the formation of geographical features has been implicated as a major driver of evolutionary diversification. Widespread species complexes provide opportunities to examine allopatric divergence across varying degrees of isolation in both time and space. In North America, several geographical features may play such a role in diversification, including the Mississippi River, Pecos River, Rocky Mountains, Cochise Filter Barrier, Gulf of California, and Isthmus of Tehuantepec. We used thousands of nuclear single nucleotide polymorphisms (SNPs) and mitochondrial DNA from several species of whipsnakes (genera *Masticophis* and *Coluber*) distributed across North and Central America to investigate the role that these geographical features have played on lineage divergence. We hypothesize that these features restrict gene flow and separate whipsnakes into diagnosable genomic clusters. We performed genomic clustering and phylogenetic reconstructions at the species and population levels using Bayesian and likelihood analyses, and quantified migration levels across geographical features to assess the degree of genetic isolation due to allopatry. Our analyses suggest that (i) major genetic divisions are often consistent with isolation by geographical features, (ii) migration rates between clusters are asymmetrical across major geographical features, and (iii) areas that receive proportionally more migrants possess higher levels of genetic diversity. Collectively, our findings suggest that multiple features of the North American landscape contributed to allopatric divergence in this widely-distributed snake group.

INTRODUCTION

Divergence in allopatry has long been considered the most common model of diversification (Dobzhansky, 1940; Mayr, 1942; Coyne & Orr, 2004; Zink, 2014). The concordance of species' boundaries with geographical features provides the strongest evidence for allopatric differentiation (Avise et al., 1987; Coyne & Orr, 2004). Dispersal to new areas or the formation of physical barriers isolates populations (Kirkpatrick & Barton, 1997; Diamond, 1977) and can lead to significant reductions in gene flow, thus promoting lineage divergence (Futuyma & Mayr, 1980). However, the genetic signal of previous isolation can be masked by gene flow and recombination at secondary contact. Recently diverged populations experiencing secondary contact can form hybrid zones, indicating that either a barrier no longer exists, such as glaciers, or that a barrier is permeable, such as non-continuous mountain ranges (Jordan, 1905; Coyne & Orr, 2004; Feder et al., 2012). Thus, evidence of hybrid zones can support a scenario where historical barriers led to temporarily isolated populations (e.g. Pleistocene glacial refugia in North America). However, complete barriers to gene flow can isolate populations permanently, leading to reproductive isolation (Pyron & Burbrink, 2010). Studying species at different temporal intervals in this process can help us understand the influence of such geographical features on limiting gene flow, and how barriers contribute to species diversification. The age and permeability of these features often determines the level of genetic differentiation that occurs between isolated populations (Pyron & Burbrink 2010).

Rivers, mountains, and geographical depressions have played important roles in the diversification of North American biota, including plants, invertebrates, and vertebrates (Fig. 2.1, Table 2.1). Seven geographical features correlate with divergence of multiple taxa across the continental United States, Mexico, and Central America. In the eastern United States a consistent

faunal break is found at the Mississippi River (MR). In the western United States, the Pecos River (PR; dividing the Chihuahuan Desert and central plains), the Cochise Filter Barrier (CFB; the division between the Chihuahuan and Sonoran Deserts), the Rocky Mountains (RM), and the Gulf of California (GC) have been identified as barriers that likely influenced the evolution of multiple plant and animal species. In Mexico, the Isthmus of Tehuantepec (IT) has been identified as an influential barrier (see citations in Table 2.1). In this study we investigate how these geographical features have promoted diversification of eight widely distributed snake species (genera *Masticophis* and *Coluber).*

Table 2.1: Organisms that support the role that focal geographical features have played in facilitating divergence in allopatry.

Feature	Species	Common Name	Evidence	Reference
Mississippi River	*Erimystax dissimilis*	Streamline chub	Intraspecific mtDNA	Strange & Burr 1997
	Kinosternon subrubrum	Eastern mud turtle	Intraspecific mtDNA	Walker et al., 1998
	Deirochelys reticularia	Chicken turtle	Intraspecific mtDNA	Walker & Avise 1998
	Chelydra serpentina	Common snapping turtle	Intraspecific mtDNA	Walker & Avise 1998
	Pantherophis obsoletus	Black rat snake	Intraspecific mtDNA	Burbrink et al., 2000
	Apalone mutica	Smooth softshell turtle	Intraspecific mtDNA	Weisrock & Janzen 2000
	Percina evides	Gilt darter	Intraspecific mtDNA	Near et al., 2001
	Pinus taeda	Loblolly pine	Intraspecific range limits	Al-Rabab'ah & Williams 2002; Eckert *et al.,* 2010
	Sceloporus undulatus	Eastern fence lizard	Intraspecific mtDNA	Leaché & Reeder 2002
	Blarina carolinensis	Southern short-tailed shrew	Intraspecific mtDNA	Brant & Orti 2002
	Pantherophis guttatus	Rat snake	Intraspecific mtDNA	Burbrink 2002

	Anaxyrus fowleri	Fowler's toad	Intraspecific mtDNA	Masta et al., 2002
	Anaxyrus woodhousii/A. americanus	North American toads	Interspecific range boundary	Masta et al., 2002
	Blarina brevicauda	Short-tailed shrew	Intraspecific mtDNA	Brant & Orti 2003
	Ambystoma maculatum	Spotted salamander	Intraspecific mtDNA	Zamudio & Savage 2003
	Lithobates pipiens	Northern leopard frog	Intraspecific mtDNA	Hoffman & Blouin 2004
	Lithobates catesbeiana	Bullfrog	Intraspecific mtDNA	Austin et al., 2004
	Pseudacris crucifer	Spring peeper	Intraspecific mtDNA	Austin et al., 2004
	Pseudacris nigrita	Southern chorus frog	Intraspecific mtDNA	Moriarty & Cannatella 2004
	Juglans nigra	Black walnut	Intraspecific cpDNA	Soltis et al., 2006
	Eumeces fasciatus	Five-lined skink	Intraspecific microsats	Howes et al., 2006
	Etheostoma caeruleum	Rainbow darter	Intraspecific mtDNA	Ray et al., 2006
	Pseudacris spp.	Trilling chorus frogs	Interspecific mtDNA	Lemmon et al., 2007; 2008
	Procyon lotor	Raccoon	Intraspecific mtDNA	Cullingham et al., 2008
	Acris spp.	Cricket frogs	Interspecific mtDNA, nuDNA	Gamble et al., 2008
	Coluber constrictor	North American racer	Instraspecific mtDNA	Burbrink et al., 2008
	Lampropeltis getula	Common kingsnake	Intraspecific mtDNA	Pyron & Burbrink 2009
	Aphonopelma hentzi	Texas brown tarantula	Intraspecific mtDNA, eastern range limit	Hamilton et al., 2011
	Mephitis mephitis	Striped skunk	Intraspecific mtDNA	Barton & Wisely 2012
	Cryptobranchus alleganiensis	Eastern hellbender	Intraspecific microsats	Unger et al., 2013
	Micrurus spp.	Eastern/Texas coralsnakes	Interspecific microsats, mtDNA, SNPs	Castoe et al., 2012; Streicher et al., 2016
	Campanulastrum americanum	American bellflower	Interspecific mtDNA, SNPs	Bernard-Kubow et al., 2015
Pecos River	*Peromyscus maniculatus*	Deer Mouse	Interspecific mtDNA	Lansman et al., 1983

	Onychomys spp.	Grasshopper Mice	Interspecific mtDNA	Riddle & Honeycutt 1990
	Chaetodipus penicillatus	Desert pocket mouse	Western range limit	Lee et al., 1996
	Peromyscus eremicus	Cactus Mouse	Western range limit	Wadpole et al., 1997
	Pituophis catenifer	Bullsnake	Interspecific mtDNA	Rodríguez-Robles & De Jesús-Escobar 2000; Myers et al., 2017
	Sceloporus undulatus	Eastern fence lizard	Intraspecific mtDNA	Leaché & Reeder 2002
	Sceloporus magister	Desert spiny lizard	Eastern range limit	Leaché & Mulcahy 2007
	Diadophis punctatus	Ring-necked snake	Intraspecific mtDNA	Fontanella et al., 2008
	Acris blanchardi	Blanchard's cricket frog	Western range limit	Gamble et al., 2008
	Sceloporus cowlesi	White sands prairie lizard	Interspecific mtDNA, nuDNA	Leaché 2009
	Sceloporus consobrinus	Southern prairie lizard	Interspecific mtDNA, nuDNA	Leaché 2009
	Crotalus atrox	Western diamondback rattlesnake	Intraspecific SNPs	Schield et al., 2015
	Rhinocheilus lecontei	Long-nosed snake	Interspecific mtDNA	Myers et al., 2017
	Arizona elegans	Glossy snake	Intraspecific mtDNA	Myers et al., 2017
Rocky Mountains	*Xerobates agassizii*	Desert tortoise	Intraspecific mtDNA	Lamb et al., 1989
	Peromyscus sp.	Deer mouse	Interspecific mtDNA	Riddle et al., 2000
	Crotalus viridis	Prairie rattlesnake	Intraspecific mtDNA	Pook et al., 2000
	Sceloporus undulatus	Eastern fence lizard	Intraspecific mtDNA	Leaché & Reeder 2002
Cochise Filter Barrier	*Onychomys* spp.	Grasshopper mice	Interspecific mtDNA	Riddle & Honeycutt 1990
	Chaetodipus intermedius	Red pocket mouse	Intraspecific mtDNA	Riddle 1995

Chaetodipus penicillatus	Desert pocket mouse	Intraspecific mtDNA	Riddle 1995; Lee et al., 1996
Sciurus aberti	Tassel-eared squirrel	Intraspecific mtDNA	Lamb et al., 1997
Peromyscus eremicus	Cactus mouse	Intraspecific mtDNA	Wadpole et al., 1997
Gambelia wislizenii	Long-nosed leopard lizard	Intraspecific mtDNA	Orange et al., 1999
Corvus corax	Common raven	Interspecific mtDNA	Omland et al., 2000
Crotalus viridis	Prairie rattlesnake	Interspecific mtDNA	Ashton & de Queiroz 2001
Toxostoma curvirostre	Curve-billed thrasher	Intraspecific mtDNA	Zink et al., 2001
Pipilo fuscus	Canyon towhee	Intraspecific mtDNA	Zink et al., 2001
Kinosternon flavescens	Yellow mud turtle	Intraspecific mtDNA	Serb et al., 2001
Lophocereus schottii	Senita cactus	Eastern range limit	Nason et al., 2002
Myotis spp.	Vesper bats	Interspecific mtDNA	Rodriguez & Ammerman 2004
Phrynosoma cornutum	Texas horned lizard	Interspecific mtDNA	Rosenthal & Forstner 2004
Rhinocheilus lecontei	Long-nosed snake	Interspecific mtDNA	Rosenthal & Forstner 2004; Myers et al., 2017
Bufo punctatus	Red-spotted toad	Intraspecific mtDNA	Jaeger et al., 2005
Moneilema appressum	Longhorn cactus beetle	Intraspecific mtDNA	Smith & Farrell 2005
Phrynosoma spp.	Horned lizards	Interspecific mtDNA, nuDNA	Leaché & McGuire 2006
Crotalus atrox	Western diamondback rattlesnake	Intraspecific mtDNA; SNPs	Castoe et al., 2007; Schield et al., 2015
Sceloporus magister	Desert spiny lizard	Eastern range limit	Leaché & Mulcahy 2007
Hypsiglena torquata	North American nightsnake	Interspecific mtDNA	Mulcahy 2008; Mucahy & Macey 2009; Myers et al., 2017

	Thomomys spp.	Gophers	Interspecific nuDNA	Belfiore et al., 2008
	Dilophotopsis spp.	Velvet ant	Interspecific mtDNA	Wilson & Pitts 2008; 2010b
	Odocoileus hemionus	American mule deer	Interspecific mtDNA	Latch et al., 2009
	Lampropeltis getula	Common kingsnake	Intraspecific mtDNA	Pyron & Burbrink 2009; Myers et al., 2017
	Melampodium leucanthum	Blackfoot daisy	AFLP, cpDNA	Rebernig et al., 2010
	Pituophis catenifer	Bullsnake	Intraspecific mtDNA	Bryson et al., 2011; Myers et al., 2017
	Gastrophryne spp.	Great Plains narrowmouth toads	Interspecific mtDNA	Streicher et al., 2012
	Crotalus molossus	Northern black-tailed rattlesnake	Intraspecific mtDNA	Anderson & Greenbaum 2012; Myers et al., 2017
	Pseudouroctonus minimus	Vaejovid scorpion	Intraspecific mtDNA	Bryson et al., 2013
	Ammospermophilus spp.	Antelope squirrels	Interspecific mtDNA, nuDNA	Mantooth et al., 2013
	Arizona elegans	Glossy snake	Intraspecific mtDNA	Myers et al., 2017
	Thamnophis marcianus	Checkered garter snake	Intraspecific mtDNA	Myers et al., 2017
	Salvadora hexalepis	Western patch-nosed snake	Intraspecific mtDNA	Myers et al., 2017
	Masticophis flagellum	Western whipsnake	Intraspecific mtDNA	Myers et al., 2017
	Pituophis catenifer	Bullsnake	Intraspecific mtDNA	Myers et al., 2017
Gulf of California	*Thomomys bottae*	Pocket gopher	Interspecific mtDNA	Smith 1998
	Polioptila spp.	Gnatcatcher	Interspecific mtDNA	Zink & Blackwell 1998
	Urosaurus spp.	Collared lizard	Intraspecific isozymes	Aguirre et al., 1999
	Quercos spp.	Oaks	Interspecific cpDNA	Manos et al., 1999
	Peromyscus spp.	Deer mouse	Interspecific mtDNA	Riddle et al., 2000
	Pituophis catenifer	Bullsnake	Interspecific mtDNA	Rodríguez-Robles & De Jesús-Escobar 2000

	Lophocereus spp.	Senita cactus	Interspecific cpDNA	Nason et al., 2002
	Ammospermophilus leucurus	Antelope ground squirrel	Intraspecific isozymes	Whorley et al., 2004
	Xantusia spp.	Night lizards	Interspecific mtDNA	Sinclair et al.,
	Phrynosoma mcallii	Flat-tailed horned lizard	Intraspecific mtDNA	Mulcahy et al., 2006
	Homalonychus sp.	Spider	Intraspecific mtDNA	Crews & Hedin 2006
	Trimorphodon biscutatus	Western lyresnake	Intraspecific mtDNA	Devitt 2006
	Sceloporus magister	Desert spiny lizard	Eastern range limit	Leaché & Mulcahy 2007
	Hypsiglena spp.	Nightsnakes	Interspecific mtDNA	Mulcahy & Macey 2009
	Odocoileus hemionus	American mule deer	Interspecific mtDNA	Latch et al., 2009
	Crotalus atrox	Western diamondback rattlesnake	Intraspecific mtDNA	Castoe et al., 2007
	Pseudouroctonus minimus	Vaejovid scorpion	Intraspecific mtDNA	Bryson et al., 2013
	Arizona elegans	Glossy snake	Intraspecific mtDNA	Myers et al., 2017
Ithmus of Tehuantepec	*Peromyscus aztecus*	Aztec mouse	Intraspecific mtDNA	Sullivan et al., 1996
	Abronia spp.	Alligator lizards	Interspecific mtDNA	Chippindale et al., 1998
	Reithrodontomys sumichrasti	Sumichrast's harvest mouse	Intraspecific mtDNA	Sullivan et al., 2000
	Habromys lophurus	Crested tailed deer mouse	Interspecific ranges	Carleton et al., 2002
	Bufo punctatus	Red-spotted toad	Intraspecific mtDNA	Mulcahy et al., 2006
	Alouatta pigra	Black howler monkey	Upper range boundary	Baumgarten & Williamson 2007
	Lampornis amethystinus	Amethyst-throated hummingbird	Interspecific mtDNA	Cortés-Rodríguez et al., 2008
	Habromys spp.	Deer mouse	Interspecific mtDNA	León-Paniagua et al., 2007
	Atropoides spp.	Jumping pitvipers	Interspecific mtDNA	Castoe et al., 2008

Cerrophidion spp.	Montane pitvipers	Interspecific mtDNA	Castoe et al., 2008
Campylopterus curvipennis	Wedge-tailed sabrewing	Intraspecific mtDNA, microsats	González et al., 2011
Palicourea padifolia	Distylous shrub	Intraspecific cpDNA	Gutiérrez-Rodríguez et al., 2011
Pituophis lineaticollis	Gopher snake	Interspecific mtDNA	Bryson et al., 2011
Aphelocoma spp.	Scrub jays	Interspecific mtDNA	McCormack et al., 2011
Bolitoglossa spp.	Tropical salamanders	Interspecific mtDNA, nuDNA	Rovito et al., 2012
Bombus ephippiatus	Polymorphic bumble bee	Intraspecific mtDNA, nuDNA	Duennes et al., 2012
Dermatemys mawii	Central American river turtle	Intraspecific mtDNA	González-Porter et al., 2013
Amazilia cyanocephala	Azure-crowned hummingbird	Intraspecific mtDNA	Rodríguez-Gómez et al., 2013
Boa constrictor	Boa constrictor	Intraspecific mtDNA, nuDNA, microsats	Suárez-Atilano et al., 2014
Aegolius acadicus	Northern Saw-whet owl	Southern range boundary	Withrow et al., 2014
Rhipsalis baccifera	Mistletoe cactus	Intraspecific cpDNA, nuDNA	Ornelas et al., 2015
Eugenes fulgens	Magnificent hummingbird	Interspecific mtDNA, nuDNA	Zamudio-Beltrán & Hernández-Baños 2015
Phaseolus vulgaris	Common bean	Intraspecific SNPs	Rodriguez et al., 2016

Both biotic and abiotic factors regulate levels of gene flow that occur across discrete geographical features (Futuyma & Mayr, 1980; Steeves et al., 2005). First, biotic factors such as a species' dispersal potential and ecological tolerance influence how often a species can cross a geographical feature (Pyron & Burbrink, 2010). Often, larger animals are more capable of dispersing larger distances (Sutherland et al., 2000). The abiotic factors intrinsic to the geographical feature also determine how much gene flow can occur, and thus the level of

population differentiation. The abiotic isolating potential of a feature is influenced by three factors. First, the age of the feature determines how long isolation has taken place, and thus the level of differentiation between populations. Second, the permeability of the feature to gene flow affects genetic divergence of allopatric populations (hard and soft barriers; Pyron & Burbrink, 2010). Finally, the intrinsic composition of a feature also influences its isolating potential. For example, rivers and mountains may isolate species differently, and some historically hard barriers today allow limited gene flow. Habitat contractions associated with Pleistocene glaciation, and the once flooded IT in Mexico would be two examples of features that once isolated populations, but today only leave an eroding signal of isolation. On the other hand, ancient features such as the MR and the RM have isolated populations since their formation, although we expect to see evidence for greater levels of historical gene flow as the features were newly formed, with low levels of contemporary gene flow (Burbrink et al., 2008; Egge & Hagbo, 2015). These factors add additional complications to hypotheses about the level of divergence and gene flow observed between isolated populations, because an ancient, yet permeable barrier may allow greater gene flow than a younger yet less permeable barrier. Additionally, while a feature such as the CFB may isolate less vagile animals (Table 2.1), the high vagility of birds has allowed many species to migrate across it (Zink et al., 2001).

A broad geographic distribution, high potential for dispersal, and high species diversity make colubrid snakes ideal models to test hypotheses of diversification because they provide natural replicates to test hypotheses across distinct geographical features (Hirth et al., 1969; Conant & Collins, 1998; Dodd & Barichivich, 2007; Burbrink et al., 2008; Halstead et al., 2009). Whipsnakes (genera *Masticophis* and *Coluber*) are a group of colubrid snakes distributed throughout North and South America spanning several important geographical features (Conant

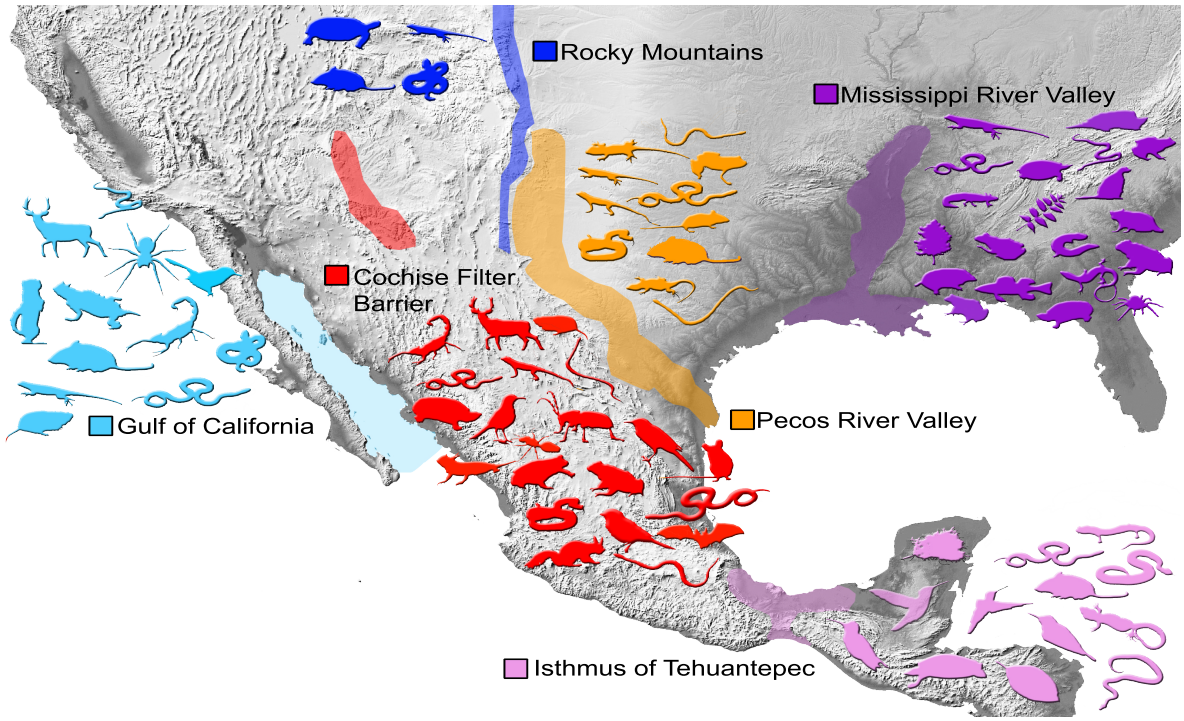

Fig. 2.1: The major geographic features discussed in this study are highlighted. Next to each feature are representatives of species with allopatric divisions at these features. References are found in Table 2.1. Age of origin is also listed next to each feature. We used the estimated date of origin for each feature to encompass its complete history. We found these dates from literature searches from the following sources: Mississippi River (Arthur & Taylor, 1998), Pecos River (Havenor, 2003), Cochise Filter Barrier (Devitt et al., 2006; and citations therein), Rocky Mountains (Riddle et al., 2006), Gulf of California (Lonsdale, 1991), Isthmus of Tehuantepec (Barrier et al., 1998).

& Collins, 1998; Utiger et al., 2005; Pyron et al., 2011). In this study, we investigate the role that barriers to dispersal have played in divergence across several widely distributed whipsnake species. Using a restriction site associated DNA sequencing (RADseq) dataset, we pursue the following questions: 1) How is genetic diversity partitioned within species across the landscape? 2) Does migration occur between populations across geographical features? We find that at least six geographical features are associated with allopatric units in whipsnakes, including the MR, the PR, the CFB, the RM, the GC, and the IT. Using comparisons that involved several species, we find evidence for asymmetric rates of migration from east to west across the MR, the PR, and the IT and from west to east across the RM and CFB. More extensive geographic sampling of mitochondrial DNA revealed corroborating evidence for many of the patterns observed in the nuclear dataset, and also several instances of intraspecific allopatric circumscription.

Collectively, these results suggest that divergence in allopatry is the predominant form of evolution among whipsnake species.

MATERIALS AND METHODS

Study system and sampling

Whipsnakes (Colubridae: Colubroidea) are large, typically diurnal, slender and active snakes that occur throughout North America and into northern South America (Dodd & Barichivich, 2007). For decades, most taxonomists placed all whipsnakes, excluding *Coluber constrictor*, in the genus *Masticophis* (Ortenburger, 1923), until Utiger et al*.,* (2005) used molecular data to demonstrate that *C. constrictor* was nested within two *Masticophis flagellum* samples. Recently, Pyron et al., (2013), and Burbrink and Myers, (2015) provided additional support for this arrangement when they found that *C. constrictor* was nested among samples of *Masticophis* species. Thus, until recently, most authorities recognized *Masticophis* as a junior synonym of *Coluber* (Uetz & Hošek, 2016). However, Myers et al*.,* (*In Press*) recovered a monophyletic *Masticophis*, and recommend the distinguishing *Masticophis* species from *C. constrictor*. In this study we use the term whipsnakes to include species pertaining to both *Masticophis* and *Coluber*. Previous work on whipsnakes has used morphology to infer species boundaries (e.g., Ortenburger, 1923; Wilson, 1970; Johnson, 1977; Grismer, 1990), but this method can underestimate diversity as a result of cryptic species (Ruane et al*.,* 2014). Whipsnakes include 12 species (11 species in *Masticophis* and one species in *Coluber*) ranging across North America, with one species extending into northern South America. We used eight species of whipsnakes to test for isolating effects of North American geographical features: 1) *M. flagellum*, a group of snakes distributed from coast to coast in the southern half of the United

States, and into northern Mexico (MR, PR, CFB), 2) *M. fuliginosus*, restricted to the Baja California Peninsula in Mexico (GC), 3) *M. mentovarius,* distributed from central Mexico to Columbia and Venezuela (IT), 4) *M. taeniatus*, distributed from the north western United States to north eastern Mexico (RM), 5) *M. schotti*, distributed from southern Texas into northern Mexico, 6) *M. lateralis*, distributed throughout California, and the Baja California Peninsula in Mexico (GC), 7) *M. bilineatus*, restricted to the Sonoran desert in the south-western United States and into central Mexico, and 8) *C. constrictor*, distributed across the continental United States, except the Chihuahuan desert, which has been studied in detail previously (Roze, 1953; Conant & Collins, 1998; Stebbins, 2003; Burbrink et al., 2008; Richmond et al*.,* 2011; Uetz & Hošek, 2016). Most of these species possess longitudinal stripes (*M. lateralis*, *M. taeniatus*, *M. schotti*, *M. bilineatus*) while three are predominantly uniform in dorsal coloration (*M. flagellum, M. fuliginosus* and *M. mentovarius*). However, within at least one of the uniformly colored species, coloration is highly polymorphic (*M. flagellum*). Subspecies have been described for all continental whipsnakes (four island/peninsula species are not included in this study): *M. flagellum* (*M. f. cingulum*, *M. f. flagellum*, *M. f. lineatulus*, *M. f. piceus*, *M. f. ruddocki*, *M. f. testaceus, M. f. fuliginosus*), *M. mentovarius* (*M. m. centralis*, *M. m. mentovarius*, *M. m. suborbatilis*, *M. m. striolatus, M. m. variolosus*), *M. schotti* (*M. s. schotti* and *M. s. ruthveni*), *M. taeniatus* (*M. t. girardi*, *M. t. taeniatus*), *M. lateralis* (*M. l. euryxanthis*, *M. l. lateralis*), and *M. bilineatus* (*M. b. bilineatus, M. b. lineolatus*, *M. b. semilineatus*), *C. constrictor* (*C. c. anthicus*, *C. c. constrictor*, *C. c. etheridgei*, *C. c. flaviventris*, *C. c. foxii*, *C. c. helvigularis*, *C. c. latrunculus*, *C. c. mormon, C. c. oaxaca*, *C. c. paludicola*, *C. c. priapus*). Several of these subspecies are at least partly delimited by the focal geographical features of this study (e.g., *M. f. flagellum* and *M. f. testaceus* [MR], *C. c. latrunculus* and *C. c. priapus* [MR]*, M. f. testaceus* and

M. f. linatulus [PR], *M. f. cingulum* and *M. f. lineatulus* [CFB], *M. fuliginosus* and *M. f. cingulum* [GC]; Wilson, 1970).

Characteristics of geographical features

We conducted a literature review of putatively important geographical features in North America (Table 2.1). We cited studies that provided evidence of species level differentiation, population structuring, or species range boundaries divided allopatrically by geographical features. Representative focal organisms from several studies are shown in Figure 1 next to the feature of interest. We primarily focused our sampling on taxonomically similar species to whipsnakes, but have also included other examples to more broadly demonstrate the contribution of these geographical features to the diversification of North American biota (Table 2.1). We also compiled ages of each feature from the literature (Table 2.3). We used the age of the origin of the feature to encompass its entire history. For example, the MR is an ancient feature (65 million years (My)) but has likely isolated species at different magnitudes since that time, depending on climatological conditions. For each feature we used the following ages: MR 65 million years ago (Mya; Arthur & Taylor, 1998), PR 1.8 Mya (Havenor, 2003), CFB 1.8 Mya (Devitt et al*.,* 2006; and citations therein), RM 45-36 Mya (Riddle et al*.,* 2006), GC 5.5-4.0 Mya (Lonsdale, 1991), IT 6 Mya (Barrier et al*.,* 1998).

DNA extraction and mitochondrial DNA sequencing

We acquired tissue samples from across much of the range of whipsnakes, as far south as Costa Rica. Our sampling included tissues from *Masticophis bilineatus* (n=2), *M. lateralis* (n=6), *M. schotti* (n=5), *M. taeniatus* (n=13), *M. mentovarius* (n=34), *M. flagellum* (n=69), and a

putatively undescribed Mexican lineage (*Masticophis* sp., n=5) (Table A2.1; Figs. 2a, 2b). We extracted DNA from muscle, liver, shed skin, or whole blood stored in SDS buffer or 70% ethanol using a standard salt extraction protocol (Sambrook & Russell, 2001). We checked the quality of our extractions using a 1% Agarose gel and quantified the DNA using QUBIT® 2.0 Fluorometer (Life Technologies, Grand Island, NY, USA). We sequenced a 770 base pair fragment of the Cytochrome b gene for 119 individuals using custom primers (Table A2.4), designed from previous *Masticophis* sequences using Geneious v.7.0 (Kearse et al*.,* 2012). Each PCR reaction occurred in a 25 µl reaction that included 10mM tris-HCl, 50 mM KCl, 1.5mM MgCl2, 0.04 mM of each dNTP, 1 U *Taq* DNA Polymerase, 0.5 µM each primer, and 10-25 ng of DNA. The amplification protocol for all PCR reactions was: 94˚C, 2 min; 40 cycles of 94˚C 30 sec, annealing temperature 54.5˚C 30 sec, 72˚C 30 sec; 72˚C 10 min; final rest at 12˚C. PCR purifications were performed using Sera-Mag Speedbeads (Rohland & Reich, 2012). Cycle sequencing reactions were conducted using PCR primers under the following conditions: 95˚C, 2 min; 40 cycles of 95˚C 15 sec, annealing temperature 50˚C 15 sec, 60˚C 4 sec; final rest at 12˚C. Sequencing products were resolved on an Applied Biosystems 3130XL at the University of Texas Arlington Genomics Core Facility (gcf.uta.edu; Arlington, TX, USA).

Mitochondrial sequence processing and phylogenetic analyses

Raw sequences were assembled into contigs and edited by eye for sequencing errors in Geneious v7.0 (Kearse et al., 2012). We also downloaded 52 sequences from Genbank, including *M. flagellum* (n = 42), *M. bilineatus* (n = 1), *C. constrictor* (n = 3), *Drymarchon corais* (n = 1), *Opheodrys aestivus* (n = 1), *Oxybelis aeneus* (n = 1), *Phyllorhynchus decurtatus* (n = 1), *Salvadora mexicana* (n = 1), *Sonora semiannulata* (n = 1), *Spilotes pullatus* (n = 1), Tantilla

relicta (n = 1; Table A2.1). For *C. constrictor* we chose one individual from each of the three primary clades identified in Burbrink et al*.,* (2008). Sequences were aligned using Geneious aligner with default settings. Prior to phylogenetic analysis, we selected the most probable models of nucleotide evolution for Likelihood and Bayesian analyses using Bayesian information criteria implemented in PartitionFinder (Lanfear et al*.,* 2012), partitioning by codon position.

We estimated phylogenetic relationships across all taxa using maximum likelihood (ML) in raxmlGUI v1.3 (Silvestro & Michalak, 2012) with 1 000 rapid bootstrap repetitions. We partitioned by codon, using GTR + Γ for each partition. We calculated mean pairwise distances (p-distance) among haplotype groups in Mega v7 (Tamura et al., 2013). We estimated divergence times of mitochondrial clades across geographical features to place our diversification events in a historical context. We randomly sampled one individual from each haplotype group identified in our ML phylogeny, including our eight outgroups, three *C. constrictor*, five *M. flagellum*, two *M. bilineatus*, one *Masticophis* sp., two *M. lateralis*, two *M. taeniatus*, one *M. schotti*, and two *M. mentovarius*. We estimated the phylogeny using a HKY model of evolution across each codon position. To estimate divergence times across geographical features, we used a relaxed clock lognormal clock model, a calibrated Yule tree prior, and a lognormal prior on our fossil calibration points. We used two fossil calibration points, following Burbrink et al., (2008). We placed a lognormal prior on the MRCA of *Masticophis* and *Coluber* with a mean of 11 My, with a standard deviation of 0.1 (Holman, 2000). This resulted in a 95% confidence interval (CI) of 9.00–13.3 My. We also placed a lognormal prior on the root age, which encompassed all North American Colubrinae, with a mean age of 19 My and a standard deviation of 0.2. This resulted in a 95% CI of 12.6–27.6 My.

This calibration corresponds to the oldest dates of the fossils *Paracoluber* (middle Miocene) and *Salvadora* (Late Miocene; Holman, 2000). We sampled 100 000 000 generations, sampling every 10 000 generations in BEAST v2.4.5 (Bouckaert et al., 2014). We checked convergence of all parameters in Tracer (Rambaut, 2015), and summarized all trees in TreeAnnotator (Bouckaert et al., 2014). We removed the first 25% of trees as burnin and estimated the maximum clade credibility tree with median node heights.

RADseq library generation and computational analysis

We prepared ddRADseq libraries for 132 individuals following the protocol described in Peterson *et al.,* (2012). This method allows for the sequencing of thousands of orthologous loci from across the genome for large sample sets and has been successfully used in the absence of a reference genome in a variety of taxa (Eaton & Ree, 2013; Wagner et al*.,* 2013; Hipp et al*.,* 2014; Streicher et al., 2014).

We conducted double digests of 200–500 ng of DNA per individual using 20 units of *Sbf*I and 20 units of *Msp*I (NEB) for eight hours at 37˚C in 1X CutSmart Buffer (NEB). We ligated barcoded Illumina TruSeq adapters at 16˚C for 30 minutes, and heat killed the enzyme at 65˚C for 10 minutes. Each adapter included an 8 bp unique molecular identifier (UMI) that helped reduce poor quality sequence at the end of sequencing reads. We pooled up to 12 uniquely barcoded individuals into a group and labeled each group with a TruSeq single index; this double-barcoding scheme allowed us to multiplex all individuals for sequencing on a single Illumina HiSeq 2500 lane. We size selected all 11 groups using the Blue Pippin electrophoresis platform (Sage Science, Beverly, MA, USA) for fragments between 435–535bp. RAD libraries were amplified using indexed Illumina® paired end PCR primers with Phusion® High Fidelity

Proofreading Taq (NEB) under the following thermocyler conditions: 98°C, 30 sec; 12–30 cycles of 98°C 30 sec, annealing temperature 55°C 30 sec, 72°C 1 min; 72°C 5 min; final rest at 12°C. We confirmed successful library preparation using a 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA, USA) with a DNA 7500 chip kit and final concentrations were verified using the Qubit 2.0®. We pooled our 12 sub-libraries in equimolar amounts and sequenced our final library (100 bp paired end sequencing) on an Illumina® HiSeq 2500 at the University of Texas Southwestern Genomics Core facility (genomics.swmed.edu).

We processed our RAD data using the STACKS v1.12 pipeline (Catchen et al*.,* 2011). We followed the recommended workflow which implemented the following scripts and programs: (i) process_radtags which filtered out reads below 90% quality score threshold, (ii) ustacks which set a maximum distance of 3 between 'stacks', (iii) cstacks, which creates a catalogue of all of the loci within all individuals (-n flag; setting of 0) (iv) sstacks which searches the stacks created in ustacks against the catalogue from cstacks, and (v) populations, which genotypes each individual according to the matched loci from sstacks. Following populations, we used custom python scripts to filter out invariant loci and loci with more than two haplotypes. We began by processing our RAD data for all species together, but recovered very few homologous loci (<100). Thus, we analyzed each species group independently to maximize the number of homologous loci retained in each dataset. In order to test the effect of missing data on our analyses we generated three SNP datasets with varying amounts of missing data (50%, 20%, and 10% missing data per locus) for *M. flagellum* and *M. mentovarius*. For *M. lateralis* and *M. taeniatus* we filtered to > 20% missing data per locus. This resulted in datasets ranging from 80–3006 loci. At the individual level, our datasets ranged from 0–59% missing data per individual. The full number loci used in each analysis is shown in Table A2.2.

Inferring patterns of genomic divergence with Bayesian clustering

We sought to identify how geographical features may have influenced genetic diversity across the landscape by analyzing population structure in STRUCTURE (Pritchard et al., 2000). We analyzed each species group separately to avoid bias from uneven sampling (Puechmaille, 2016). Our sampling for each analysis included 36 *M. flagellum*, 24 *M. mentovarius,* five *M. lateralis* and four *M. taeniatus*. We ran STRUCTURE using all three missing data thresholds for *M. flagellum* and *M. mentovarius*. We analyzed K = 1–10, with five iterations at each K value. Each analysis was run for 500 000 generations with a burn-in of 100 000 MCMC generations. We used the independent allele frequency and the admixture ancestry model. We evaluated the results of our STRUCTURE analyses using the Evanno method (Evanno et al., 2005) implemented in STRUCTURE HARVESTER (Earl & vonHoldt, 2012). We used the highest DeltaK value to identify the best value of K for each species group.

Bayesian estimation of migration across geographical *features*

To quantify the level of isolation caused by each geographical feature, we estimated migration between populations across four features of varying permeability using Migrate-n V.3.6.9 (Beerli, 2009). To generate input files we called nuclear SNPs using default parameters in pyRAD v3.0.5 (Eaton, 2014). We generated four input files for four population pairs (see below). Each population pair required different clustering thresholds specified in the pyRAD params file depending on the number of shared loci between the populations. The clustering thresholds and the number of loci used in each Migrate-n run are reported in the Table A2.2. We conducted five independent analyses: 1) *M. flagellum* east (n=12) and west (n=23) separated by

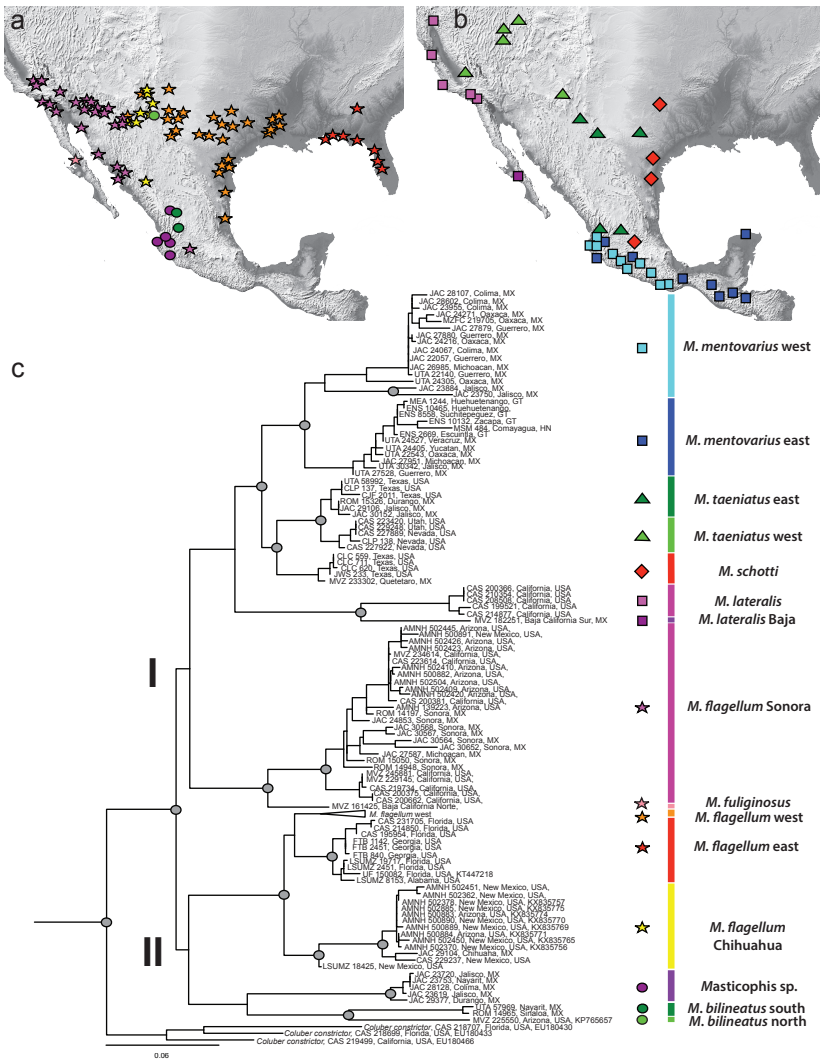

Fig. 2.2: Phylogenetic analysis of eight whipsnake species based on Cytochrome b sequencing. A) Map showing location of haplotypes for *M. flagellum, M. bilineatus*, and *Masticophis* sp. B) Map of localities for mitochondrial haplotype groups of *M. lateralis*, *M. taeniatus*, *M. schotti*, and *M. mentovarius*. C) Maximum likelihood phylogeny of mitochondrial whipsnake relationships. Dark gray circles represent nodes with ≥ 70% bootstrap support. Shapes on the map and next to the phylogeny differentiate species divisions, while different colors represent different clades.

the MR, 2) *M. flagellum* west (n=26) and Chihuahua (n=6) separated by the PR, 3) *M. flagellum* Sonoran (n=4) and Chihuahua (n=6) separated by the CFB, 4) *M. mentovarius* east (n=28) and west (n=20) of the Isthmus of Tehuantepec, 5) *M. taeniatus* individuals east (n=6) and west (n=2) of the Rocky Mountains. We used a Bayesian inference model with uniform priors for θ (mutation scaled population size; 0-0.1) and *M* (mutation scaled immigration rate; 0-10 000). After a burn in of 50 000 steps, we sampled 5000 states from the Markov chain, one every 100 steps. We sampled four heated chains at four temperatures (1, 1.5, 3, and 100 000) to thoroughly search the parameter space. We calculated migrants per generation (*Nm*) by multiplying θ with *M* and dividing by four.

Visualizing Estimated Effective Migration Surfaces

We used the program EEMS (Petkova et al*.,* 2016) to visualize how nuclear DNA-inferred migration rates were spatially distributed in select species of whipsnakes. EEMS estimates effective migration by visualizing regions where genetic dissimilarity decays quickly. It relates effective migration rates to expected genetic dissimilarities to clarify spatial features of population structure across the landscape. We ran six analyses with EEMS to estimate gene flow across the range of *M. flagellum* (n=36) and *M. mentovarius* (n=24) using our three missing data thresholds. We did not use *M. taeniatus*, *M. schotti*, *M. bilineatus*, *M. lateralis*, or the identified lineage, because of small sample sizes. We ran three independent chains for each analysis, with 500 demes, for 8 000 000 MCMC iterations, with 3 200 000 iterations of burnin and 9 999 thinning iterations. We checked convergence by analyzing the trace file produced by the accompanying plotting program, rEEMSplots.

RESULTS

Mitochondrial phylogenetic analyses and divergence dating support allopatric divergence

Our ML analysis included wide geographic and taxonomic sampling, and recovered 19 clades among all sampled species (Fig. 2c). We recovered high support ($\geq$ 70% bootstrap value) for relationships within species, but low support for many of the nodes between species. We found that *Masticophis* were monophyletic with respect to *C. constrictor.* Among these species, we recovered strong support for two large groups (excluding *C. constrictor*). The first group includes *M. flagellum* west of the CFB (including *M. fuliginosus*), *M. lateralis*, *M. mentovarius*,

M. taeniatus, and *M. schotti*. The second group included *M. flagellum* east of the CFB, *M. bilineatus* and *Masticophis* sp.

Within Clade I (Fig. 2.

2c), we recovered two clades pertaining to *M. flagellum* west of the CFB. The first clade included individuals ranging from Arizona in the east, to California in the west, and Sonora and Michoacán, Mexico in the south (*M. flagellum* Sonora). The second clade included the *M. fuliginosus* sample. These two clades were sister to seven clades pertaining to *M. lateralis, M. mentovarius*, *M. taeniatus*, and *M. schotti*, although this relationship was poorly supported. We recovered two clades within *M. lateralis*, one on mainland California, and the other on Baja California, Mexico (*M. lateralis* and *M. lateralis* Baja). Within *M. schotti*, we recovered one clade. In *M. taeniatus* we recovered two clades from the east and west of the RM (*M. taeniatus* east and west). The western clade included individuals from Utah, Nevada, and New Mexico, while the eastern clade included individuals from Texas, and Jalisco and Durango, Mexico, with additional substructure observed between the Texas and Mexico samples. Within *M. mentovarius* we observed two clades divided by the IT (*M. mentovarius* east and west). The western clade included individuals from the Pacific coast of Mexico, from Jalisco to Oaxaca, with additional substructure observed in Jalisco. The eastern clade included individuals from the Atlantic coast of Mexico, and nuclear Central America, as far south as northern Costa Rica. However, the eastern clade also included several individuals from western Mexico.

Within Clade II (Fig. 2.2c), we recovered a sister relationship between *M. bilineatus* and *Masticophis* sp. Within *M. bilineatus* we observe two clades, one from Arizona, and the other from Nayarit and Sinaloa, Mexico (*M. bilineatus* north and south). These species were sister to three clades of *M. flagellum* east of the CFB. The first clade pertained to individuals between the

CFB and the PR in the Chihuahuan Desert (*M. flagellum* Chihuahua). This clade was sister to two reciprocally monophyletic clades divided by the MR, *M. flagellum* east and west. *Masticophis flagellum* east included all samples east of the MR as far north as Georgia. *Masticophis flagellum* west included all samples between the MR and the PR, although a few samples with this haplotype came from between the CFB and the PR. Table 2.2 shows uncorrected pairwise distances between each clade recovered in the ML analysis. Inter-clade divergences ranged from 3.4% between *M. flagellum* east and west to 16.1% between *M. lateralis* Baja and *M. flagellum* Sonora.

Table 2.2. Mean between group divergences generated from uncorrected *p* distances among Cytochrome b haplogroups in the whipsnake species complex.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1. *M. flagellum* east															
2. *M. flagellum* west	3.4														
3. *M. flagellum* Chihuahua	4.9	5													
4. *M. flagellum* Sonora	10.7	10.3	9.7												
5. *M. fuliginosus*	8.3	9.8	8.6	7.3											
6. *M. bilineatus* north	9.5	9.6	6.6	13.3	12.5										
7. *M. bilineatus* south	10.5	8.9	7.9	14.2	13.8	5.8									
8. *Masticophis* sp.	10.6	10.9	9.4	11.6	10.3	8.9	11.2								
9. *M. mentovarius* west	9.2	9.2	7.5	12.3	11	11.6	12.9	12.5							
10. *M. mentovarius* east	8.9	8.6	8.6	11.8	10.5	12.3	11.3	12.6	8.8						
11. *M. taeniatus* west	9.9	10	7.7	11.6	10.1	10	10.9	10.3	8.8	7.4					
12. *M. taeniatus* east	9.8	9.8	8.4	12.3	10.4	10.8	10.8	11.8	9.8	8.4	2				
13. *M. schotti*	10.6	9.9	9	11.4	9.8	11.8	12.2	11.2	9.4	6.9	5.2	6.6			

14. *M. lateralis*	12.1	11.7	10.3	14.1	13.7	11.9	10	14.6	12.6	12	11.3	11.8	10.3		
15. *M. lateralis* Baja	14.6	14.7	11.9	15.2	13.4	13.4	13.4	15.2	13.7	12.8	10.7	10.5	11	8.2	
16. *C. coluber*	11.3	11.9	9.1	11.7	10.6	12.2	13.4	11.5	13.5	12.7	11.7	12.6	12.2	13.8	15.3

Our Bayesian inference (BI) of phylogenetic relationships revealed similar phylogenetic structure between haplotype groups as the ML analysis with the two exceptions of *M. flagellum* Sonora and *M. fuliginosus*, and the relationship between *M. bilineatus* and *Masticophis* sp. (Fig. 2.3). In the BI analysis, we recovered *M. flagellum* west of the CFB as sister to all species in group I, rather than group II (0.90 PP). We also recovered *M. bilineatus* as sister to *M. flagellum* east of the CFB, instead of to *Masticophis* sp. (0.68 PP). However, the relationship of *M. bilineatus* was not recovered with high support, underscoring the phylogenetic uncertainty of this species.

Our estimates of divergence dates placed the oldest divergence event between *C. constrictor* and all other species at 10.8 Mya (95% HPD 9.02–12.82; 1.00 PP; Table 2.3; Fig. 2.3). We found that clades I and II diverged 8.56 Mya (6.83–10.48; 1.00 PP). Within group I, *M. lateralis* diverged from *M. mentovarius*, *M. taeniatus*, and *M. schotti* 7.05 Mya (5.23–9.09; 0.98 PP).

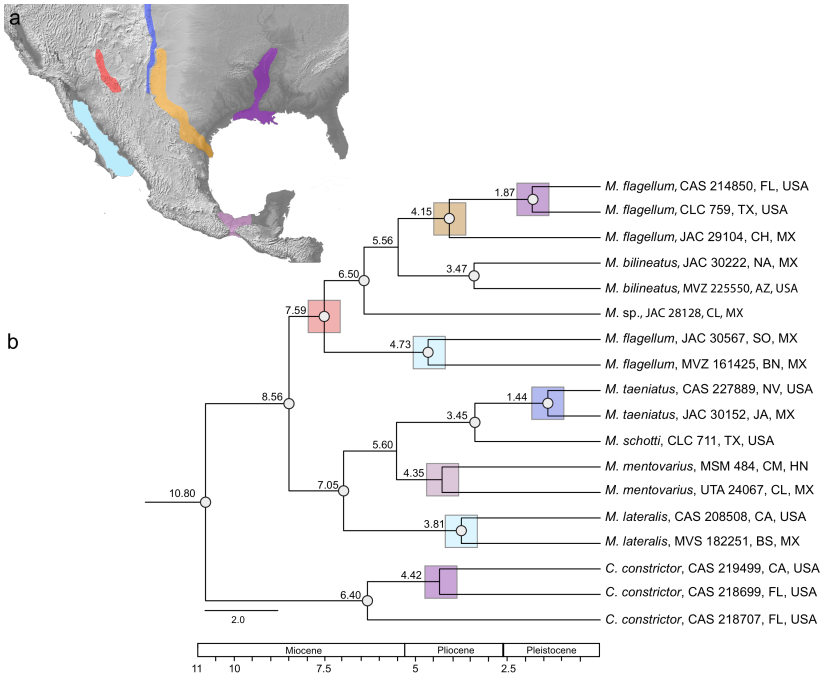

Fig. 2.3: Bayesian phylogenetic analysis and divergence time estimation. A) Map showing the geographical features of interest from Fig. 1. B) Bayesian phylogeny generated in BEAST. Nodes with ≥ 90% posterior probability are colored with a gray circle. The colored boxes behind the nodes signal phylogenetic breaks that correspond to geological features. The mean divergence time is shown above each node.

Masticophis lateralis clades were split by the GC 3.81 Mya (2.34–5.53; 1.00 PP). *Masticophis mentovarius* diverged from *M. taeniatus* and *M. schotti* 5.60 Mya (3.87–7.43; 0.85 PP), and was split by the IT 4.35 Mya (2.90–6.28; 0.67 PP). *Masticophis taeniatus* diverged from *M. schotti* 3.45 Mya (2.07–5.12; 0.99 PP), and was split by the RM 1.44 Mya (0.66–2.40; 1.00 PP). Within group II, we found the oldest divergence event at the CFB, where *M. flagellum* Sonora split from the eastern lineages 7.59 Mya (5.90–9.49; 0.90 PP). In addition, the Sonoran lineage of *M. flagellum* diverged at the GC 4.73 Mya (3.05–6.76; 0.99 PP). *Masticophis* sp. diverged from *M. bilineatus* and *M. flagellum* 6.50 Mya (4.89–8.38; 0.97 PP), and *M. bilineatus* diverged from *M. flagellum* 5.56 Mya (4.04–7.12; 0.68 PP). The northern and southern clades of *M. bilineatus* diverged 3.47 Mya (2.19–4.90; 1.00 PP). *Masticophis flagellum* diverged at the PR 4.15 Mya (2.77–5.84; 0.99 PP), and at the MR 1.87 Mya (1.00–2.89; 1.00 PP). This places the majority of divergence events in whipsnakes within the late Miocene and the Pliocene, with only two events occurring during the Pleistocene. However, including our 95% HPD, several events may have occurred in the early Pleistocene. Of the eight clades that are separated by geographical features, only two divergence events were older than the date of formation of the current geographical feature (CFB and PR), providing additional support for the role of geographical features in promoting diversification in allopatry. We note that the divergence events across the IT by *M. mentovarius* and across the MR by *C. constrictor* are not strongly supported in this analysis, and thus should be interpreted with caution.

Table 2.3. Estimated divergence times of whipsnake clades across geographical features. Clades are defined by ML haplotype groups shown in Fig. 2.2.

Feature	**Haplotype groups**	**Divergence age (my)**	**95% CI**	**Date of formation (my)**
Mississippi River	*M. flagellum* east/west	1.87	1.00-2.89	65
Mississippi River	*C. constrictor* east/west	4.4	2.86-6.21	65

Pecos River Valley	*M. flagellum* west/Chihuahua	4.15	2.77-5.84	1.8
Cochise Filter Barrier	*M. flagellum* Sonora/Chihuahua	7.59	5.90-9.49	1.8
Rocky Mountains	*M. taeniatus* east/west	1.44	0.66-2.40	45
Gulf of California	*M. flagellum* Sonora/*M. fuliginosus*	4.73	3.05-6.76	5.5
Gulf of California	*M. lateralis* mainland/Baja	3.82	2.34-5.53	5.5
Isthmus of Tehuantepec	*M. mentovarius* east/west	4.35	2.90-6.28	6

Genomic variation forms discrete allopatric clusters at multiple scales

We used the Evanno method to infer K values from our STRUCTURE analyses for *M. flagellum* (K = 5), *M. mentovarius* (K = 2)*, M. taeniatus* (K = 2)*,* and *M. lateralis* (K = 3). Results shown in Fig. 2.3 correspond to the datasets with 20% missing data; results for the other missing data thresholds show similar patterns and are summarized in Fig. A2.1. The five clusters of *M. flagellum* individuals corresponded to samples from 1) the Baja California Peninsula 2) west of the CFB 3) between the CFB and the PR 4) between the PR and the MR 5) east of the MR (Figs. 4a, 4e). The two *M. mentovarius* clusters corresponded to individuals west and east of the IT (Figs. 4b, 4f). *Masticophis taeniatus* clustering corresponded to samples east and west of the RM (Figs. 4c, 4f). *Masticophis lateralis* populations divided between the California mainland and Baja California, and the southern California sample showed evidence for an intermediate population (Figs. 4d, 4f). Notably, all the major genomic clusters inferred from our nuclear SNP sampling occurred on opposite sides of our focal geographical features (Fig. 2.4). Our analyses that utilized different missing data thresholds recovered similar results. In *M. flagellum*, we found that allowing up to 50% missing data at the locus level, and up to 59.6% missing data at the individual level recovered very similar results to the dataset shown in Fig. 2.4 (Fig. A2.1a).

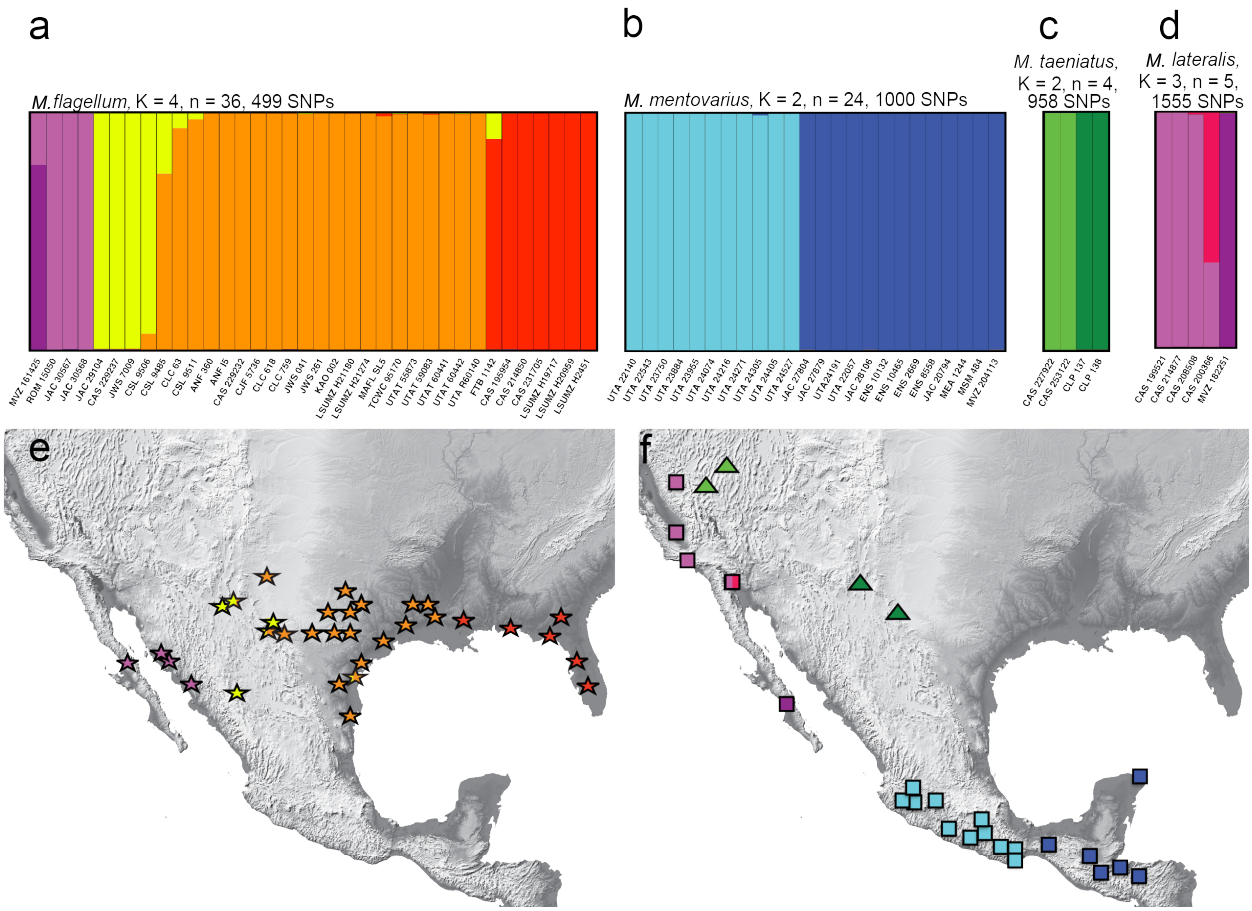

Fig. 2.4: Graphical results of the nuclear analyses. A-D) Genomic clustering results for each species group. Species name, inferred value of K, sample size, and number of SNPs used is labelled above each STRUCTURE plot. E) Maps showing locations of genomic clusters for *M. flagellum*. F) Map showing location of genomic clusters for *M. mentovarius*, *M. taeniatus*, and *M. lateralis*.

Allowing only 10% missing data did not recover *M. fuliginosus* as an independent population, and included an additional population within *M. flagellum* west that may not correspond to real genetic structure. In *M. mentovarius* we recovered congruent population assignments across missing data thresholds (Fig. A2.1b.

Migration occurs asymmetrically across some geographical features

Our Migrate-n analyses supported migration across five primary geographical features (Fig. 2.5). All analyses reached convergence. We report the mean values for each parameter in Table A2.3. Across the CFB we found that that Sonoran samples exchanged 0.800 migrants per

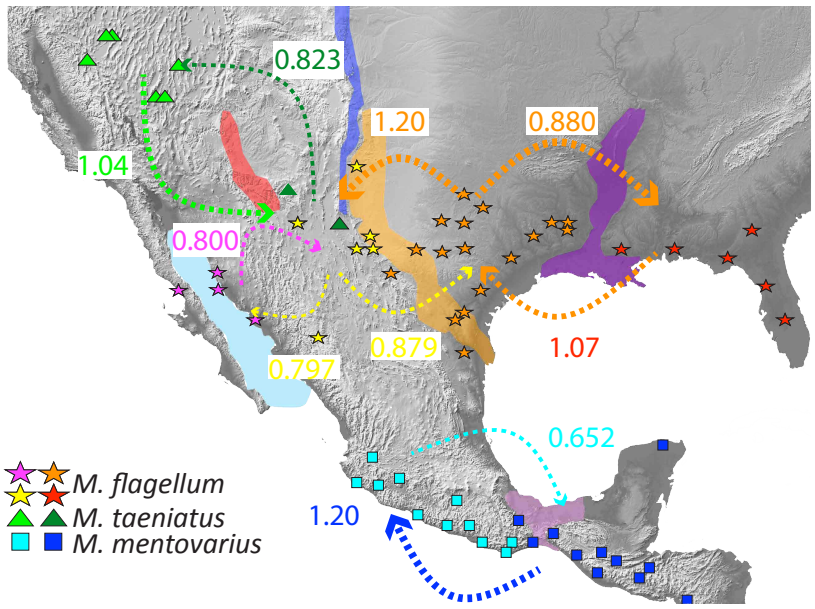

Fig. 2.5: Results of the Migrate-n analyses. Values are given for migrants per generation between genomic clusters. Arrows are sized to indicate the strength of migration in each direction. Geographical features are shown behind each migration estimate.

generation (*Nm*) with the Chihuahuan clade, which exchanged 0.797 *Nm* in return. Considering the large divergence between the Sonoran clade and the other *M. flagellum,* we consider this to be low (and equal) levels of migration that helped us contextualize our other comparisons. Across the PR, we found evidence for asymmetric gene flow from east to west, with the western clade exchanging 1.20 *Nm*, and the Chihuahuan clade exchanging 0.879 *Nm*. We found a similar east to west pattern across the MR, with the eastern clade exchanging 1.07 *Nm*, and the western clade returning 0.880 *Nm*. At the RM we found that northern *M. taeniatus* exchanged 0.860 and southern *M. taeniatus* exchanged 0.072 *Nm*. At the IT we again found asymmetrical migration from east to west, with *M. mentovarius* east and *M. mentovarius* west exchanging 1.20 and 0.652 *Nm*, respectively. Across the MR, PR, and the IT, we found stronger migration from east to west, while across the RM we found asymmetrical rates from west to east. Across the CFB we found symmetrical rates of migration.

Estimated effective migration reveals additional population structure and centres of genetic diversity

Estimated effective migration surfaces analyses for *M. flagellum* supported four populations, with strong barriers to gene flow at the MR, the northern PR, and the CFB. We recovered evidence for weak isolation between the *M. flagellum* west and Chihuahuan population in the southern portion of their putative contact area (Fig. 2.6a). We found no evidence for gene flow between the Sonoran population and the other three populations. We found that the centre of diversity for this species lies in the western clade, but that regions of genetic diversity existed in the Chihuahuan Desert, and in the northern range of the eastern clade. We observed low levels of genetic diversity in the southern part of the eastern clade where gene flow was more prevalent, as well as in the Sonoran clade (Fig. 2.6b). Figure A2.3 regresses genetic distance against Euclidian geographic distance to show that genetic distance is partitioned into four groups and is not equal across the landscape.

Our analyses of *M. mentovarius* showed a reduction of gene flow at the IT as we observed in STRUCTURE, but also showed additional population structuring on both sides of the IT (Fig. 2.6c). The centre of diversity for this species was recovered to the west of the IT on the Pacific coast (Fig. 2.6d). We recovered very low diversity estimates for the population east of the IT. We found that *M. mentovarius* is not isolated by distance, but that genetic variation and diversity are partitioned at the Isthmus of Tehuantepec (Fig. A2.3). Analyses under three missing data thresholds are shown in Fig. A2.2.

DISCUSSION

We used genome wide SNPs and mitochondrial sequence data to evaluate major genomic divisions within whipsnakes to quantify migration and structure between clusters associated with geographical features. We found that whipsnake genomic clusters largely corresponded to

geographical features, indicating that these features played a notable role in the diversification of whipsnakes. Our genomic data supported twelve clusters of whipsnakes, and our expanded mitochondrial sampling revealed extensive diversification within each species. Divergence dating suggested that most diversification events in whipsnakes occurred during the late Miocene or early Pliocene. We tested migration across four geographical features that partitioned genetic clusters and found evidence for asymmetric gene flow occurring from east to west in *M. flagellum* across the MR and the PR, and in *M. mentovarius* across the IT (Fig. 2.5). We observed a west to east pattern of migration across the RM in *M. taeniatus* and symmetrical rates across the CFB in *M. flagellum* (Fig. 2.5). However, our sampling for *M. taeniatus* was limited. Our estimated effective migration surfaces revealed strong differentiation at the MR, the CFB, the PR, and at the IT (Fig. 2.6). We observed that populations that received more migrants had higher levels of genetic diversity (Figs. 2.5 and 2.6b, 2.6d). These results add more evidence of the importance of geographical features in driving diversification of North American biota by isolating populations in allopatry.

Whipsnake phylogenetics, phylogeography, and taxonomy

We recovered extensive geographical structuring within each group. While much of our nuclear clustering assigned individuals to distinct clusters with high probability, the support values for some mitochondrial relationships were low at deeper nodes. This result could be explained by an initial rapid radiation in this group, or be indicative of substitution saturation of evolving mitochondrial DNA (Rothfels et al., 2012; Streicher et al., 2014). While we observed substantial similarity between the groups recovered by the mitochondrial and nuclear analyses, some results were discordant. Our phylogenetic analyses with mitochondrial data recovered *M.*

flagellum as non-monophyletic with respect to *M. bilineatus* and *Masticophis* sp. (Fig. 2.2c). More extensive sampling of *M. bilineatus* would help resolve its relationship with *M. flagellum*. Likewise, while *M. flagellum* from west of the CFB clustered with group II in the ML analysis, rather than with the other *M. flagellum,* but in the BI analysis we recovered western *M. flagellum* as sister to all other *M. flagellum*, *M. bilineatus*, and *Masticophis* sp. In light of these findings, much work remains to resolve the relationships of all whipsnake species. More extensive nuclear sampling of *M. flagellum* west of the PR, *M. bilineatus*, and *Masticophis* sp. would help toward this objective.

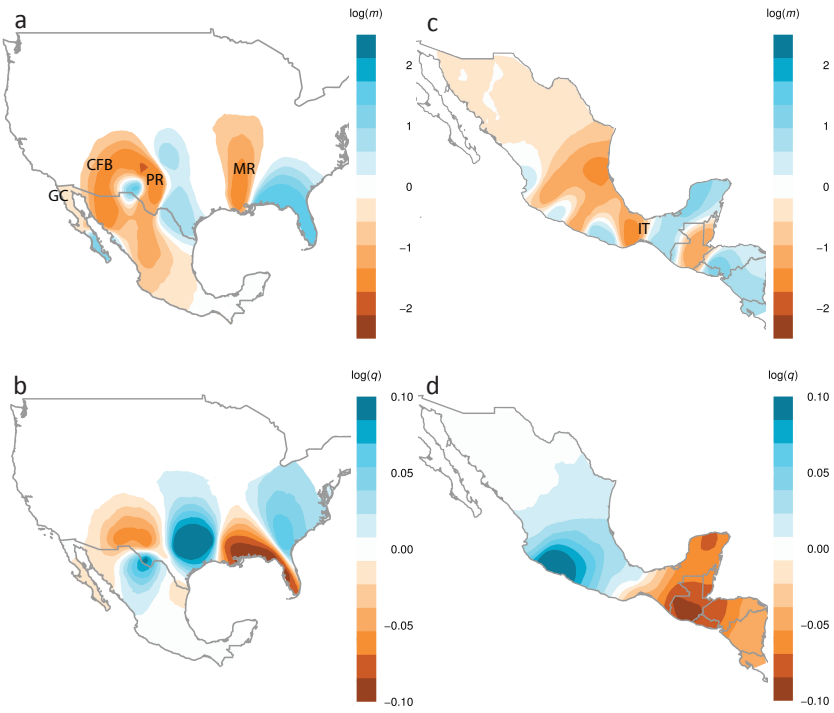

Fig. 2.6: Graphical representations of estimated effective migration and diversity surfaces (EEMS). High values are represented by shades of blue; low values are represented by red-orange shades. A-B show results for *M. flagellum,* and C-D show results for *M. mentovarius*. A, C) Estimated effective migration surfaces. B, D) Estimated effective diversity surfaces. Focal geographical features are labeled in A and C: GC = Gulf of California, CFB = Cochise Filter Barrier, PR = Pecos River, MR = Mississippi River, IT = Isthmus of Tehuantepec.

Migration occurs asymmetrically across geographical features in whipsnakes

Our two migration analyses show largely concurrent results that differ in scale. Lower rates of migration inferred by Migrate-n appear as much darker breaks in the EEMS analyses (Figs. 5 and 6a, 6d). Not surprisingly, the populations with the highest levels of genetic diversity are

those that receive more migrants in both *M. flagellum* and *M. mentovarius*. However, the western clade of *M. flagellum* appears to have contributed higher levels of migrants to the Chihuahuan clade than to the eastern clade, yet shows little evidence for migration in Figure 6a. Our data also suggest that the Sonoran clade exchanges few migrants with the Chihuahuan clade (Figure 6a). Low migration may explain the ~10% mitochondrial divergence between the Sonoran clade and other *M. flagellum*.

Migration patterns observed in *M. mentovarius* are consistent with the distribution of mitochondrial haplotypes, where the eastern haplotype was present to the west of the IT, but the western haplotype was not recovered east of the IT. This may suggest that eastward migration (inferred using Migrate-n) has carried the eastern mtDNA haplotype westward. The higher level of genetic diversity inferred in the west by EEMS (Fig. 2.6d) suggests that westward-biased migration has increased genetic diversity in the west disproportionately (Fig. 2.5). Unlike the MR, the CFB, and the CD, the IT changed from a shallow embayment to a land bridge within the last 2 my, allowing for previously isolated populations to experience secondary contact (Barrier et al., 1998).

Other recent studies have also inferred migration in reptiles using molecular data. Grummer et al*.,* (2015) estimated migration between fence lizard populations in the Mexican highlands (using IMA2; Hey, 2010), and Suárez-Atilano et al., (2014) used Migrate-n to estimate migration between *Boa constrictor* populations using microsatellite markers and found high levels of gene flow between them (7.78–31.1 *Nm*). Alternatively, Ruane et al*.,* (2014) used Migrate-n to estimate rates of migration between milksnake species, rather than populations, and recovered very low rates of gene flow (0.00–1.22 *Nm*). The rates observed in our study are considerably lower than those found in Suárez-Atilano et al*.,* (2014), but higher than those found

in Ruane et al., (2013), reflecting the various levels of divergence that we investigated across geological features. Similar studies have also inferred high rates of migration between bird populations, including mallards (0.42–8.26 *Nm*), flamingos (0.40–7.88 *Nm*), and black footed albatrosses (0.02–4.5 *Nm*; Kraus et al., 2012; Geraci et al., 2012; Dierickx et al., 2015). While it is unsurprising that rates of migration are higher in flying vertebrates than in less mobile snakes from this study, rates of migration estimated for birds were lower than those estimated for boas (Suárez-Atilano et al., 2014). More studies of migration using genomic data are needed to identify 'normal' rates of migration between reptile populations and lineages. However, migration among whipsnake groups provides evidence that migration associated with geographical features may be lower than in other species.

Geographical features promoted diversification at multiple timescales in whipsnakes

To place our findings in a historical context, we compared the influence of each geographical feature on whipsnakes to that of past studies. Certain features (i.e. MR, IT) consistently separated populations or sister species into discrete groups. This could be due to the nature of the features, as these two represent water crossings, while the CFB and PR represent habitat transitions, and generally more recent histories of isolation. Our study found support for six geographical features that limit gene flow in North American biota (Fig. 2.1).

The Mississippi River has exerted a strong isolating force on a variety of taxa including plants, amphibians, fish, reptiles, and mammals (Fig 1, Table 2.1). The MR serves as the primary isolating boundary for *M. flagellum* in the east. The MR formed as long ago as 65 my (Arthur & Taylor, 1998), implicating this as an ancient isolating boundary for snakes (Castoe et al., 2012; Streicher et al., 2016). We found that whipsnakes diverged across this feature 1.87 Mya, which is

the most recent diversification event in *M. flagellum*. Yet despite this recent divergence, we recover completely sorted lineages in the mitochondrial and nuclear datasets at this feature. This differs from the older divergence at the PR that shows higher levels of admixture or incomplete lineage sorting (ILS; Fig 2a, 4a). *Coluber constrictor* also exhibits lineage divergence at the MR. Burbrink et al*.,* (2008) found that *C. constrictor* diverged 6.09 Mya at the MR, while our divergence time estimates placed this split at 4.42 Mya. This discrepancy may be due to our reduced lineage sampling for *C. constrictor*, despite utilizing the same calibration points. However, our estimates fall well within the confidence limits of this split estimated by Burbrink et al., (2008). In *C. constrictor,* Burbrink et al., (2008) did not find the northern end of the MR to be an effective barrier. For *M. flagellum*, which does not extend as far north, the MR remains a consistent barrier. Thus, we found that the MR serves as a strong barrier for many taxa, and that *M. flagellum* and *C. constrictor* have diverged across it at different times and have exhibited distinct biogeographic histories.

The Cochise Filter Barrier has a complex geological and climatological history, which may have isolated species asynchronously (Myers et al., 2017). The uplift of the Sierra Madre Occidental during the late Miocene, the formation of the Sonoran and Chihuahuan Deserts during the Pliocene, and isolation in Pleistocene glacial refugia may have all influenced species diversification at this feature (Morafka, 1977; Axelrod, 1979; Moore & Jansen, 2006; Wilson & Pitts, 2010a). Myers et al., (2017) identified 12 snake population or species pairs that were genetically differentiated at the CFB. They found that most snake species diverged at this feature during the Pleistocene or Pliocene, but that two species diverged ~6 Mya (*M. flagellum* and *Hypsiglena torquata*). Our divergence dating suggested that *M. flagellum* diversified at the CFB at 7.59 My, an event likely influenced by Miocene mountain building, and later reinforced by

Pliocene desert formation. This Miocene diversification differs from most other co-distributed snakes, which diverged at the CFB due to Pleistocene glacial cycles or Pliocene aridification (Myers et al*.,* 2017). In fact, divergence at the CFB is the oldest within-species divergence event observed in whipsnakes, supporting a Sonoran origin and subsequent eastward colonization for *M. flagellum*. Interestingly, the opposite pattern was observed in *C. constrictor,* which seems to have an eastern origin followed by westward expansion (Burbrink et al*.,* 2008).

The Pecos River separates the Chihuahuan Desert from the North American Grasslands (Morafka, 1977). This region between the PR and the CFB has been largely shaped by Pleistocene era processes, as glacial cycles created refugia that separated the Sonoran and Chihuahuan Deserts (Riddle & Hafner, 2006). However, the formation of the Chihuahuan Desert during the Pliocene may have also isolating species into discrete habitats (Wilson & Pitts, 2010). The PR inhibits gene flow of several other taxa, including rattlesnake populations (*Crotalus atrox*; Schield et al*.,* 2015), several species of fence lizards (*Sceloporus magister, S. undulates, S. cowlesi, S. consobrinus*; Leaché & Reeder, 2002; Leaché *et al.,* 2007; Leaché, 2009), and several species of mice (*Chaetodipus penicillatus*, *Peromyscus maniculatus*, *Peromyscus eremicus*, and *Onychomys* spp.; Lansman et al., 1983; Riddle & Honeycutt, 1990; Lee et al., 1996; Wadpole et al., 1997). While past studies have found that individuals occupying the Chihuahuan desert region are most closely related to either Sonoran Desert, or Colorado Plateau populations (west of the PR; Leaché & Mulcahy, 2007), our study found that the Chihuahuan clade of *M. flagellum* was most closely related to the western and eastern clades (east of the PR; Fig. 2.3a). This difference suggests that either the PR is more permeable to whipsnakes than the CFB, or it could reflect the more recent divergence between the western, eastern and Chihuahuan clade. We also recovered higher rates of migration (and admixture or ILS) across the PR than the CFB (Figs. 2a,

4a, 5). Finally, we found that the timing of divergence of this clade (4.15 Mya) likely corresponds to Pliocene desertification, rather than the more recent formation of the PR.

Our mitochondrial and nuclear clustering analyses found that populations of *M. taeniatus* on either side of the RM were genetically distinct. We also found asymmetrical rates of migration between these two populations from west to east. Our divergence dating of the two *M. taeniatus* clades estimated very recent divergence for these clades, suggesting that the ancient formation of the feature did not separate an already widespread species. However, we emphasize caution in the interpretation of our results regarding *M. taeniatus* due to the small sampling sizes. Our mitochondrial and nuclear datasets are consistent with the RM having isolated this species, but more complete sampling is necessary.

The Gulf of California separated the Baja California Peninsula from western Mexico 5.5-4.0 Mya (Lonsdale, 1991). This barrier has isolated many taxa; mammals, birds, snakes, and insects, and many species are endemic to the peninsula (Grismer, 2000; Rodríguez-Robles & de Jesús-Escobar, 2000; Castoe et al., 2007). Our mitochondrial data were consistent with past studies in that both our species sampled from the peninsula, *M. lateralis* and *M. flagellum,* had unique haplotypes found there. Our divergence dating estimated divergence at this barrier for *M. flagellum* at 4.73 My, and at 3.81 Mya for *M. lateralis*. Both these data estimates are after the formation of the feature, indicating that these species likely invaded the isthmus from the North after the isthmus had separated from the mainland.

The Isthmus of Tehuantepec has been implicated in the diversification of birds, amphibians, reptiles, mammals, and plants (Table 2.1). The IT was submerged until the late Miocene or early Pliocene (~6 Mya; Barrier *et al.,* 1998; Ornelas *et al* 2013). Therefore, unlike the MR, the IT represents an ancient barrier that likely no longer isolates terrestrial species. We

found that the IT separates two distinctive nuclear clusters of *M. mentovarius* (Fig. 2.3b). Our migration analyses suggested that migration has proceeded largely from east to west across this feature (Figs. 5 and 6c). For this reason, we expected smaller effective population sizes in the west, but our Migrate-n analysis estimated equal effective population sizes (Table A2.3). This may indicate that further population subdivision occurs to the south of Guatemala and Honduras, but we lacked substantial sampling there.

Our study has emphasized the role of geographical features such as rivers, an isthmus, mountains, and depressions as forces of diversification based on their ability to divide populations into isolated units. We quantified the influence specifically of the Mississippi River, the Pecos River, the Cochise Filter Barrier, the Rocky Mountains, and the Isthmus of Tehuantepec. We found that each of these features has likely played a role in the diversification of snake species that are distributed across them. This study supports the tenant that allopatric divergence is the predominant mode of diversification for terrestrial vertebrates, even among relatively vagile and widely distributed animals like whipsnakes.

ACKNOWLEDGMENTS

We would like to thank the three anonymous reviewers for their feedback on this manuscript. We would also like to sincerely thank the American Museum of Natural History for funding this work through the Theodore Roosevelt Memorial Grant. Wet lab work was funded through the UT Arlington start up fund for MFK. We would like to thank the following museum collections for providing tissues and support towards this project: California Academy of Sciences (Jens Vindum), American Museum of Natural History Ambrose Monell Cryo Collection (Julie Feinstein), Louisiana State University Museum of Natural Science (Robb

Brumfield, Christopher Austin, Eric Rittmeyer), Kansas University Biodiversity Institute and Natural History Museum (Rafe Brown), Museum of Vertebrate Zoology Berkeley (Jim McGuire, Carol Spencer), Royal Ontario Museum (Robert Murphy), Texas A&M University (Toby Hibbits), University of Central Florida (Chris Parkinson, Jason Strickland), University of Texas El Paso (Carl Lieb), Tarleton State University (Jesse Meik), University of Texas Arlington (Carl Franklin, Corey Roelke). USDA Forest Service, Southern Research Station (Josh Pierce). University of Nebraska Lincoln (Dennis Ferraro), College of Staten Island (Frank Burbrink, Ed Myers), Georgia Southern University (Christian Cox), Smithsonian Institution (Dan Mulcahy). We would also like to thank D. Petkova for help with EEMS. We used the reptile database (www.reptiledatabse.org) to aid our literature searches.

REFERENCES

Aguirre G, Morafka DJ, Murphy RW (1999) The peninsular archipelago of Baja California: A thousand kilometers of three lizard genetics. *Herpetologica*, 55, 369–381.

Al-Rabab'ah MA, Williams CG (2004) An ancient bottleneck in the Lost Pines of central Texas. *Molecular Ecology*, 13, 1075-1084.

Amman BR, Bradley RD (2004) Molecular evolution in Baiomys (Rodentia: Sigmodontinae): evidence for a genetic subdivision in *B. musculus*. *Journal of Mammalogy*, 85, 162-166.

Anderson CG, Greenbaum E (2012) Phylogeography of northern populations of the black-tailed rattlesnake (*Crotalus molossus* Baird And Girard, 1853), with the revalidation of *C. ornatus* Hallowell, 1854. *Herpetological Monographs*, 26, 19–57.

Arnold B, Corbett-Detig RB, Hartl D, Bomblies K (2013) RADseq underestimates diversity and introduces genealogical biases due to nonrandom haplotype sampling. *Molecular Ecology*, 22, 3179-3190.

Arthur JK, Taylor RE (1998) Ground-water flow analysis of the Mississippi Embayment aquifer system, South-Central United States. *USGS Professional Paper,* 1416-I, 1-48.

Ashton KG, de Queiroz A (2001) Molecular Systematics of the Western Rattlesnake, *Crotalus viridis* (Viperidae), with comments on the utility of the D-Loop in phylogenetic studies of snakes. *Molecular Phylogenetics and Evolution*, 21, 178-189.

Austin JD, Lougheed SC, Boag PT (2004) Discordant temporal and geographic patterns in maternal lineages of eastern north American frogs, *Rana catesbeiana* (Ranidae) and *Pseudacris cricifer* (Hylidae). *Molecular Phylogenetics and Evolution*, 32, 799-816.

Avise JC, Arnold J, Ball RM, Bermingham E, Lamb T, Neigel JE, Reeb CA, Saunders NC (1987) Intraspecific phylogeography: the mitochondrial DNA bridge between population genetics and systematics. *Annual Review of Ecology and Systematics*, 18, 489-522.

Axelrod DI (1979) Age and origin of Sonoran Desert vegetation. *Occasional papers of the California Academy of Sciences*, 132, 1–74.

Barrier E, Velasquillo L, Chavez M, Gaulon R (1998) Neotectonic evolution of the Isthmus of Tehuantepec (southeastern Mexico). *Tectonophysics*, 287, 77-96.

Barton HD, Wisely SM (2012) Phylogeography of striped skunks (*Mephitis mephitis*) in North America: Pleistocene dispersal and contemporary population structure. *Journal of Mammalogy*, 93, 38-51.

Baumgarten A, Williamson GB (2007) The distributions of howling monkeys (Alouatta pigra and A. palliata) in southeastern Mexico and Central America. *Primates*, 48, 310-315.

Beerli P (2009) How to use MIGRATE or why are Markov chain Monte Carlo programs difficult to use. *Population Genetics for Animal Conservation*, 17, 42-79.

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.

Belfiore NM, Liu Liang, Mortiz C (2008) Multilocus phylogenetics of a rapid radiation in the genus *Thomomys* (Rodentia: Geomyidae). *Systematic Biology*, 57, 294-310.

Barnard-Kubow KB, Debban CL, Galloway LF (2015) Multiple glacial refugia lead to genetic structuring and the potential for reproductive isolation in a herbaceous plant. *American Journal of Botany*, 102, 1842-1853.

Bouckaert R (2010) DensiTree: making sense of sets of phylogenetic trees. *Bioinformatics*, 26, 1372–1373.

Bouckaert R, Heled J, Kühnert D, Vaughan T, Wu CH, Xie D, Suchard MA, Rambaut A, Drummond AJ (2014) BEAST 2: a software platform for Bayesian evolutionary analysis. *PloS Computational Biology*, 10, e1003537.

Brant SV, Orti G (2002) Molecular phylogeny of short-tailed shrews *Blarina* (Insectivora: Soricidae) *Molecular Phylogenetics and Evolution*, 22, 163-173.

Brant SV, Orti G (2003) Phylogeography of the Northern short-tailed shrew, *Blarina brevicauda* (Insectivora: Soricidae): past fragmentation and postglacial recolonization. *Molecular Ecology*, 12, 1435-1449.

Brunsfeld S, Sullivan, J, Soltis DE, Soltis PS (2001) Comparative phylogeography of northwestern North America: a synthesis. *Special Publication-British Ecological Society*, 14, 319-340.

Bryant D, Bouckaert R, Felsenstein J, Rosenberg NA, RoyChoudhury A (2012) Inferring species trees directly from biallelic genetic markers: bypassing gene trees in a full coalescent analysis. *Molecular Biology and Evolution*, 29, 1917-1932.

Bryson RW, García-Vázquez UO, Riddle BR (2011) Phylogeography of Middle American gophersnakes: mixed responses to biogeographical barriers across the Mexican Transition Zone. *Journal of Biogeography*, 38, 1570-1584.

Bryson RW, Murphy RW, Lathrop A, Lazcano-Villareal D (2011) Evolutionary drivers of phylogeographical diversity in the highlands of Mexico: a case study of the *Crotalus triseriatus* species group of montane rattlesnakes. *Journal of Biogeography*, 38, 697-710.

Bryson RW, Riddle BR, Graham MR, Tilston Smith B, Prendini L (2013) As old as the hills: montane scorpions in south-western North America reveal ancient associations between biotic diversification and landscape history, *PloS one*, 8.1 e52822.

Bryson RW, Savary WE, Prendini L (2013) Biogeography of scorpions in the Pseudouroctonus minimus complex (Vaejovidae) from south-western North America: implications of ecological specialization for pre-Quaternary diversification. *Journal of Biogeography*, 40, 1850-1860.

Burbrink FT (2002) Phylogeographic analysis of the cornsnake (*Elaphe guttata*) complex as inferred from maximum likelihood and Bayesian analyses. *Molecular Phylogenetics and Evolution*, 25, 465-476.

Burbrink FT, Lawson R, Slowinski JB (2000) Mitochondrial DNA phylogeography of the polytypic North American rat snake (*Elaphe bsolete*): a critique of the subspecies concept. *Evolution*, 54, 2107-2118.

Burbrink FT, Fontanella F, Pyron AR, Guiher TJ, Jimenez C (2008) Phylogeography across a continent: The evolutionary and demographic history of the North American racer (Serpentes: Colubridae: *Coluber constrictor*). *Molecular Phylogenetics and Evolution*, 47, 274-288.

Burbrink FT, Myers EA (2015) Both traits and phylogenetic history influence community structure in snakes over steep environmental gradients. *Ecography*, 38, 1036-1048.

Castoe TA, Spencer CL, Parkinson CL (2007) Phylogeographic structure and historical demography of the western diamondback rattlesnake (Crotalus atrox): a perspective on North American desert biogeography. *Molecular Phylogenetics and Evolution*, 42, 193-212.

Castoe TA, Daza JM, Smith EN, Sasa MM, Kuch U, Campbell JA, Chippindale PT, Parkinson C L (2008) Comparative phylogeography of pitvipers suggests a consensus of ancient Middle American highland biogeography. *Journal of Biogeography*, 36, 88-103.

Castoe TA, Streicher JW, Meik JM, Ingrasci MJ, Poole AW, de Koning APJ, Campbell JA, Parkinson CL, Smith EN, Pollock DD (2012) Thousands of microsatellite loci from the venomous coralsnake (*Micrurus fulvius*) and variability of select loci across populations and related species. *Molecular Ecology Resources*, 12, 1105-1113.

Catchen JM, Amores A, Hohenlohe P, Cresko W, Postlethwait JH (2011) Stacks: building and genotyping loci de novo from short-read sequences. *G3: Genes, Genomes, Genetics*, 1, 171-182.

Carleton MD, Sanchez O, Urbano Vidales G (2002) A new species of Habromys (Muroidea: Neotominae) from México, with generic review of species definitions and remarks on diversity patterns among Mesoamerican small mammals restricted to humid montane forests. *Proceedings of the Biological Society of Washington*, 115, 488-533.

Chippindale PT, Ammerman LK, Campbell JA (1998) Molecular approaches to phylogeny of *Abronia* (Anguidae: Gerrhonotinae), with emphasis on relationships in subgenus *Auriculabronia*. *Copeia*, 1998, 883-892.

Conant R, Collins JT (1998) Reptiles and Amphibians: Eastern/Central North America. Houghton Mifflin Co, Boston.

Cortés-Rodríguez N, Hernández-Baños BE, Navarro-Sigüenza AG, Peterson AT, García-Moreno J (2008) Phylogeography and population genetics of the Amethyst-throated Hummingbird (*Lampornis amethystinus*). *Molecular Phylogenetics and Evolution*, 48, 1-11.

Coyne JA, Orr HA (2004) Speciation. Sinauer Associates, Sunderland.
Crews SC & Hedin M (2006) Studies of morphological and molecular phylogenetic divergence in spiders (Araneae: Homalonychus) from the American southwest, including divergence along the Baja California Peninsula. *Molecular Phylogenetics and Evolution*, 38, 470-487.
Cullingham CI, Kyle CI, Pond BA, White BN (2008) Genetic structure of raccoons in eastern North America based on mtDNA: implications for subspecies designation and rabies disease dynamics. *Canadian Journal of Zoology*, 86, 947-958.
Devitt T (2006) Phylogeography of the Western Lyresnake (*Trimorphodon biscutatus*): testing aridland biogographical hypotheses across the Nearctic-Neotropical transition. *Molecular Ecology*, 15, 4387-4407.
Diamond JM (1977) Continental and insular speciation in pacific island birds. *Systematic Zoology*, 26, 263-268.
Dierickx EG, Shultz AJ, Sato F, Hiraoka T, Edwards SV (2015) Morphological and genomic comparisons of Hawaiian and Japanese Black-footed Albatrosses (Phoebastria nigripes) using double digest RADseq: implications for conservation. *Evolutionary Applications*, 8, 662-678.
Dobzhansky T (1940) Speciation as a stage in evolutionary divergence. *American Naturalist,* 74, 302-321.
Dodd CK, Barichivich WJ (2007) Movements of large snakes (*Drymarchon*, *Masticophis*) in north-central Florida. *Florida Scientist*, 70, 83-94.
Drummond AJ, Suchard MA, Xie D, Rambaut A (2012) Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Molecular Biology and Evolution*, 29, 1969-1.
Duennes MA, Lozier JD, Hines HM, Cameron SA (2012) Geographical patterns of genetic divergence in the widespread Mesoamerican bumble bee Bombus ephippiatus (Hymenoptera: Apidae). *Molecular Phylogenetics and Evolution*, 64, 219-231.
Eaton DAR (2014) PyRAD: assembly of de novo RADseq loci for phylogenetic analyses. *Bioinformatics,* btu121.
Eaton DAR, Ree R. (2013) Inferring phylogeny and introgression using RADseq Data: an example from flowering plants (Pedicularis: Orobanchaceae). *Systematic Biology*, 62, 689-706.
Earl D, vonHoldt B (2012) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources*, 4, 359-361.
Eckert, AJ, van Heerwaarden J, Wegrzyn JL, Nelson CD, Ross-Ibarra J, González-Martínez SC, Neale DB (2010) Patterns of population structure and environmental associations to aridity across the range of loblolly pine (Pinus taeda L., Pinaceae). *Genetics*, 185, 969-982.
Edgar R (2004) MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics*, 5, 113.
Egge JD, Hagbo TJ (2015) Comparative phylogeography of Mississippi embayment fishes. *PloS one,* 10, e0116719.
Emadzade K, Lebmann MJ, Hoffmann MH, Tkach N, Lone FA, Hörandl E (2015) Phylogenetic relationships and evolution of high mountain buttercups (*Ranuculus*) in North America and Central Asia. *Perspectives in Plant Ecology, Evolution and Systematics*, 17, 131-141.

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.

Feder LJ, Egan SP, Nosil P (2012) The genomics of speciation-with-gene-flow. *Trends in Genetics*, 28, 342-350.

Fontanella FM, Feldman CR, Siddall ME, Burbrink FT (2008) Phylogeography of *Diadophis punctatus*: extensive lineage diversity and repeated patterns of historical demography in a trans-continental snake. *Molecular Phylogenetics and Evolution*, 46, 1049-1070.

Franks SJ, Munshi-South J (2014) Go forth, evolve and prosper; the genetic basis of adaptive evolution in an invasive species. *Molecular Ecology*, 23, 2137-2140.

Futuyma DJ, Mayer GC (1980) Non-allopatric speciation in animals. *Systematic Biology*, 29, 254-271.

Gamble T, Berendzen PB, Shaffer HB, Starkey DE, Simons AM (2008) Species limits and phylogeography of North American cricket frogs (*Acris*: Hylidae). *Molecular Phylogenetics and Evolution*, 48, 112-125.

Geraci J, Béchet A, Cézilly F, Ficheux S, Baccetti N, Samraoui B, Wattier R (2012) Greater flamingo colonies around the Mediterranean form a single interbreeding population and share a common history. *Journal of Avian Biology*, 43, 341-354.

González –Porter GP, Maldonado JE, Flores-Villela O, Vogt RC, Janke A, Fleischer RC, Hailer F (2013) Cryptic population structuring and the role of the Isthmus of Tehuantepec as a gene flow barrier in the critically endangered Central American River Turtle. *PloS one*, 8, e71668.

González C, Ornelas JF, Gutiérrez-Rodríguez C (2011) Selection and geographic isolation influence hummingbird speciation: genetic, acoustic and morphological divergence in the wedge-tailed sabrewing (*Campylopterus curvipennis*). *BMC Evolutionary Biology*, 11, 38.

Grismer LL (1990) Relationships, taxonomy, and biogeography of the *Masticophis lateralis* complex in Baja California, Mexico. *Herpetologica*, 66-77.

Grismer LL (2000) Evolutionary biogeography on Mexico's Baja California peninsula: A synthesis of molecules and historical geology. *Proceedings of the National Academy of Sciences*, 97, 14017-14018.

Grummer JA, Calderón-Espinosa ML, Nieto-Montes de Oca A, Smith EN, Méndez-de la Cruz FR, Leaché AD (2015) Estimating the temporal and spatial extent of gene flow among sympatric lizard populations (genus *Sceloporus*) in the southern Mexican highlands. *Molecular Ecology*, 24, 1523-1542.

Gutiérrez-Rodríguez C, Ornelas JF, Rodríguez-Gómez F (2011) Chloroplast DNA phylogeography of a distylous shrub (Palicourea padifolia, Rubiaceae) reveals past fragmentation and demographic expansion in Mexican cloud forests. *Molecular Phylogenetics and Evolution*, 61, 603-615.

Hahn MW, Nakhleh L (2015) Irrational exuberance for resolved species trees. *Evolution*, 70, 7-17.

Hamilton CA, Formanowicz DR, Bond JE (2011) Species delimitation and phylogeography of *Aphonopelma hentzi* (Araneae, Mygalomorphae, Theraphosdae): Cryptic diversity in North American Tarantulas. *Plos One*, 6, 326207.

Havenor K (2003) The geological framework of the Pecos Valley and the evolution of the Roswell groundwater basin in Chaves and northern Eddy Counties, New Mexico. *AAPG Search and Discovery*, 90023.

Hirth HF, Pendleton RC, King AC, Downard TR (1969) Dispersal of snake from a hibernaculum in northwestern Utah. *Ecology*, 50, 332-339.

Halstead BJ, Mushinsky HR, McCoy ED (2009) *Masticophis flagellum* selects Florida scrub habitat at multiple spatial scales. *Herpetologica*, 65, 268-279.

Hedin M, Starret J, Hayashi C (2012) Crossing the uncrossable: novel trans-valley biogeographic patterns revealed in the genetic history of low-dispersal mygalomorph spiders (Antrodiaetidae, *Antrodiaetus*) from California. *Molecular Ecology,* 22, 508-526.

Hey J (2010) Isolation with migration models for more than two populations. *Molecular Biology and Evolution*, 27,905–920.

Hipp AL, Eaton DAR, Cavender-Bares J, Fitzek E, Nipper R, Manos PS (2014) A Framework Phylogeny of the American Oak Clade Based on Sequenced RAD Data. *PloS one,* 9, e93975.

Hoffman EA, Blouin MS (2004) Evolutionary history of the northern leopard frog: reconstruction of phylogeny, phylogeography, and historical changes in population demography from mitochondrial DNA. *Evolution*, 58, 145-159.

Holman JA (2000) Fossil snakes of North America, Origin, Evolution, Distribution, Paleoecology. Indiana University Press, Indianapolis.

Howes BJ, Lindsay B, Lougheed SC (2006) Range-wide phylogeography of a temperate lizard, the five-lined skink (*Eumeces fasciatus*). *Molecular Phylogenetics and Evolution*, 40, 183-194.

Huelsenbeck J, Ronquist F (2001) MRBAYES: Bayesian inference of phylogeny. *Bioinformatics*, 17, 754-755.

Hulsey CD, Garcí FJ, Johnson YS, Hendrickson DA, Near TJ (2004) Temporal diversification of Mesoamerican cichlid fishes across a major biogeographic boundary. *Molecular Phylogenetics and Evolution*, 31, 754-764.

Jaeger JR, Riddle BR, Bradford DF (2005) Cryptic Neogene vicariance and Quaternary dispersal of the red-spotted toad (*Bufo punctatus*): insights on the evolution of North American warm desert biotas. *Molecular Ecology*, 14, 3033-3048.

Johnson JD (1977) The taxonomy and distribution of the Neotropical whipsnake *Masticophis mentovarius* (Reptilia, Serpentes, Colubridae). *Journal of Herpetology*, 11, 287-309.

Jordan DS (1905) The origin of species through isolation. *Science*, 22, 545-562.

Kass RE, Raftery AE (1995) Bayes factors. *Journal of the American Statistical Association*, 90, 773-795.

Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A, Markowitz S, Duran C, Thierer T, Ashton B, Mentjies P, Drummond A (2012). Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics*, 28, 1647-1649.

Kirkpatrick M, Barton NH (1997) Evolution of a species' range. *The American Naturalist*, 150, 1–23.

Kraus RHS, van Hooft P, Megengs HJ, Tsvey A, Fokin SY, Ydenberg RC, Prins HHT (2012) Global lack of flyway structure in a cosmopolitan bird revealed by a genome wide survey of single nucleotide polymorphisms. *Molecular Ecology*, 22, 41-55.

Kuhner MK, Yamato J, Felsenstein J (1998) Maximum likelihood estimation of population growth rates based on the coalescent. *Genetics*, 149, 429-434.

Lamb T, Avise JC, Gibbons JW (1989) Phylogeographic patterns in mitochondrial DNA of the desert tortoise (*Xerobates agassizi*), and evolutionary relationships among the North American gopher tortoises. *Evolution*, 76-87.

Lamb T, Jones TR, Wettstein PJ (1997) Evolutionary genetics and phylogeography of tassel-eared squirrels (*Sciurus aberti*). *Journal of Mammology*, 78, 117-133.

Lande R (1980) Genetic variation and phenotypic evolution during allopatric speciation. *American Naturalist*, 116, 463-479.

Lanfear R, Calcott B, Ho SYW, Guindon S (2012) PartitionFinder: Combined Selection of Partitioning Schemes and Substitution Models for Phylogenetic Analysis. *Molecular Biology and Evolution*, 29, 1695-1701.

Lansman RA, Avise JC, Aquadro CF (1983) Extensive genetic variation in mitochondrial DNA's among geographic populations of the deer mouse, *Peromyscus maniculatus*. *Evolution*, 37, 1-16.

Latch EK, Heffelfinger JR, Fike JA, Rhodes OE (2009) Species-wide phylogeography of North American mule deer (*Odocoileus hemionus*): cryptic glacial refugia and postglacial recolonization. *Molecular Ecology,* 18, 1730-1745.

Leaché AD. (2009) Species Tree Discordance Traces to Phylogeographic Clade Boundaries in North American Fence Lizards (Sceloporus). *Systematic Biology*, 58, 547-559.

Leaché AD, McGuire JA (2006) Phylogenetic relationships of horned lizards (Phrynosoma) based on nuclear and mitochondrial data: evidence for a misleading mitochondrial gene tree. *Molecular Phylogenetics and Evolution*, 39, 628-644.

Leaché AD, Reeder TW (2002) Molecular systematics of the eastern fence lizard (*Sceloporus 51ndulates*): a comparison of parsimony, likelihood, and Bayesian approaches. *Systematic Biology*, 51, 44-68.

Leaché AD, Mulcahy DG (2007) Phylogeny, divergence times and species limits of spiny lizards (Sceloporus magister species group) in western North American deserts and Baja California. *Molecular Ecology*, 16, 5216-5233.

Lee TE, Riddle BR, Lee PL (1996) Speciation in the desert pocket mouse (*Chaetodipus penicillatus* Woodhouse). *Journal of Mammalogy*, 77, 58-68.

Lemmon EM, Lemmon AR, Collins JT, Lee-Yaw JA, Cannatella DC (2007) Phylogeny-based delimitation of species boundaries and contact zones in the trilling chorus frogs (*Pseudacris*). *Molecular Phylogenetics and Evolution*, 44, 1068-1082.

Lemmon EM, Lemmon AR, Cannatella DC (2007) Geological and climatic forces driving speciation in the continentally distributed trilling chorus frogs (*Pseudacris*). *Evolution*, 61, 2086-2103.

León-Paniagua L, Navarro-Sigüenza AG, Hernández-Baños BE, Morales JC (2007) Diversification of the arboreal mice of the genus Habromys (Rodentia: Cricetidae: Neotominae) in the Mesoamerican highlands. *Molecular Phylogenetics and Evolution,* 42, 653-664.

Lonsdale P (1991) Structural patterns on the Pacific floor offshore of peninsular California. In: The Gulf and Peninsular Province of the Californias. *American Association of Petroleum Geologists Memoir* (eds Dauphin P, Simoneit BRT), pp. 87–125.

Malaney JL, Conroy CJ, Moffit LA, Spoonhunter HD, Patton JL, Cook JA (2013) Phylogeography of the western jumping mouse (*Zapus princeps*) detects deep and persistent allopatry with expansion. *Journal of Mammology,* 94, 1016-1029.

Manos PS, Doyle JJ, Nixon KC (1999) Phylogeny, biogeography, and processes of molecular differentiation in Quercus subgenus Quercus (Fagaceae). *Molecular Phylogenetics and Evolution*, 12, 333-349.

Mantooth SJ, Hafner DJ, Bryson RW, Riddle BR (2013) Phylogeographic diversification of antelope squirrels (Ammospermophilus) across North American deserts. *Biological Journal of the Linnean Society*, 109, 949-967.

Masta SE, Sullivan BK, Lamb T, Routman EJ (2002) Molecular systematics, hybridization, and phylogeography of the *Bufo americanus* complex in Eastern North America. *Molecular Phylogenetics and Evolution*, 24, 302-314.

Mayr E (1942) Systematics and the Origin of Species. Columbia University Press, New York.

McCormack JE, Peterson AT, Bonaccorso E, Smith TB (2008) Speciation in the highlands of Mexico: genetic and phenotypic divergence in the Mexican jay (*Aphelocoma nalyseds*). *Molecular Ecology*, 17, 2505-2521.

McCormack JE, Heled J, Delaney KS, Peterson AT, Knowles LL (2011) Calibrating divergence times on species trees versus gene trees: implications for speciation history of Aphelocoma jays. *Evolution*, 65, 184-202.

Moore MJ, Jansen RK (2006) Molecular evidence for the age, origin, and evolutionary history of the American desert plant genus Tiquilia (Boraginaceae). *Molecular Phylogenetics and Evolution*, 39, 668-687.

Moriarty EC, Cannatella DC (2004) Phylogenetic relationships of the North American chorus frogs (*Pseudacris*: Hylidae). *Molecular Phylogenetics and Evolution*, 30, 409-420.

Morafka DJ (1977) *A biogeographical analysis of the Chihuahuan Desert through its herpetofauna*. Vol. 9. Springer Science & Business Media.

Mulcahy DG (2008) Phylogeography and species boundaries of the western North American Nightsnake (*Hypsiglena torquata*): revisiting the subspecies concept. *Molecular Phylogenetics and Evolution*, 46, 1095-1115.

Mulcahy DG, Macey JR (2009) Vicariance and dispersal form a ring distribution in nightsnakes around the Gulf of California. *Molecular Phylogenetics and Evolution*, 53, 537–546.

Mulcahy DG, Spaulding AW, Mendelson JR, Brodie ED (2006) Phylogeography of the flat-tailed horned lizard (Phrynosoma mcallii) and systematics of the P. mcallii–platyrhinos mtDNA complex. *Molecular Ecology*, 15, 1807-1826.

Mulcahy DG, Morrill BH, Mendelson JR (2006) Historical biogeography of lowland species of toads (Bufo) across the Trans-Mexican Neovolcanic Belt and the Isthmus of Tehuantepec. *Journal of Biogeography*, 33, 1889-1904.

Mulcahy DG, Macey RJ (2009) Vicariance and dispersal form a ring distribution in nightsnakes around the Gulf of California. *Molecular Phylogenetics and Evolution*, 53, 537-546.

Myers ED, Burgoon JL, Ray MJ, Martínez-Gómez JE, Matías-Ferrer N, Mulcahy DG, Burbrink FT. Coalescent species tree inference of *Coluber* and *Masticophis. In Press.*

Myers ED, Hickerson MJ, Burbrink FT (2017) Asynchronous diversification of snakes in the North American warm deserts. *Journal of Biogeography*, 44, 461-474.

Nason JD, Hamrick JL, Flaming TH (2002) Historical vicariance and postglacial colonization effects on the evolution of genetic structure in *Lophocereus*, a Sonoran Desert columnar cactus. *Evolution*, 56, 2214-2226.

Near TJ, Page LM, Mayden RL (2001) Intraspecific phylogeography of *Percina evides* (Percidae: Etheostomatinae): an additional test of the Central Highlands pre-Pleistocene vicariance hypothesis. *Molecular Ecology*, 10, 2235-2240.

Omland KE, Tarr CL, Boarman WI, Marzluff JM, Fleischer RC (2000) Cryptic genetic variation and paraphyly in ravens. *Proceedings of the Royal Society of London B: Biological Sciences*, 267, 2475-2482.

Orange DI, Riddle BR, Nickle DC (1999) Phylogeography of a wide-ranging desert lizard, *Gambelia wislizenii* (Crotaphytidae). *Copeia*, 1999, 267-273.

Ornelas JF, Rodríguez-Gómez F (2015) Influence of Pleistocene Glacial/Interglacial Cycles on the Genetic Structure of the Mistletoe Cactus *Rhipsalis baccifera* (Cactaceae) in Mesoamerica. *Journal of Heredity*, 106, 196-210.

Ornelas JF, Sosa V, Soltis, DE, Daza JM, González C, Soltis PS, Carla Gutiérrez-Rodríguez, Espinosa de los Monteros A, Castoe TA, Bell C, Ruiz-Sanchez E (2013) Comparative phylogeographic analyses illustrate the complex evolutionary history of threatened cloud forests of northern Mesoamerica. *PloS one*, 8, e56283.

Ortenburger AI (1923) A note on the genera *Coluber* and *Masticophis*, and a description of a new species of *Masticophis. Occasional Papers of the Museum of Zoology, University of Michigan*, 139, 1-14.

Pauly GB, Hillis DM, Cannatella DC (2004) The history of a nalysed colonization: molecular phylogenetics and biogeography of the nalysed toads (*Bufo*). *Evolution*, 58, 2517-2535.

Pelletier TA, Crisafulli C, Wagner S, Zellmer AJ, Carstens BC (2014) Historical Species Distribution Models Species Limits in Western *Plethodon* Salamanders. *Systematic Biology*, 64, 909-925.

Peterson BK, Weber JN, Kay EH, Fisher HS, Hoekstra HE (2012) Double digest RADseq: an inexpensive method for de novo SNP discovery and genotyping in model and non-model species. *PloS one*, 7, e37135.

Petkova D, Novembre J, Stephens M (2016) Visualizing spatial population structure with estimated effective migration surfaces. *Nature Genetics*, 48, 94-100.

Pook CE, Wüster W, Thorpe RS (2000). Historical Biogeography of the Western Rattlesnake (Serpentes: Viperidae: *Crotalus viridis*), Inferred from Mitochondrial DNA Sequence Information. *Molecular Phylogenetics and Evolution*, 15, 269–282.

Pfennig KS (2003) A test of alternative hypotheses for the evolution of reproductive isolation between spadefoot toads: support for the reinforcement hypothesis. *Evolution*, 57, 2842-2851.

Pritchard JK, Stephens M, Donnelly P (2000) Inference of Population Structure Using Multilocus Genotype Data. *Genetics*, 155, 945-959.

Puechmaille SJ (2016) The program STRUCTURE does not reliably recover the correct population structure when sampling is uneven: sub-sampling and new estimators alleviate the problem. *Molecular Ecology Resources*, 16, 608-627.

Pyron RA, Burbrink FT (2009) Lineage diversification in a widespread species: roles for niche divergence and conservatism in the common kingsnake, Lampropeltis getula. *Molecular Ecology*, 18, 3443-3457.

Pyron RA Burbrink FT (2010). Hard and soft allopatry: physically and ecologically mediated modes of geographic speciation. *Journal of Biogeography*, 37, 2005-2015.

Pyron RA, Burbrink FT, Colli CR, Nieto-Montes de Oca A, Vitt LJ, Kuczynki CA, Wiens JJ (2011) The phylogeny of advanced snakes (Colubroidea) with discovery of a new subfamily and comparison of support methods for likelihood trees. *Molecular Phylogenetics and Evolution*, 58, 329-342.

Pyron RA, Burbrink FT, Wiens JJ (2013). A phylogeny and revised classification of Squamata, including 4161 species of lizards and snakes. *BMC Evolutionary Biology*, 13, 93.

Ray JM, Wood RM, Simons, AM (2006) Phylogeography and post-glacial colonization patterns of the rainbow darter, *Etheostoma caeruleum* (Teleostei: Percidae). *Journal of Biogeography*, 33, 1550-1558.

Rebernig CA, Schneeweiss GM, Bardy KE, Schoenswetter P, Villasenor JL, Obermayer R, Stuessy TF, Weiss-Schneeweis H (2010). Multiple Pleistocene refugia and Holocene range expansion of an abundant southwestern American desert plant species (*Melampodium leucanthum*, Asteraceae). *Molecular Ecology*, 19, 3421-3443.

Remington CL (1968) Suture-zones of hybrid interaction between recently joined biotas. Pages 321–428 in T. Dobzhansky, M. K. Hecht, and W. C. Steere, eds. Evolutionary biology. Plenum, New York.

Rambaut A, Drummond AJ (2007) Tracer v1.4, Available from http://beast.bio.ed.ac.uk/Tracer.

Richmond JQ, Wood DA, Hoang C, Vandergast AG (2011) Quantitative assessment of population genetic structure and historical phylogeography of the Alameda Whipsnake *Masticophis lateralis euryxanthus*. U.S. Geological Survey, Western Ecological Research Center. Data Summary prepared for USFWS Sacramento FWO. 81 pp.

Riddle BR (1995) Molecular biogeography in the pocket mice (Perognathus and Chaetodipus) and grasshopper mice (Onychomys): the late Cenozoic development of a North American aridlands rodent guild. *Journal of Mammalogy*, 76, 283-301.

Riddle BR, Hafner DJ, Alexander LF, Jaeger JR (2000) Cryptic vicariance in the historical assembly of a Baja California Peninsular Desert biota. *Proceedings of the National Academy of Sciences*, 97, 14438-14443.

Riddle BR, Honeycutt RL (1990) Historical biogeography of North American arid regions: An approach using mitochondrial-DNA phylogeny in Grasshopper Mice (Genus *onychomys*). *Evolution*, 44, 1-15.

Riddle BR, Hafner DJ (2006) A step-wise approach to integrating phylogeographic and phylogenetic biogeographic perspectives on the history of a core North American warm deserts biota. *Journal of Arid Environments*, 66, 435-461.

Rodriguez M, Rau D, Bitocchi E, Bellucci E, Biagetti E, Carboni A, Gepts P, Nanni L, Papa R, Attene G (2016) Landscape genetics, adaptive diversity and population structure in Phaseolus vulgaris. *New Phytologist*, 209, 1781-1794.

Rodriguez RM, Ammerman LK (2004) Mitochondrial DNA divergence does not reflect morphological difference between *Myotis californicus* and *Myotis ciliolabrum*. *Journal of Mammalogy*, 85, 842-851.

Rodríguez-Robles JA, De Jesús-Escobar JM (2000) Molecular systematics of New World gopher, bull, and pinesnakes (Pituophis: Colubridae), a transcontinental species complex. *Molecular Phylogenetics and Evolution*, 14, 35-50.

Rodríguez-Gómez F, Gutiérrez-Rodríguez C, Ornelas JF (2013) Genetic, phenotypic and ecological divergence with gene flow at the Isthmus of Tehuantepec: the case of the azure-crowned hummingbird (*Amazilia cyanocephala*). *Journal of Biogeography*, 40, 1360-1373.

Rohland N, Reich D (2012) Cost-effective, high-throughput DNA sequencing libraries for multiplexed target capture. *Genome research*, 22, 939-946.

Rosenthal J, Forstner MRJ (2004) Effects of Plio-Pleistocene Barrier on Chihuahuan Desert Herpetofauna. *Proceedings of the Sixth Symposium on the Natural Resources of the Chihuahuan Desert Region*, 269-282.

Rothfels CJ, Larsson A, Kuo, LY, Korall P, Chiou WL, Pryer KM (2012) Overcoming deep roots, fast rates, and short internodes to resolve the ancient rapid radiation of eupolypod II ferns. *Systematic Biology*, 61, 490-509.

Rovito SM, Parra-Olea G, Vásquez-Almazán CR, Luna-Reyes R, Wake DB (2012) Deep divergences and extensive phylogeographic structure in a clade of lowland tropical salamanders. *BMC Evolutionary Biology*, 12.

RoyChoudhury A, Felsenstein J, Thompson EA (2008) A two-stage pruning algorithm for likelihood computation for a population tree. *Genetics*, 180, 1095-1105.

Roze JA (1953) The *Rassenkreis Coluber* (*Masticophis*) *mentovarius* (Dumeril, Bibron et Dumeril), 1854. *Herpetologica*, 9, 113-120.

Ruane S, Bryson R, Pyron A, Burbrink F (2014) Coalescent species delimitation in Milksnakes (Genus *Lampropeltis*) and impacts on phylogenetic comparative analyses. *Systematic Biology*, 63, 231-250.

Sambrook J, Russell D (2001) Molecular Cloning: A Laboratory Manual, 3rd Edition. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.

Satler JD, Carstens BC, Hedin M (2013) Multilocus Species Delimitation in a Complex or Morphologically Conserved Trapdoor Spiders (Mygalomorphae, Antrodiaetidae, *Aliatypus*). *Molecular Ecology,* 62, 805-823.

Schield DR, Card DC, Adams RH, Jezkova T, Reyes-Valasco J, Proctor FN, Specer CL, Herrmann HW, Mackessy SP, Castoe TA (2015) Incipient speciation with biased gene flow between two lineages of the Western Diamondback Rattlesnake (*Crotalus atrox*). *Molecular Phylogenetics and Evolution,* 83, 213-223.

Silvestro M, Michalak I (2012). raxmlGUI: a graphical front-end for RaxML. *Organisms Diversity and Evolution*, 12, 335-337.

Sinclair EA, Bezy RL, Bolles K, Camarillo RJL, Crandall KA, Sites JW (2004) Testing species boundaries in an ancient species complex with deep phylogeographic history: genus Xantusia (Squamata: Xantusiidae). *The American Naturalist*, 164, 396-414.

Serb JM, Phillips CA, Iverson JB (2001) Molecular phylogeny and biogeography of *Kinosternon flavescens* based on complete mitochondrial control region sequences. *Molecular Phylogenetics and Evolution*, 18, 149-162.

Smith MF (1998) Phylogenetic relationships and geographic structure in pocket gophers in the genus *Thomomys*. *Molecular Phylogenetics and Evolution*, 9, 1-14.

Smith CI, Farrell BD (2005) Phylogeography of the longhorn cactus beetle *Moneilema appressum LeConte* (Coleoptera: Cerambycidae): was the differentiation of the Madrean sky islands driven by Pleistocene climate changes? *Molecular Ecology*, 14, 3049-3065.

Soltis DE, Morris AB, McLachlan JS, Manos PS, Soltis PS (2006) Comparative phylogeography of unglaciated eastern North America. *Molecular Ecology*, 15, 4261–4293.

Stebbins RC (2003) Western Reptiles and Amphibians. Houghton Mifflin Co, New York.

Steeves TE, Anderson DJ, Friesen VL (2005) The Isthmus of Panama: a major physical barrier to gene flow in a highly mobile pantropical seabird. *Journal of Evolutionary Biology*, 18, 1000–1008.

Strange RM, Burr BM (1997) Intraspecific phylogeography of North American highland fishes: a test of the Pleistocene vicariance hypothesis. *Evolution*, 51, 885-897.

Streicher JW, Cox CL, Campbell JA, Smith EN, de Sá RO (2012) Rapid range expansion in the Great Plains narrow-mouthed toad (Gastrophryne olivacea) and a revised taxonomy for North American microhylids. *Molecular Phylogenetics and Evolution*, 64, 645-653.

Streicher JW, Devitt TJ, Goldberg CS, Malone JH, Blackmon H, Fujita MK (2014) Diversification and asymmetrical gene flow across time and space: lineage sorting and hybridization in polytypic barking frogs. *Molecular Ecology*, 23, 3273-3291.

Streicher JW, McEntee JP, Drzich LC, Card DC, Schield DR, Smart U, Parkinson CL, Jezkova T, Smith EN, Castoe TA (2016) Genetic surfing, not allopatric divergence, explains spatial sorting of mitochondrial haplotypes in venomous coralsnakes. *Evolution,* 70, 1435-1449.

Suárez-Atilano M, Burbrink F, Vázquez-Domínguez E (2014) Phylogeographical structure within Boa constrictor imperator across the lowlands and mountains of Central America and Mexico. *Journal of Biogeography*, 41, 2371-2384.

Sullivan J, Arellano E, Rogers DS (2000) Comparative phylogeography of Mesoamerican highland rodents: concerted versus independent response to past climatic fluctuations. *The American Naturalist*, 155, 755-768.

Sullivan J, Markert JA, Kilpatrick CW (1997) Phylogeography and molecular systematics of the Peromyscus aztecus species group (Rodentia: Muridae) inferred using parsimony and likelihood. *Systematic Biology*, 46, 426-440.

Sutherland GD, Harestad AS, Price K, Lertzman KP (2000) Scaling of natal dispersal distances in terrestrial birds and mammals. *Ecology and Society*, 4, 16.

Swenson NG, Howard DJ (2005) Clustering of Contact Zones, Hybrid Zones, and Phylogeographic Breaks in North America. *The American Naturalist*, 166, 581–591.

Tamura K, Stecher G, Peterson D, Filipski A, Kumar S (2013) Mega6: Molecular Evolutionary Genetics Analysis Version 6.0. *Molecular Biology and Evolution,* 30, 2725-2729.

Templeton AR, Routman E, Phillips CA (1995) Separating population structure from population history: a nalyseds analysis of the geographical distribution of mitochondrial DNA haplotypes in the tiger salamander, *Ambystoma tigrinum*. *Genetics*, 140, 767-782.

Uetz P and Hošek J. The Reptile Database, http://www.reptile-database.org, accessed Jan 8, 2014

Unger SD, Rhodes OE, Sutton TM, Williams RN (2013) Population Genetics of the Eastern Hellbender (*Cyrptobranchus alleganiensis alleganiensis*) across Multiple Spatial Scales*,* 8.10: e74180.

Utiger U, Schatti B, Helfenberger N (2005) The oriental colubrine genus Coelognathus FITZINGER, 1843 and classification of old and new world racers and ratsnakes (Reptilia, Squamata, Colubridae, Colubrinae). *Russian Journal of Herpetology,* 12, 39-60.

Unmack PJ, Dowling TE, Laitinen NJ, Secor CL, Mayden RL, Shiozawa DK, Smith GR (2014) Influence of introgression and geological processes on phylogenetic relationships of western North American mountain suckers (Pantosteus, Catostomidae). *PloS one*, *9*, e90061.

Walpole DK, Davis SK, Greenbaum IF (1997) Variation in mitochondrial DNA in populations of Peromyscus eremicus from the Chihuahuan and Sonoran deserts. *Journal of Mammalogy*, 78, 397-404.

Walker D, Moler PE, Buhlmann KA, Avise JC (1998) Phylogeographic patterns in Kinosternon subrubrum and K. baurii based on mitochondrial DNA restriction analyses. *Herpetologica*, 174-184.

Walker D, Avise JC (1998) Principles of phylogeography as illustrated by freshwater and terrestrial turtles in the southeastern United States. *Annual Review of Ecology and Systematics*, 29, 23-58.

Wagner CE, Keller I, Wittwer S, Selz OM, Mwaiko S, Greuter L, Sivasundar A, Seehausen O (2013) Genome-wide RAD sequence data provide unprecedented resolution of species boundaries and relationships in the Lake Victoria cichlid adaptive radiation. *Molecular Ecology*, 22, 787-798.

Weisrock DW, Janzen FJ (2000) Comparative molecular phylogeography of North American softshell turtles (Apalone): implications for regional and wide-scale historical evolutionary forces. *Molecular Phylogenetics and Evolution*, 14, 152-164.

Whorley JR, Alvarez-Casteneda ST, Kenagy GJ (2004) Genetic structure of desert ground squirrels over a 20-degree-latitude transect from Oregon through the Baja California peninsula. *Molecular Ecology*, 13, 2709-2720.

Wilson LD (1970) The coachwhip snake *Masticophis flagellum* (Shaw): Taxonomy and distribution. *Tulane Studies in Zoology and Botony,* 16, 31-99.

Wilson JS, Pitts JP (2010a) Illuminating the lack of consensus among descriptions of earth history data in the North American deserts: a resource for biologists. *Progress in Physical Geography*, 34, 419–441.

Wilson JS, Pitts JP (2010b) Phylogeographic analysis of the nocturnal velvet ant genus *Dilophotopsis* (Hymenoptera: Mutillidae) provides insights into diversification in the Nearctic deserts. *Biological Journal of the Linnean Society*, 101, 360-375.

Withrow JJ, Sealy SG, Winker K (2014) Genetics of divergence in the Northern Saw-whet Owl (Aegolius acadicus). *The Auk*, 131, 73-85.

Zaldívar-Riverón A, León-Regagnon V, de Oca ANM (2004) Phylogeny of the Mexican coastal leopard frogs of the *Rana berlandieri* group based on mtDNA sequences. *Molecular Phylogenetics and Evolution*, 30, 38-49.

Zamudio KR, Savage WK (2003) Historical isolation, range expansion, and secondary contact of two highly divergent mitochondrial lineages in spotted salamanders (*Ambystoma maculatum*). *Evolution*, 57, 1631-1652.

Zamudio-Beltrán LE, Hernández-Baños BE (2015) A multilocus analysis provides evidence for more than one species within Eugenes fulgens (Aves: Trochilidae). *Molecular Phylogenetics and Evolution*, 90, 80-84.

Zink RM, Blackwell RC (1998) Molecular systematics and biogeography of aridland gnatcatchers (genus *Polioptila*), and evidence supporting species status of the California Gnatcatcher (*P. californica*). *Molecular Phylogenetics and Evolution*, 9, 26-32.

Zink RM, Kessen AE, Line TV, Blackwell-Rago RC (2001) Comparative phylogeography of some aridland bird species. *Condor*, 103, 1-10.

Zink RM (2014) Homage to Hutchinson, and the role of ecology in lineage divergence and speciation. *Journal of biogeography*, 41, 999-1006.

DATA ACCESSIBILITY

Mitochondrial sequence data is available on Genbank with accessions KT713652-KT713738. Genomic sequences are available on the Sequence Read Archive SRS1047195-SRS1047343. See Table A2.1 for full details.

THE EFFECT OF MISSING DATA ON COALESCENT SPECIES DELIMITATION AND A TAXONOMIC REVISION OF WHIPSNAKES (COLUBRIDAE: *COLUBER*)

Kyle A. O'Connell*[1,2], Eric N. Smith[1,2]

[1]*Department of Biology, The University of Texas at Arlington, Arlington, Texas 76019, USA.*
[2]*The Amphibian and Reptile Diversity Research Center; University of Texas at Arlington, Arlington, Texas 76010, USA.*

Submitted, 2017

CHAPTER 3

THE EFFECT OF MISSING DATA ON COALESCENT SPECIES DELIMITATION AND A TAXONOMIC REVISION OF WHIPSNAKES (COLUBRIDAE: *COLUBER*)

ABSTRACT

A stable alpha taxonomy is essential to understanding evolutionary processes and achieving effective conservation aims. Taxonomy depends on the identification of independently evolving lineages, and the delimitation of these lineages based on multiple lines of evidence. Coalescent species delimitation within an integrative framework has increased the rigor of the delimitation process. Here we use genome-wide SNP data and coalescent species delimitation to explore lineage relationships within several North American whipsnake species, and to test the species status of several lineages. We find support for the elevation of previous subspecies to full species status, and formally elevate two species. This study demonstrates the power of molecular data, paired with model-based delimitation methods, to identify evolutionary relationships, and to delimit previously overlooked species within well-studied taxa.

INTRODUCTION

The field of species delimitation has received increased attention in recent years (Sites & Marshall, 2003). Since the foundational work of De Queiroz (2007), the definition of the general lineage species concept has decoupled species conceptualization from species delimitation. As such, various lines of evidence can be used to assess lineage independence, but the status of the species is not dependent on any one type of evidence (De Queiroz, 2007). In the pre-molecular era, species delimitation primarily depended on morphological data, although ecological,

distributional, or other types of data have been used support a species' status (Padial et al., 2010; Sites & Marshal, 2004). In the case of allopatrically distributed species, reproductive isolation is demonstrable, but in species with overlapping ranges, researchers traditionally relied on morphological differences as a proxy for reproductive isolation (Fujita et al., 2012). However, morphological or ecological variation may not accurately represent the evolutionary history of a species (Ruane et al., 2013). The advent of molecular data revolutionized taxonomy and species delimitation, but a dependence on single locus data can often mislead inferences regarding the number of species, or the relationships of those species, due to incomplete lineage sorting and hybridization (Knowles & Carstens, 2007; Streicher et al., 2016). Fortunately, genomic data, and the subsequent increase in available loci, have helped to mitigate many of these shortcomings by estimating species trees more accurately, and by allowing for robust model testing of species hypotheses (Liu et al., 2015; Leaché et al., 2014; Faircloth et al., 2012).

Species delimitation methods attempt to accurately quantify independently evolving lineages (Knowles & Carstens, 2007; Petit & Excoffier, 2009; Sites & Marshall, 2003). The species delimitation process is comprised of two steps: lineage identification, and hypothesis testing (Reid & Satler, 2013). Lineage identification relies on a variety of methods, including morphological or ecological variation, disjunct geographic distributions, or molecular phylogenies (Wiens, 2007). However, lineages identified by one or more of these methods may not reflect the accurate evolutionary history of lineages, creating the need to test hypotheses regarding species composition and relationships (Fontaneto et al., 2015). Several recent techniques leverage coalescent theory to test species delimitation hypotheses (Fujita et al., 2012; Pante et al., 2015). Bayes Factor Delimitation (with genomic data; BFD*) is one method for testing hypotheses of species relationships that utilizes genome-wide SNP data (Leaché et al.,

2014). This method is advantageous to other coalescent species delimitation methods because it does not require a guide tree, but rather, it iterates over gene trees to directly estimate the species tree, and to calculate a marginal likelihood estimate (MLE) for each species model. It also accommodates genomic SNP data, is robust to missing data, and does not require a minimum number of individuals per species (Leaché et al., 2014).

Here, we utilize a species delimitation framework with molecular data to explore species relationships within North American whipsnakes, and to test the validity of several subspecies. Whipsnakes are a widespread clade of slender colubrid snakes distributed across North America into Columbia and Venezuela (Johnson, 1977; Wilson, 1970; Fig. 3.1A). These species were previously classified as the genus *Masticophis* (Johnson, 1977), but recent phylogenetic analyses showed that *Masticophis* was paraphyletic with respect to *Coluber constrictor* (Burbrink & Myers, 2015, but see O'Connell et al., 2017; Myers et al., In Press). As such, *Masticophis* is recognized as a junior synonym of *Coluber* (Uetz & Hošek, 2017). This study focuses on the systematics of three species of whipsnakes, *C. flagellum, C. bilineatus* and *C. mentovarius*. *Coluber flagellum* (Shaw, 1802) is a large bodied snake distributed across North America, and across several diverse ecoregions (Roze, 1953; Johnson, 1977; Conant & Collins, 1998; Uetz & Hošek, 2017). Color pattern and scale count variation is dramatic in this species group across geographic space (Wilson, 1970), leading to the recognition of six subspecies for *C. flagellum* (Fig. 3.1A): 1) *C. f. flagellum* (Type locality, Carolina and Virginia), 2) *C. f. testaceus* (Type locality, Pueblo County, Colorado, USA), 3) *C. f. lineatulus* (Type locality, Chihuahua, MX), 4) *C. f. cingulum*, (Type locality, Moctezuma, Sonora, MX), 5) *C. f. piceus* (Type locality, Graham County, Arizona, USA), and 6) *C. f. ruddocki* (Type locality, Kern County, California, USA). *Coluber fuliginosus* was classified as a seventh subspecies of *C. flagellum* from Baja California,

Mexico, until Grismer (1994) elevated it to evolutionary species status. In this study, we refer to *C. fuliginosus* as part of the *C. flagellum* group, but retain its classification as a separate species.

The second species we investigate, *C. mentovarius,* has experienced a very complex taxonomic history. This species currently encompasses five recognized subspecies, including *C. m. mentovarius* (Type locality, Chisec, Guatemala), *C. m. centralis* (Type locality, Guajira, Colombia), *C. m. suborbitalis* (Type locality, Caracas, Venezuela), *C. m. striolatus* (Type locality, Colima, Mexico), and *C. m. variolosus* (Type locality, Maria Magdalena Island, Mexico). However, Peters & Orejas-Miranda (1970) only listed three subspecies (*Masticophis m. mentovarius*, *M. m. centralis*, and *C. m. suborbitalis*), considering all Mexican *M. mentovarius* to pertain to *M. m. mentovarius*. Zweifel (1960) considered whipsnakes from the Tres Marias Islands to belong to *M. lineatus.* He recommended synonymizing *Masticophis striolatus*, and *M. variolosus*, now both recognized as subspecies of *C. mentovarius*, as *M. lineatus*.

Finally, *Coluber bilineatus* inhabits a much smaller geographic range than both *C. flagellum* and *C. mentovarius;* it is primarily restricted to the Sonoran Desert. Originally, *C. bilineatus* was divided into two subspecies. The first, *C. b. slevini* is from the Isla San Esteban in the Gulf of California, Sonora, Mexico (Lowe & Norris, 1955), but was elevated to species status by Grismer (1999). The second subspecies was *C. b. lineolatus* (Hensley, 1950), which has since been synonymized with *C. bilineatus*. The ranges of *C. bilineatus* and several *C. flagellum* lineages overlap in the Cochise Filter Barrier, and Stebbins (2003), expressed uncertainty of subspecies limits of *C. flagellum* in this region, where *C. f. testaceus*, *C. f. piceus*, *C. f. cingulum*, and *C. bilineatus* all overlap (Fig. 3.1A).

Here we explore the number of lineages within *C. flagellum*, *C. mentovarius*, *and C. bilineatus* using mitochondrial and genomic data, and conduct species delimitation using BFD*. We also collect morphological data for each lineage. We use these data to address four longstanding questions in whipsnake systematics: 1) Do subspecies of *C. flagellum* represent independently evolving lineages? 2) Which subspecies are present in and around the Cochise Filter Barrier? 3) What lineages pertain to the *Coluber* of western Mexico? 4) How well does morphological variation correspond to phylogenetic structure?

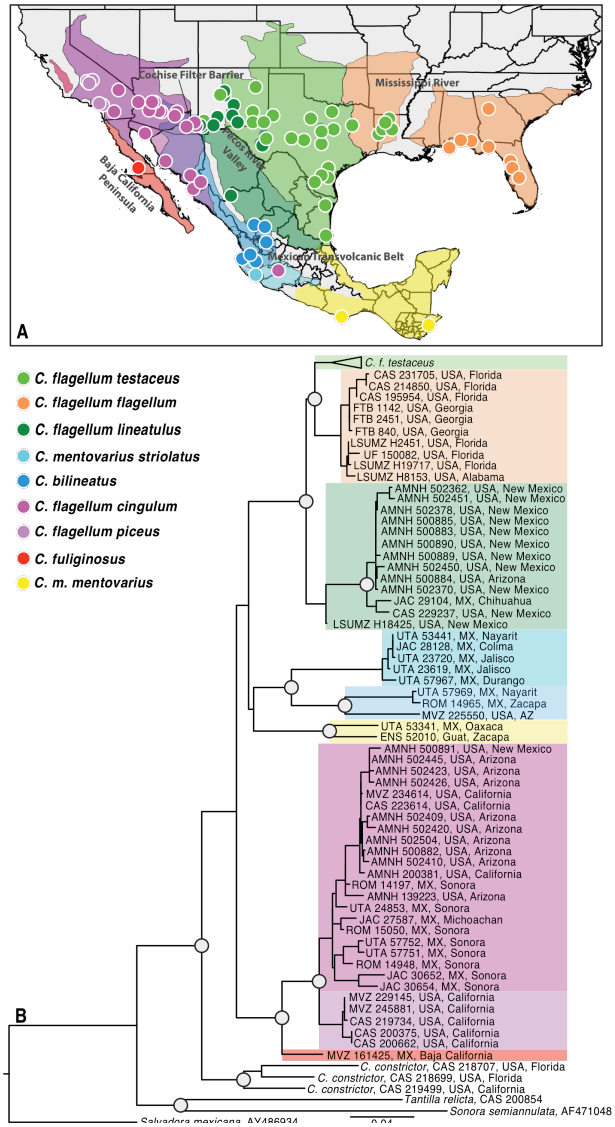

Fig. 3.1: A) Map showing the approximate distributional ranges of each subspecies investigated in this study as described by Wilson (1970) and Stebbins (2003). Circles represent sampling localities for mitochondrial data. B) Maximum likelihood phylogeny including several species of whipsnakes. Grey circles show nodes with at least 70% bootstrap support. Colors on each clade correspond to the colors used in the range map. We collapsed the clade pertaining to *C. flagellum testaceus* to save space; the full phylogeny is shown in Fig. A3.1.

MATERIALS AND METHODS

Mitochondrial sequencing and phylogenetic analyses

We utilized 65 mitochondrial sequences of the cytochrome B oxidase gene from O'Connell *et al.,* (2017), available on GenBank (KT713652-KT713738), as well as 46 additional cytochrome B sequences that we downloaded from Genbank (Table A3.1. Sequences used in this study included *C. f.*

flagellum (n = 10), *C. f. testaceus* (n = 44), *C. f. lineatulus* (n = 13), *C. f. cingulum* (n = 22), *C. f. piceus* (n = 5), *C. fuliginosus* (n = 1), *C. bilineatus* (n = 3), *C. m*. striolatus (n = 5), and *C. m. mentovarius* (n = 2), *Coluber constrictor* (n = 3), *Salvadora mexicana* (n = 1), *Tantilla relicta* (n = 1), and *Sonora semiannulata* (n = 1; Table A4.2). We aligned all sequences with the Geneious Aligner using default settings (Kearse et al., 2012). We calculated uncorrected average pairwise distance between lineages in Mega v7 (Kumar et al., 2016). We selected the most probable model of nucleotide evolution for Likelihood analyses using Bayesian information criteria implemented in PartitionFinder (Lanfear et al., 2012), partitioning by codon position. We estimated a maximum likelihood phylogeny using raxmlGUI v1.3 with 1000 rapid bootstrap iterations (Silvestro & Michalak, 2012) and visualized our final phylogeny in FigTree v.1.4.3 (Rambaut, 2017). We considered nodes with bootstrap values $\geq$ 70 as strongly supported.

Genomic sequence generation and computational analysis

We utilized double-digest restriction associated DNA sequencing (ddRADseq) data for 49 individuals from O'Connell et al., (2017) to evaluate relationships between mitochondrial lineages using nuclear data. Our sampling included *C. f. flagellum* (n = 6), *C. f. testaceus* (n = 6), *C. f. lineatulus* (n = 4), C. *fuliginosus* (n = 1), and *C. bilineatus* (n = 1), *C. f. cingulum* (n = 3), *C. m. striolatus* (n = 3), and *C. m. mentovarius* (n = 6).

We processed our RAD data using the STACKS v1.12 pipeline (Catchen et al., 2013). We followed the recommended workflow which implemented the following scripts and programs: (i) process_radtags, which filtered out reads below 90% quality score threshold, (ii) ustacks, which set a maximum distance of 4 between 'stacks', (iii) cstacks, which creates a catalogue of all loci within all individuals (-n flag; setting of 0) (iv) sstacks, which searches the

stacks created in ustacks against the catalogue from cstacks, and (v) populations, which genotypes each individual according to the matched loci from sstacks. After running populations, we used custom python scripts (available at https://github.com/dportik/Stacks_pipeline) to filter out invariant loci, and loci with more than two haplotypes. We produced several SNP datasets that differed in the species included as well as the percent missing data (Table A4.2). Dataset A included between one to three individuals from all species in our study. We limited the number of individuals to maximize taxonomic diversity while minimizing allelic dropout. Thus, dataset A included 15 individuals and 365 loci, including three *C. f. cingulum*, *C. m. striolatus*, *C. m. mentovarius*, and one *C. fuliginosus*, *C. f. flagellum*, *C. f. testaceus*, *C. f. lineatulus*, and *C. bilineatus*. We allowed up to 30% missing data per locus. Next we created species-specific datasets that used two different missing data thresholds to test species limits within *C. flagellum* (datasets B–C), and within *C. mentovarius* and *C. bilineatus* (datasets D-E). Dataset B included four *C. f. lineatulus*, six *C. f. flagellum*, six *C. f. testaceus*, three *C. f. cingulum*, one *C. fuliginosus*, and six *C. m. mentovarius*. We allowed up to 50% missing loci, resulting in 2079 loci. We also created dataset C with the same individuals that only allowed up to 20% missing loci resulting in 325 loci. Dataset D included three *C. m. mentovarius*, three *C. m. striolatus*, and one *C. f. flagellum*, *C. f. lineatulus*, *C. f. testaceus*, and *C. f. bilineatus*. We allowed up to 50% missing loci, resulting in 1464 loci. Finally, dataset E included the same 10 individuals, but allowed up to 20% missing loci resulting in 216 loci (Table A4.2.

Investigating species relationships within Coluber using neighbor networks

We investigated phylogenetic relationships between all study species using dataset A in SPLITSTREE v4.13.1 (Huson and Bryant, 2006; Fig. 3.2). SPLITSTREE4 uses a distance-based

method to estimate a neighbor network, rather than estimating a strict phylogeny. We used default settings and visualized the network using EqualAngle distances.

Coalescent species delimitation and species tree estimation

Our mitochondrial and nuclear phylogenetic analyses identified several lineages that may represent evolutionary species, but have historically been recognized as subspecies. To test if these lineages should be elevated to species status, we conducted Bayes Factor Delimitation with genomic data following Leaché et al., (2014; BFD*) and Grummer et al., (2013; BFD). Bayes Factor delimitation utilizes the SNAPP (Bryant et al., 2012) plugin of the BEAST2 platform (Kühnert et al., 2014) to calculate a MLE for alternative models using path sampling. One advantage of this method for SNP data is that it can accommodate missing data between individuals (among species), and allows for varying numbers of individuals per species. Using a Bayes Factor (BF), we were able to compare and rank models to determine the best-supported species hypothesis. We calculated the BF by subtracting the absolute value of the MLE of the model representing the current taxonomic classification of each dataset from each alternative model. Following Kass and Raftery (1995) we considered a BF over 10 to provide strong support for a model. We subsequently ranked each model and chose the model with the highest BF (Table 3.1). In addition to testing lineage limits, we wanted to test the effect of missing data and locus type on species delimitation. Thus, we conducted four sets of analyses, two for each species group using the $\leq$ 50% and 20% missing loci datasets (Table 3.1; Fig. 3.3). We assigned individuals to lineages based on the classifications of O'Connell et al., (2017), and confirmed these classifications based on additional SPLITSTREE analyses (results not shown). Our first two analyses tested species limits within *C. flagellum* using datasets B and C (Fig. 3.3). We

tested the following models: a) current taxonomy, all *C. flagellum* were lumped, and split from *C. fuliginosus* and *C. m. mentovarius*, b) lumped all *C. flagellum* with *C. fuliginosus*, split *C. mentovarius*, c) split *C. f. cingulum*, *C. fuliginosus* and *C. m. mentovarius,* lumped *C. f. flagellum*, *C. f. testaceus*, and *C. f. lineatulus*, d) lump *C. f. cingulum* with *C. fuliginosus*, lump all other *C. flagellum*, split *C. m. mentovarius*, e) lump *C. f. testaceus* with *C. f. lineatulus*, split all other lineages, f) lump *C. f. flagellum* with *C. f. testaceus*, split all other lineages, g) split all lineages, h) split all lineages, but mix *C. f. flagellum*, *C. f. testaceus*, *C. f. lineatulus*, and *C. f. cingulum* randomly. Our second set of analyses utilized datasets D and E and tested the following models (Fig. 3.3): a) current taxonomy, where we split *C. flagellum*, *C. m. mentovarius*, *C. m. striolatus*, and *C. bilineatus*, b) lump *C. m. striolatus* and *C. bilineatus*, c) lump *C. m. striolatus* with *C. m. mentovarius*, d) lump *C. m. striolatus* with *C. flagellum*, e) split all lineages but mix them randomly. We allowed BEAUti to estimate the mutation rate, and confirmed that both U and V were approximately equal to one. We assigned a Gamma distribution to our Lambda prior, with an Alpha of 1 and a Beta of 77. On our Snap prior we assigned an Alpha of 1, a Beta of 100, and a Lambda of 77. We performed 48 path sampling steps, with 100 000 MCMC generations, and 10,000 burnin generations. We calculated the Bayes Factor by subtracting the absolute value of the MLE of the all models from the current taxonomic classification (Model A).

We estimated the species tree for each dataset using SNAPPv1.0. We assigned species identities based on the best supported model from our BFD* analysis. We utilized the same parameters as above, but we ran our analyses for 10 000 000 MCMC generations, sampling

every 1000 generations. We visualized our complete tree sets in DENSITREE (Bouckaert, 2010), and removed the first 10% of trees as burn-in.

Table 3.1. Bayes Factor delimitation results are shown for each analysis. The number of species represents the number of species included in each analysis after lumping or splitting lineages. The number of loci represents the number of loci shared between all species in each analysis.

Model	Species	Loci	MLE	BF	Rank
Analysis 1, *Coluber flagellum cingulum* & *C. fuliginosus* 50% missing data					
a. Current taxonomy	3	832	7992.089	-	5
b. lump *C. flagellum* with *C. fuliginosus*	2	1409	13807.758	-5815.669	8
c. split off *C. f. cingulum*	4	770	13807.7376	-5815.6486	7
d. lump *C. f. cingulum* and *C. fuliginosus*	3	1239	10577.2167	-2585.1277	6
e. split off *C. f. flagellum*	5	737	5446.333	2545.756	2
f. split off *C. f. lineatulus*	5	734	5744.23	2247.859	3
g. split all subspecies	6	698	4957.0188	3035.0702	1
h. split all, mix all *C. flagellum*	6	810	7768.476	223.613	4
Analysis 2, *Coluber flagellum cingulum* & *C. fuliginosus* 20% missing data					
a. Current taxonomy	3	268	2648.984	-	6
b. lump *C. flagellum* with *C. fuliginosus*	2	321	3302.128	-653.144	8
c. split off *C. f. cingulum*	4	264	2298.313	350.671	4
d. lump *C. f. cingulum* and *C. fuliginosus*	3	312	2783.404	-134.42	7
e. split off *C. f. flagellum*	5	264	2081.4077	567.5763	2
f. split off *C. f. lineatulus*	5	264	2186.11	462.874	3
g. split all subspecies	6	264	2003.415	645.569	1
h. split all, mix all *C. flagellum*	6	268	2640.044	8.94	5
Analysis 3, *Coluber mentovarius striolatus* 50% missing data					
a. Current taxonomy, all split	4	366	1883.26	-	1
b. Lump *C. m striolatus* with *C. bilineatus*	3	912	4468.79	-2585.53	4
c. Lump *C. m. striolatus* with *C. m. mentovarius*	3	456	2649.06	-765.8	2
d. Lump *C. m. striolatus* with *C. flagellum*	3	548	3999.75	-2116.49	3
e. Split all lineages but mix randomly	4	1151	7994.76	-6111.5	5
Analysis 4, *Coluber mentovarius striolatus* 80% missing data					
a. Current taxonomy, all split	4	159	879.786	-	1
b. Lump *C. m striolatus* with *C. bilineatus*	3	216	1192.552	-312.766	4
c. Lump *C. m. striolatus* with *C. m. mentovarius*	3	159	1049.436	-169.65	2
d. Lump *C. m. striolatus* with *C. flagellum*	3	159	1090.767	-210.981	3
e. Split all lineages but mix randomly	4	159	1713.235	-833.449	5

Morphological data collection

We collected ventral and subcaudal counts from the literature and from three museum specimens. We recorded count data for the *C. flagellum* group (n = 1452) from Wilson (1970), for *C. m. striolatus* (n = 91), *C. m. variolosus* (n = 39), and *C. m. mentovarius* (n = 92) from

Zweifel (1960) and Johnson (1977), and for *C. bilineatus* (n = 4) from Hensley (1960). While the counts for most species pertained to large portions of the species' ranges, the counts for *C. bilineatus* were from one locality in Arizona, and only from four individuals. We summarized counts for each subspecies based on current subspecies distributions, except in *C. f. flagellum*, where we classified all *C. flagellum* west of the Mississippi River as *C. f. testaceus* based on our molecular results. We also counted ventral and subcaudal scales for three individuals that we sequenced, including one male and one female *C. m. striolatus*, and one male *C. bilineatus*. The *C. bilineatus* was from central Mexico, on the southern end of the range from the *C. bilineatus* measured by Hensley (1960).

RESULTS

Phylogenetic analyses revealed mito-nuclear discordance and suggested unrecognized species diversity

We used mitochondrial data to explore the number of lineages within whipsnakes (Fig. 3.1). Our mitochondrial analyses suggested that *C. flagellum* is composed of an eastern and western radiation, but that this species may be paraphyletic with respect to *C. bilineatus* and *C. m. striolatus*. We found that *C. f. testaceus*, *C. f. flagellum*, and *C. f. lineatulus* formed a monophyletic group with *C. bilineatus*, *C. m. striolatus*, and *C. m. mentovarius*. We refer to *C. f. flagellum, C. f. testaceus*, and *C. f. lineatulus* as the eastern *C. flagellum*. Within eastern *C. flagellum*, we found that clades did not strictly adhere to traditional subspecies range boundaries (Figs. 1A; A3.1). Specifically, *C. f. flagellum* was traditionally thought to extend west of the Mississippi River into the east Texas pine forests, but we found a clear distinction between *C.*

flagellum to the east and west of the Mississippi River. Likewise, we recovered two clades composed of samples pertaining to *C. f. lineatulus*. Both clades are restricted to the Chihuahua Desert. While the division between *C. f. lineatulus* and *C. f. testaceus* is clearly defined by the Pecos River Valley, where the Great Plains transition into the Chihuahua Desert we sampled several individuals west of the Pecos River Valley with the *C. f. testaceus* haplotype. We recovered strong support (bootstrap $\geq$ 70) for relationships between eastern *C. flagellum*, namely, a sister relationship between *C. f. testaceus* and *C. f. flagellum*, and the inclusion of *C. f. lineatulus* to form a monophyletic group (Fig. 3.1B). We recovered *C. m. mentovarius*, *C. bilineatus*, and *C. m. striolatus* as sister to the eastern *C. flagellum*. However, this sister relationship did not receive high support, nor did the sister relationship between *C. m. mentovarius* and the other two clades. We did recover strong support for the sister relationship between *C. m. striolatus* and *C. bilineatus*, and for the split between eastern and western *C. m. mentovarius*.

Finally, we recovered a group of clades that we refer to as western *C. flagellum* as sister to the eastern *C. flagellum*, *C. bilineatus*, *C. m. striolatus* and *C. m. mentovarius*. This western *C. flagellum* included three clades that represent *C. f. cingulum*, *C. f. piceus*, and *C. fuliginosus.* We recovered one monophyletic clade comprised of all individuals pertaining to *C. f. cingulum* including one individual from Michoacán. This *C. f. cingulum* clade also included individuals from Arizona, New Mexico, and California that were traditionally classified as *C. f. piceus* (Fig. 3.1). The *C. f. cingulum* clade was sister to Californian samples classified as *C. f. piceus*. These two clades were sister to *C. fuliginosus* from Baja California, MX.

Table 3.2. Mean between-group divergences generated from uncorrected *p* distances among Cytochrome b haplogroups using Mega 7.

	1	2	3	4	5	6	7	8	9

1. *C. f. flagellum*									
2. *C. f. testaceus*	3								
3. *C. f. lineatulus*	5.5	6							
4. *C. f. cingulum*	9.6	10.3	9.3						
5. *C. f. piceus*	9.6	9.8	9.8	4.6					
6. *C. fuliginosus*	9	9.9	9.1	6.6	6.5				
7. *C. m. striolatus*	9.3	9.8	9.2	12	11.3	10.8			
8. *C. m. mentovarius*	10.1	9.8	10.1	12.1	12.8	10.3	12		
9. *C. bilineatus* AZ	8.4	9.4	8.1	12	11.6	11.2	9.4	11.5	
10. *C. bilineatus* MX	9.7	8.9	8.4	11.4	11.8	11.5	10.3	10.6	6.4

Interspecific genetic divergences ranged from 3.0% between *C. f. flagellum* and *C. f. testaceus,* to 12.8% between *C. m. mentovarius* and *C. f. piceus* (Table 3.2). Three primary distinctions are clear from the distance matrix in Table 3.2. First, *C. f. flagellum*, *C. f. testaceus*, and *C. f. lineatulus* are closely related. Second, *C. f. cingulum, C. f. piceus*, and *C. fuliginosus* are more closely related than any other taxa in the matrix, and are less genetically distant from *C. bilineatus* than from the western *C. flagellum*. We note that *C. f. piceus* is only 4.6% divergent from *C. f. cingulum,* which is less divergent than *C. f. lineatulus* is from *C. f. flagellum* and *C. f. testaceus*. We found that *C. bilineatus* and *C. m. striolatus* are both distantly related to each other, as well as to all other whipsnakes. Finally, we found deep divergence of 6.4% between the northern and southern *C. bilineatus*, which is equivalent to the divergence between *C. f. lineatulus* and *C. f. testaceus*, or between *C. f. piceus* and *C. f. fuliginosus*.

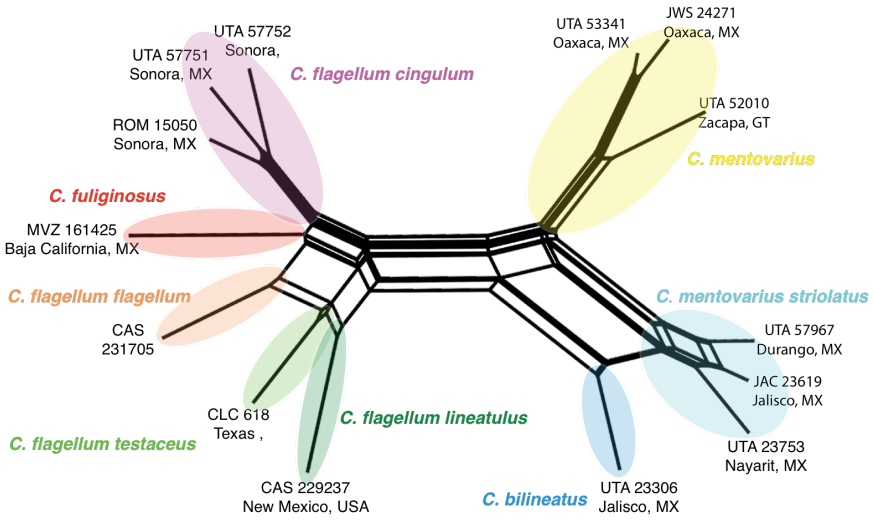

Fig. 3.2: Neighbor network generated using SPLITSTREE from 356 SNPs and 15 individuals representing each lineage for which we had nuclear sampling. The colors correspond to the range map in Fig. 1A.

Our neighbor network analysis using genomic SNP data (dataset A) recovered similar relationships within species as the mtDNA results, but very different relationships between species (Fig. 3.2). We recovered two genetic groups, one group including all individuals pertaining to *C. flagellum* and the other pertaining to *C. bilineatus* and *C. mentovarius*. Within *C. flagellum*, *C. fuliginosus* clustered closely to *C. f. cingulum* (*C. f. piceus* was not included). These two species were related to a cluster that included *C. f. flagellum*, *C. f. testaceus*, and *C. f. lineatulus*. *Coluber m. striolatus*, and *C. bilineatus* were more closely related to each other than to any other species, and clustered more closely to *C. m. mentovarius* than to *C. flagellum*. Thus, we found potential evidence for two cases of mito-nuclear discordance. First, the relationship between *C. f. flagellum*, *C. f. testaceus*, and *C. f. lineatulus* was different between the mitochondrial and nuclear datasets, where *C. f. testaceus* was sister to *C. f. flagellum* in the mitochondrial data, but to *C. f. lineatulus* in the nuclear data. Second, in the mitochondrial data *C. flagellum* was paraphyletic with respect to *C. bilineatus* and *C. mentovarius*, but in the nuclear data, *C. flagellum* appears to be monophyletic, with *C. bilineatus* and *C. m. striolatus* clustering more closely with *C. m. mentovarius*.

Species delimitation supports the elevation of several subspecies to species status

We tested eight species delimitation models to identify the number of lineages in *C. flagellum* (Table 3.1; Fig. 3.3). With both of our missing data thresholds, we recovered a general pattern where more split models were better supported than those that lumped discrete lineages. With dataset B our best-supported model involved splitting all possible lineages (BF = 3035.07), followed by splitting off *C. f. flagellum* (BF = 2545.76), and then by splitting off *C. lineatulus* (BF = 2247.86). Interestingly, the fourth best-supported model involved splitting all lineages, but mixing all *C. flagellum* subspecies randomly. With dataset C, our best-supported model again

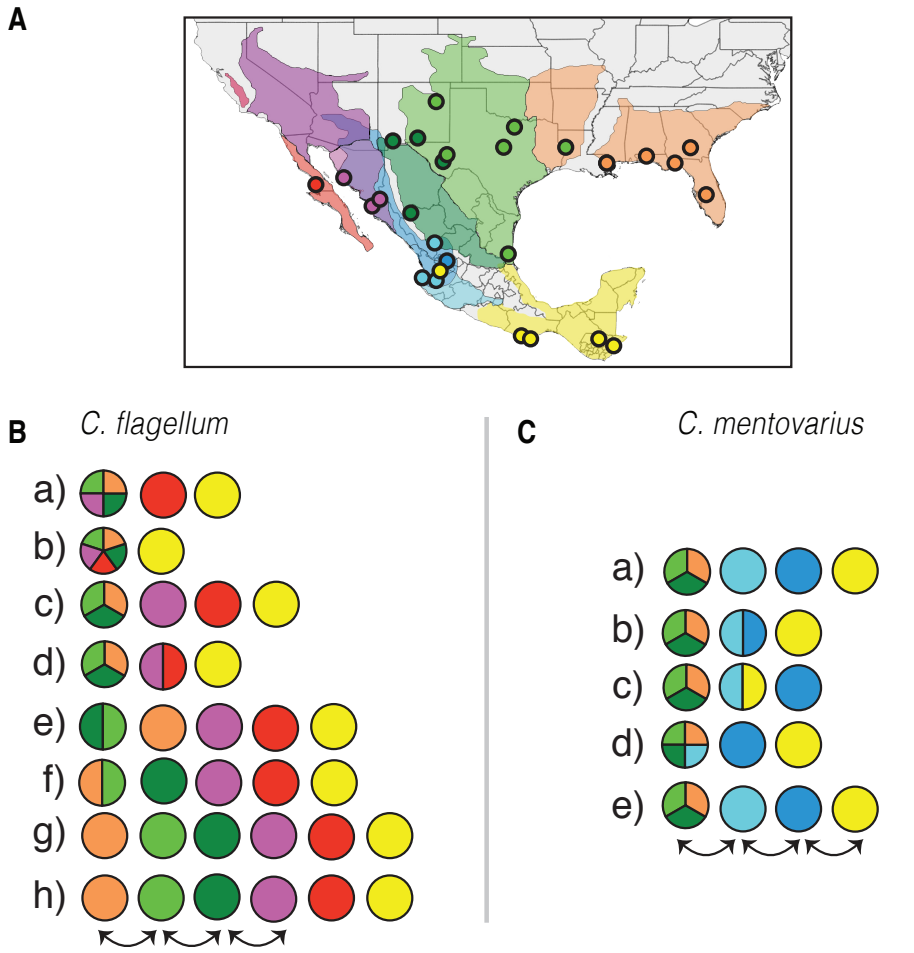

Fig. 3.3: Cartoon representing the different species relationship models we tested using Bayes Factor Delimitation. Pie charts represent models where lineages were lumped in the model. A) Map showing the distribution of whipsnake lineages. Circles represent samples with nuclear data used in the species delimitation analysis. B) Delimitation models for the *C. flagellum* species group. Orange = *C. f. flagellum*, light green = *C. f. testaceus*, dark green = *C. f. lineatulus*, purple = *C. f. cingulum*, red = *C. fuliginosus*, yellow = *C. m. mentovarius*. C) Delimitation models testing *C. m. striolatus*. The three *C. flagellum* lineages were lumped in each model. Light blue = *C. m. striolatus*, dark blue = *C. bilineatus*, yellow = *C. m. mentovarius*.

split all lineages (BF = 645.57), while the second and third best split off *C. f. flagellum* (BF = 576.58) and split off *C. f. lineatulus* (BF = 462.87). Between both analyses, the ranking of models that were lumped and mixed varied, but the top three models remained consistent. The dataset with higher levels of missing data, but more available loci (dataset B), recovered much lower ML estimates, and subsequently higher BF values than dataset C, which had less missing data.

We tested five species delimitation models for *C. m. striolatus* using datasets D and E, results for dataset E are shown in brackets (Table 3.1). The best supported model involved splitting all lineages. The other four models were ranked as follows: lumping *C. m. striolatus* with *C. m. mentovarius* (BF = -765.8 [-169.65]), lumping *C. m. striolatus* with *C. flagellum* (-2116.49 [-210.98]), lumping *C. m. striolatus* with *C. bilineatus* (BF = -2585.53 [-312.77]), and mixing all lineages randomly (BF = -611.5 [-813.45]). In summary, we found that BFD* supported the elevation of several lineages currently recognized as subspecies to species status. However, we note that method may have a bias towards splitting lineages rather than lumping them.

We estimated the species trees for the two best-supported model in each species group using datasets B-E. We found that nodes received higher support with datasets B and D, where $\leq$ 50% of loci where missing, but where more loci were available (Fig. 3.4; A4.2). Our analyses with datasets C and E had less missing data, but far fewer loci (Table 3.1). Our species trees of the *C. flagellum* group revealed a number of differences with the mitochondrial data. First, we found that the nuclear data supported the monophyly of *C. flagellum* relative to *C. mentovarius*. We found two sister clades within *C. flagellum*. In one clade we recovered a sister relationship between *C. f. cingulum* and *C. fuliginosus*. In the other clade we recovered a strongly supported

sister relationship between *C. f. testaceus* and *C. lineatulus*, which were sister to *C. f. flagellum*. All nodes had greater than 99% support. We recovered the same topology between datasets B and C, but dataset C only received 91% support for the sister relationship between *C. f. testaceus* and *C. f. lineatulus*, compared with over 99% support.

Our species tree analyses for *C. mentovarius* and *C. bilineatus* recovered the same topology for dataset D and E, but with large differences in support values at deeper nodes (Figs. 4; A4.2). In dataset D, we recovered a well resolved phylogeny with *C. m. striolatus* closely related to *C. bilineatus*, and these two lineages were sister to *C. m. mentovarius* with 99.27% support. In dataset E, we recover the same topology, but the sister relationship between *C. m. mentovarius* and *C. m. striolatus/C. bilineatus* was only supported by 84.50% support.

Morphological variation corresponds to discrete lineages

We collected ventral and subcaudal scale counts for each lineage investigated in this study; results are shown in Table 3.3. We find very little variation in the mean values of subcaudal counts between species, indicating that this character does not effectively differentiate whipsnake lineages. However, we do find substantial variation in ventral scale counts between lineages. We find that *C. bilineatus* has the highest mean number of ventral scales for males with a count of 203.8 (198–205.25; data lacking for females). *Coluber f. flagellum* also has a high ventral scale count, with a mean in males of 202.7 (201–203.7), and in females of 200.5 (196–203). This contrasts with the lowest number of ventral scales in *C. m. striolatus*, which has a mean count in males of 187 (176–195), and in females of 186.5 (166–202). It should also be noted that while *C. f. flagellum* has a high ventral count, *C. f. testaceus* has a much lower count, with a difference between the means of 10.3 scales in males and 8.2 scales in females. This corresponds to the

discontinuity observed at the Mississippi River between these two lineages. Our scale counts from museum specimens differed somewhat from the means gathered from the literature. Our male *C. bilineatus* specimen from Nayarit, Mexico, was lower than the *C. bilineatus* from Arizona reported by Hensley (1950), with a ventral count of 198. This agrees with the mitochondrial phylogenetic results that showed a deep split between the Arizonan and Mexican samples. Our male of *C. m. striolatus*, an adult specimen, had a ventral count of 193, and a subcaudal count of 121. Our female, a juvenile, had a ventral count of 188, and a subcaudal of 114. These measurements fell into the upper end of the ranges of previous *C. m. striolatus* counts (Table 3.3).

Table 3.3. Morphological data is summarized for each species. The mean for each taxon is shown for males and females for ventral and subcaudal scale counts as collected from the literature and our own specimen counts. Species are sorted by male ventral count in descending order. In parentheses is the range, followed by the sample size. We collected data for *Coluber flagellum* from Wilson (1970), for *C. bilineatus* from Hensley, for *C. m. striolatus* and *C. m. variolosus* from Zweifel (1960) and Johnson (1977), and for *C. m. mentovarius* from Johnson (1977).

	Ventral		Subcaudal	
	Male	Female	Male	Female
C. bilineatus	205.25 (203.00–204.00; 4)	–	–	–
C. f. flagellum	202.70 (201.00–203.70; 114)	200.50 (196.00–203.00; 117)	112.81 (108.00–116.00; 41)	109.26 (106.60–113.60; 50)
C. f. cingulum	195.30 (193.80–197.20; 174)	195.10 (185.00–205.00; 45)	108.10 (101.20–112.20; 91)	104.50 (99.80–106.50; 40)
C. m. variolosus	194.85 (190.00–204.00; 33)	194.50 (190.00–197.00; 6)	125.40 (119.00–132.00; 6)	115.70 (113.00–120.00; 6)
C. f. ruddocki	193.40 (193.40–193.40; 71)	194.00 (192.80–196.70; 100)	107.30 (107.30–107.30; 6)	108.00 (104.20–115.00; 50)
C. f. lineatulus	193.30 (191.10–197.00; 62)	193.30 (193.30–193.30; 7)	105.90 (104.30–108.10; 41)	102.00 (98.00–104.50; 31)
C. fuliginosus	193.30 (186.00–199.20; 82)	192.90 (187.30–198.00; 58)	117.86 (109.50–123.00; 32)	114.70 (108.80–119.20; 33)
C. f. testaceus	192.40 (188.00–196.00; 470)	192.85 (190.10–196.50; 73)	108.59 (105.50–115.10; 184)	103.30 (99.20–107.60; 181)
C. f. piceus	191.90 (189.10–195.30; 71)	192.60 (192.60–189.90; 54)	110.40 (105.30–115.30; 42)	112.10 (104.00–123.00; 45)
C. m. mentovarius	191.30 (181.00–203.00; 47)	192.30 (106.00–113.60; 384)	111.90 (102.00–120.00; 47)	–
C. m. striolatus	187 (176.00–195.00; 47)	186.50 (166.00–202.00; 46)	118.50 (111.00–123.00; 31)	113.60 (107.00–121.00; 29)

DISCUSSION

We use genetic and genomic data to explore lineage diversity within whipsnakes, and conduct species delimitation with genomic data to test species delimitation models for several lineages. We find that species diversity within whipsnakes is currently underdescribed. Namely, we found that *C. flagellum* is composed of an eastern and western radiation divided by the Cochise Filter Barrier. Within the eastern radiation, we found support for three lineages corresponding to *C. f. flagellum*, *C. f. testaceus*, and *C. f. lineatulus*. In the western group we found support for two lineages, corresponding to *C. f. cingulum*, and *C. fuliginosus*, although *C. f. piceus* was not sampled in the nuclear data. Within *C. mentovarius* we found that *C. m. striolatus* is most closely related to *C. bilineatus*, rather than *C. m. mentovarius*. Our species delimitation analyses supported the elevation of *C. f. cingulum* to evolutionary species, which we recommend elevating to species status in further work. We also found support for the elevation of *C. m. striolatus* to full species status, which we recommend elevating in future work. Our results underscore the utility of comprehensive sampling paired with coalescent species delimitation to identify and quantify evolutionary lineages.

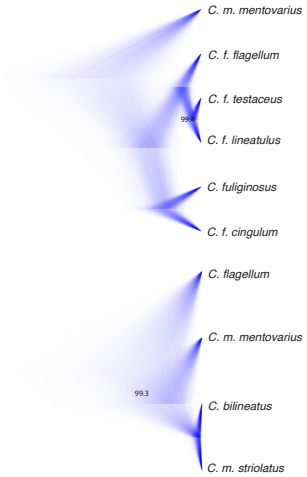

Fig. 3.4: Species trees generated using SNAPP based on the best-supported models from our Bayes Factor delimitation analysis shown in Fig. 3.3 for datasets B and D (≤ 50% missing loci). Support values are labeled for each node that is not fully supported.

Mito-nuclear discordance

Our two phylogenetic analyses demonstrated high levels of mito-nuclear discordance. Within *C. flagellum* we found support for an eastern and western radiation. These two radiations were not monophyletic in the mitochondrial analysis, rendered as such by *C. bilineatus*, *C. m. striolatus* and *C. m. mentovarius*. However, in both the neighbor network and species tree analyses, the *C. flagellum* lineages were monophyletic (with *C. fuliginosus* included). Mitochondrial data suggests that *C. f. cingulum* is on average 9.7% divergent from the eastern lineages of *C. flagellum*, indicating deep divergence between the eastern and western radiations. One other difference between the mitochondrial and nuclear results was the relationship of the three subspecies of eastern *C. flagellum*. In the mitochondrial data, we found support for a sister relationship between *C. f. flagellum* and *C. f. testaceus*, but nuclear data supported a sister relationship between *C. f. testaceus* and *C. f. lineatulus*. In all analyses, we recovered *C. m. striolatus* as most closely related to *C. bilineatus,* rather than to *C. m. mentovarius*. While the mitochondrial data placed this subspecies with *C. bilineatus* as sister to the eastern radiation of *C. flagellum*, the nuclear data placed these two species as closely related to *C. m. mentovarius*. These analyses demonstrate the utility of mitochondrial analyses for exploring lineage composition, but the power of genomic data to more fully resolve relationships between lineages. These results also highlight why caution is needed when interpreting phylogenies based on a single locus.

Missing data and species delimitation

Allelic dropout has been discussed extensively in the literature (Arnold et al., 2013). At deeper divergences, mutations in the digestion cut site lead to a reduction in homologous loci shared between species. This can lead to large amounts of missing data in more divergent taxa,

which can present a challenge when conducting analyses at the phylogenetic level (Rubin et al., 2012; Cariou et al., 2013; Huang & Knowles, 2014; Streicher et al., 2014; Collins & Hrbek, 2015; Leaché et al., 2015; Eaton et al., 2016). As a result, our analyses that included more divergent taxa (*C. bilineatus* and *C. m. striolatus*) resulted in fewer loci. Another challenge with the ddRADseq method is that it is less effective with lower quality DNA samples which may not digest well, or may fail size selection (Suchan et al., 2015). We hypothesize that DNA degradation due to the collection on roads of dead specimens reduced the number of available loci in several samples, especially with *C. bilineatus* and *C. m. striolatus*.

Conducting species delimitation using BFD* has several advantages, including the computational savings produced by the direct estimation of the species tree, the flexibility to vary the number of samples within each species, and the absence of a guide tree. However, this method does have several challenges. First, SNAPP does not accommodate missing loci between species. This creates a situation where the inclusion of more individuals per species can be advantageous, but in species with limited sampling, such as *C. bilineatus* in this study, only the SNPs present in that one individual can be included in the analysis. In analyses with few loci this may result in a locus bias to the exclusion of more variable (likely lineage specific) loci being excluded (Huang & Knowles, 2014). This is likely why in our species trees that include *C. bilineatus* and *C. m. striolatus* (Fig. 3.4; A3.2) we observe poor resolution of deeper nodes. Thus, we emphasize that ddRADseq presents challenges for phylogenetic analysis due to allelic dropout caused by deep divergences, as well as the effects of enzymatic digestion on poor quality samples, yet remains a useful method due to its low cost and rapid library preparation.

Much of the discussion regarding the effects of missing SNP data have focused on likelihood analyses of concatenated datasets (Eaton et al., 2016), but few studies have examined

the effects of missing data when using SNAPP, which does not accommodate missing data between species. We found that the inclusion of more loci, even at the expense of very high amounts of missing data (Tables 3.1; A3.2) led to higher BF and better resolved species trees than datasets with less missing data but fewer loci. This is likely because less stringent filtering is more likely to retain lineage-specific loci, which may help coalescent methods better delimit lineages (Huang & Knowles, 2014). Thus, we advocate that SNP based analyses should focus on maximizing total loci and lineage-specific (highly variable) loci, although the filtering regime will be different for each study.

Whipsnake taxonomy

We make several taxonomic recommendations for whipsnakes, but will leave these suggestions for future published work to avoid any duplication of names.

CONCLUSIONS

We demonstrate the power of using genetic and morphological data to explore lineage composition, and the use of genomic data to test models of species relationships to resolve recalcitrant taxonomic classifications, exemplified by *C. m. striolatus*. Our phylogenetic analyses recovered support for several lineages within *C. flagellum*, all of which pertain to recognized subspecies. We support the elevation of several whipsnake lineages, which will be described in a forthcoming paper. We found that *C. m. striolatus* was most closely related to *C. bilineatus.* We encourage further genomic sampling of western whipsnake lineages to further understand their phylogeny, and to investigate potential admixture at putative contact zones.

ACKNOWLEDGMENETS

We sincerely thank Ed Myers and Frank Burbrink for the use of their extensive sequencing data for western US coachwhips. We also TJ Firneno, D. Rivera, E. Wostl, DM Portik, and MK Fujita for their feedback on this manuscript. This work was supported by a Theodore Roosevelt Grant from the American Museum of Natural History.

REFERENCES

Arnold B, Corbett Detig RB, Hartl D, Bomblies K (2013) RADseq underestimates diversity and introduces genealogical biases due to nonrandom haplotype sampling. *Molecular Ecology* 22, 3179–3190. doi:10.1111/mec.12276.

Baird SF, Girard C (1853) *Catalogue of North American Reptiles in the Museum of the Smithsonian Institution. Part I. Serpents*. Washington: Smithsonian Institution.

Battstrom BH, Warren JW (1953) A new subspecies of racer, *Masticophis flagellum* from the San Joaquin Valley of California. Herpetologica 9: 177–179.

Bogert CM, Oliver JA (1945) A preliminary analysis of the herpetofauna of Sonora. *American Museum of Natural History*, 83, 297–426.

Bryant D, Bouckaert R, Felsenstein J, Rosenberg NA, RoyChoudhury A (2012) Inferring species trees directly from biallelic genetic markers: bypassing gene trees in a full coalescent analysis. *Molecular Biology and Evolution*, 29, 1917–1932. doi:10.1093/molbev/mss086

Burbrink FT, Myers EA (2015) Both traits and phylogenetic history influence community structure in snakes over steep environmental gradients. *Ecography*, 38, 1036–1048.

Burt CE (1935) Further records of the ecology and distribution of amphibians and reptiles in the middle west. *American Midland Naturalist*, 16, 311–336.

Cariou M, Duret L, Charlat S (2013) Is RAD-seq suitable for phylogenetic inference? An in silico assessment and optimization. *Ecology and Evolution*, 3, 846–852. doi:10.1002/ece3.512

Catchen J, Hohenlohe PA, Bassham S, Amores A, Cresko WA (2013) Stacks: an analysis tool set for population genomics. *Molecular Ecology*, 22, 3124–3140. doi:10.1111/mec.12354

Collins RA, Hrbek T (2015) An in silico comparison of reduced-representation and sequence-capture protocols for phylogenomics. bioRxiv 032565.

Conant R, Collins JT (1998) Reptiles and Amphibians: Eastern/Central North America. Houghton Mifflin Co, Boston.

Cope ED (1892) A critical review of the characters and variation of the snakes of North America. *Proceedings of the U.S. Natural History Museum*, 14, 589–694.

Crother BI (2012) Standard Common and Current Scientific Names for North American Amphibians, Turtles, Reptiles, and Crocodilians, Seventh Edition. SSAR Herpetological Circulation, 39, 1–92

De Queiroz K (2007) Species concepts and species delimitation. *Sytematic Biology,* 56, 879–886.
Eaton DA, Spriggs EL, Park B, Donoghue MJ (2016) Misconceptions on Missing Data in RAD-seq Phylogenetics with a Deep-scale Example from Flowering Plants. *Systematic Biology,* syw092. doi: 10.1093/sysbio/syw092
Faircloth BC, McCormack JE, Crawford NG, Harvey MG, Brumfield RT, Glenn TC (2012) Ultraconserved elements anchor thousands of genetic markers spanning multiple evolutionary timescales. *Systematic Biology*, 61, 717–726. doi:10.1093/sysbio/sys004
Fontaneto D, Flot J-F, Tang CQ (2015) Guidelines for DNA taxonomy, with a focus on the meiofauna. *Marine Biodiversity,* 45, 433–451.
Fujita MK, Leaché AD, Burbrink FT, McGuire JA, Moritz C (2012) Coalescent-based species delimitation in an integrative taxonomy. *Trends Ecology and Evolution*, 27, 480–488.
Grismer LL (1994) The origin and evolution of the peninsular herpetofauna of Baja California, Mexico. *Herpetological Natural History*, 2, 51–106.
Grimer LL (1999) Checklist of amphibians and reptiles on islands in the Gulf of California. *Bulletin of the Southern California Academy of Sciences,* 98, 45–56.
Huang H, Knowles LL (2014) Unforeseen consequences of excluding missing data from next-generation sequences: simulation study of RAD sequences. *Systematic Biology*, 65, 357–365.
Huson DH, Bryant D (2006) Application of phylogenetic networks in evolutionary studies. *Molecular Biology and Evolution*, 23, 254–267. doi:10.1093/molbev/msj030
Hykin SM, Bi K, McGuire JA (2015) Fixing Formalin: A Method to Recover Genomic-Scale DNA Sequence Data from Formalin-Fixed Museum Specimens Using High-Throughput Sequencing. PLOS ONE 10, e0141579–16. doi:10.1371/journal.pone.0141579
Johnson JD (1977) The Taxonomy and Distribution of the Neotropical Whipsnake Masticophis mentovarius (Reptilia, Serpentes, Colubridae). *Journal of Herpetology,* 11, 287–309.
Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A, Markowitz S, Duran C, Thierer T, Ashton B, Meintjes P, Drummond A (2012) Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics 28, 1647–1649. doi:10.1093/bioinformatics/bts199.
Klauber LM (1942) The status of the black whipsnake. Copeia 1942, 88–97.
Knowles LL, Carstens BC (2007) Delimiting species without monophyletic gene trees. *Systematic Biology,* 56, 887–895.
Kühnert D, Vaughan T, Wu C-H (2014) BEAST 2: A Software Platform for Bayesian Evolutionary Analysis. *Plos Computational Biology,* 10, e1003537–6. doi:10.1371/journal.pcbi.1003537
Kumar S, Stecher G, Tamura K (2016) MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets. *Molecular Biology and Evolution*, 33, 1870–1874.
Leaché AD, Banbury BL, Felsenstein J, de Oca AN-M, Stamatakis A (2015) Short tree, long tree, right tree, wrong tree: new acquisition bias corrections for inferring SNP phylogenies. *Systematic Biology,* 64, 1032–1047
Leaché AD, Fujita MK, Bouckaert RR, Minin VN (2014) Species delimitation using genome-wide SNP data. *Systematic Biology*, 63, 534–542.
Liu L, Xi Z, Wu S, Davis CC, Edwards SV (2015) Estimating phylogenetic trees from genome-scale data. *Annals of the New York Academy of Sciences*, 1360, 36–53.
Lowe CH, Norris KS (1955) Analysis of the herpetofauna of Baja California, Mexico. III. New and revived reptilian subspecies of Isla de San Esteban, Gulf of California, Sonora, Mexico,

with notes on other satellite island of Isla Tiburon. *Herpetologica*, 11, 89–96.
Lowe CH, Woodin WH (1954) A new Racer (Genus *Masticophis*) from Arizona and Sonora, Mexico. *Proceedings of the Biological Society of Washington,* 67, 247–250.
Mertens R (1934) Die Insel-Reptilian, ihre Ausbreitenung, Variation und Artbildung, *Zooligcal, Stuttgart*, 84, 1–209.
Nagy ZT, Lawson R, Joger U, Wink M (2004) Molecular systematics of racers, whipsnakes, and relatives (Reptilia: Colubridae) using mitochondrial and nuclear markers. *Journal of the Zoology, Systematics, and Evolutionary Resources,* 42, 223–233.
O'Connell KA, Streicher JS, Smith EN, Fujita MK (2017) Geographical features are the predominant driver of molecular diversification in widely distributed North American Whipsnakes. *Molecular Ecology,* 26: 5729–5751.
Ortenburger AI (1923) A note on the genera *Coluber* and *Masticophis* and a description of a new species of Masticophis. *Occasional papers of the Museum of Zoology, University of Michigan*, 139, 1–14.
Padial JM, Miralles A, De la Riva I, Vences M (2010) The integrative future of taxonomy. *Frontiers in Zoology,* 7, 16. doi:10.1186/1742-9994-7-16
Pante E, Abdelkrim J, Viricel A, Gey D, France SC, Boisselier M-C, Samadi S (2015) Use of RAD sequencing for delimiting species. *Heredity,* 114, 450–459.
Peters JA, Orejas-Miranda BR (1970) Catalogue of neotropical Squamata: Part 1, snakes, *Bulletin of the US Natural History Museum,* 297, 347.
Petit RJ, Excoffier L (2009) Gene flow and species delimitation. *Trends in Ecology and Evolution,* 24, 386–393.
Rambaut A (2017) http://tree.bio.ed.ac.uk/software/figtree/. Accessed 12/15/16.
Reid NM, Satler JD (2013) How to fail at species delimitation. *Molecular Ecology*, 22, 4369–4383.
Roze JA (1953) The *Rassenkreis Coluber* (*Masticophis*) *mentovarius* (Dumeril, Bibron et Dumeril), 1854. *Herpetologica*, 9, 113–120.
Ruane S, Bryson R, Pyron A, Burbrink F (2013) Coalescent species delimitation in milksnakes (genus Lampropeltis) and impacts on phylogenetic comparative analyses. *Systematic Biology,* 63, 231–250. doi:10.1093/sysbio/syt099
Rubin BE, Ree RH, Moreau CS (2012) Inferring phylogenies from RAD sequence data. *PlOS ONE*, 7, e33394.
Shaw G (1802) General Zoology, or Systematic Natural History. Vol.3, part 2. G. Kearsley, Thomas Davison, London, 313–615
Sites JW, Marshal JC (2004) Operational criteria for delimiting species. *Annual Review Ecology Evolution and Systematics,* 35, 199–227.
Sites JW, Marshall JC (2003) Delimiting species: a Renaissance issue in systematic biology. *Trends in Ecology and Evolution,* 18, 462–470.
Smith HM (1941) Notes on Mexican snakes of the genus *Masticophis*. *Journal of the Washington Academy of Sciences,* 31, 388–398.
Smith HM (1943) Summary of the collections of snakes and crocodilians made in Mexico under the Walter Rathbone Bacon Traveling Scholarship. Ibid 93, 393–504.
Smith HM, Taylor EH (1945) An annotated checklist and key to the snakes of Mexico. *Bulletin of US Natural History Museum,* 187, 1–239.
Stebbins RC (2003) Western Reptiles and Amphibians. Houghton Mifflin Co, New York.
Streicher JW, Devitt TJ, Goldberg CS, Malone JH, Blackmon H, Fujita MF (2014)

Diversification and asymmetrical gene flow across time and space: lineage sorting and hybridization in polytypic barking frogs. *Molecular Ecology,* 23, 3273–3291. doi:10.1111/mec.12814

Streicher JW, McEntee JP, Drzich LC, Card DC, Schield DR, Smart U, Parkinson CL, Jezkova T, Smith EN, Castoe TA (2016) Genetic surfing, not allopatric divergence, explains spatial sorting of mitochondrial haplotypes in venomous coralsnakes. *Evolution*, 70, 1435–1449.

Suchan T, Pitteloud C, Gerasimova N, Kostikova A, Arrigo N, Pajkovic M, Ronikier M (2015) Hybridization capture using RAD probes (hyRAD), a new tool for performing genomic analyses on museum collection specimens. *PLOS ONE*, doi:10.1101/025551

Taylor EH (1938) Notes on the herpetological fauna of the Mexican state of Sonora. *University of Kansas Science Bulletin,* 24, 475–503.

Uetz P, Hošek J (2017) The Reptile Database, http://www.reptile-database.org, accessed Jan 15, 2017.

Utiger U, Schatti B, Helfenberger N (2005) The oriental colubrine genus Coelognathus FITZINGER, 1843 and classification of old and new world racers and ratsnakes (Reptilia, Squamata, Colubridae, Colubrinae). *Russian Journal of Herpetology*, 12, 39–60.

Wiens JJ (2007) Species delimitation: new approaches for discovering diversity. *Systematic Biology,* 56, 875–878.

Wilson LD (1970) The coachwhip snake, *Masticophis flagellum* (Shaw): taxonomy and distribution. *Tulane Studies of Zoological Bulletin,* 16, 31–99.

Zweifel RG (1960) Herpetology of the Tres Maria Islands. Results of the Puritan-American Museum of Natural History expedition to Western Mexico. *Bulletin of the American Museum of Natural History*, 119, 77–128.

Zweifel RG, Norris KS (1955) Contribution to the herpetology of Sonora, Mexico: descriptions of new subspecies of snakes (*Micruroides euryxanthus* and *Lampropeltis getulus*) and miscellaneous collecting notes. *American Midland Naturalist,* 54, 230–248. WITHIN-ISLAND DIVERSIFICATION UNDERLIES PARACHUTING FROG (*RHACOPHORUS*) SPECIES ACCUMULATION ON THE SUNDA SHELF

Kyle A. O'Connell[1,2], Utpal Smart[1,2], Eric N. Smith[1,2], Amir Hamidy[3], Nia Kurniawan[4], Matthew K. Fujita[1,2]

[1]*Department of Biology, The University of Texas at Arlington, Arlington, Texas 76019, USA.*

[2]*The Amphibian and Reptile Diversity Research Center; University of Texas at Arlington, Arlington, Texas 76010, USA.*

[3]*Research and Development Center for Biology, Indonesian Institute of Science (LIPI), Widyasatwaloka Building, Cibinong 16911, West Java, Indonesia.*

[4]*Department of Biology, Universitas Brawijaya, Jl. Veteran, Malang 65145, East Java, Indonesia.*

Citation: O'Connell, KA, Smart, U, Smith, EN, Hamidy, A, Kurniawan, N, Fujita, MF. *In Press*. Within-island diversification underlies parachuting frog (*Rhacophorus*) species accumulation on the Sunda shelf. *Journal of Biogeography.* DOI:10.1111/jbi.13162

CHAPTER 4

WITHIN-ISLAND DIVERSIFICATION UNDERLIES PARACHUTING FROG (*RHACOPHORUS*) SPECIES ACCUMULATION ON THE SUNDA SHELF

ABSTRACT

This study seeks to understand the geological and climatological processes that have promoted biodiversity on the Sunda Shelf in Southeast Asia. Using the parachuting frog genus *Rhacophorus*, we estimate divergence times and quantify the respective contributions of between and within-island diversification to species richness and endemism. We generated a concatenated mitochondrial and nuclear DNA sequence alignment for 40 species of *Rhacophorus*. We estimated phylogenetic relationships and divergence times, constructed lineage-through-time plots, and reconstructed ancestral ranges. We found that *Rhacophorus* originated 33.0 Ma, and diversified at a slower-than-constant rate through time. Dispersal was important to early *Rhacophorus* evolution, but subsequent *in situ* diversification produced most species diversity on Sumatra and Borneo. Clades that diversified via *in situ* processes contained higher proportions of endemic species. Species diversification on the Sunda Shelf is ancient and has occurred slowly. Both dispersal and *in situ* diversification have promoted Sundaland species accumulation, but within-island phylogenesis has produced a greater proportion of endemic species on Sumatra and Borneo.

INTRODUCTION

Classic island biogeography theory has focused on how the immigration-speciation-extinction dynamic has influenced species richness and endemism on islands (Weigelt et al.,

2016; Patino et al., 2017). The size, age, and level of isolation of islands have all been implicated as important determinates of species richness and endemism (MacArthur & Wilson, 1967; Heaney, 1986; Rabosky & Glor, 2010). More connected islands should have higher immigration (dispersal) and higher species richness, but lower endemism and phylogenesis (MacArthur & Wilson, 1967). In contrast, large islands exhibit higher endemism and *in situ* diversification driven by greater habitat complexity (Losos & Schluter, 2000; Kisel & Barraclough, 2010; Borregaard et al., 2016).

Diversification rates on islands should reflect this dispersal-speciation-extinction dynamic; with high diversification rates following colonization, but low rates once species density dependence is reached (Esselstyn et al., 2009). Thus, a trajectory of declining rates is expected on older islands where niche space has already been filled in the ancient past. In addition, historical events, such as sea level fluctuations, can propel or inhibit diversification, leaving an imprint on the rates of diversification through time (Roberts et al., 2011; Klaus et al., 2012).

Diversification on islands occurs by two processes: dispersal and divergence in allopatry or within-island (*in situ*) diversification (MacArthur & Wilson, 1963; Losos & Ricklefs, 2009). While many past studies have focused on the role of dispersal-driven diversification, within-island diversification has also been implicated as a key component of island species accumulation (Gillespie, 2004; Cornuault et al., 2013; Warren et al., 2015; Wittaker et al., 2017). Over geological timescales, clades that have diversified via within-island processes are expected to contain higher levels of richness and endemism than clades that diversified via dispersal and allopatry (Heaney, 2000).

The Sunda Shelf in the Malay Archipelago (Sundaland) is a shallow continental shelf encompassing the islands of Sumatra, Java, Borneo, and the Malay Peninsula (MP; Fig. 4.1). Sundaland is ideal for studying island diversification processes because it has experienced dynamic tectonic processes, volcanism, dramatic sea level changes, and extensive connectivity between landmasses during the Tertiary and Quaternary (Hall, 2009; 2012a; 2012b; Morley, 2012). Past studies have found that three historical processes have most influenced species diversification in Sundaland: sea level changes and volcanic uplift during the Miocene-Pliocene, and sea level fluctuations during the Pleistocene (Inger & Voris, 2001; Meijaard, 2004; Lohman et al., 2011; Roberts et al., 2011). From the late Oligocene until the middle Miocene, sea levels gradually rose, peaking at ~15 Ma (Haq et al., 1987; Meijaard, 2004; Barber, 2005). High sea levels isolated Sumatran volcanoes into multiple islands, and reduced the land area of Borneo and the MP (Lohman et al., 2011; Hall 2012a). Sea levels subsided from the late Miocene to the late Pliocene, although additional sea level high stands may have persisted, especially ~5 Ma (Haq et al., 1987; Meijaard, 2004). In the late Miocene and early Pliocene mountain uplift driven by accelerated volcanic and tectonic activity expanded the land positive areas of Sumatra and Borneo (Barber, 2005; Hall 2012a). Conversely, Java emerged as a series of volcanic islands beginning ~10 Ma, but only assumed its current form ~5 Ma. During late Pliocene and Pleistocene interglacial periods all four landmasses experienced periodic connectivity (Voris, 2000).

Thus, the MP and Borneo demonstrated ancient stability and connectivity compared with the isolation of smaller islands on present day Sumatra and Java (Hall, 2012b; de Bruyn et al., 2014). Despite this isolation, two land bridges may have connected several landmasses in the early and late Miocene and again in the early Pliocene (van Bemmelen, 1943; Meijaard, 2004;

Barber, 2005). These include the Asahan High, which connected northern Sumatra to the MP, and the Lampung High, which connected southern Sumatra to west Java, Borneo, and the MP (van Bemmelen, 1949; Barber, 2005). These highs may have permitted dispersal between landmasses despite isolation by marine incursions.

To investigate island diversification dynamics, we conducted phylogenetic analyses of parachuting frogs from the genus *Rhacophorus*. *Rhacophorus* includes ~90 species distributed from the Indian subcontinent to eastern China, south to the Sunda Shelf, and east to Sulawesi. The Sunda Shelf contains 26 species of *Rhacophorus*, with four species on the MP, 15 on Borneo, 15 on Sumatra, and two on Java (Frost, 2016). Sunda Shelf species exhibit high endemism to single islands (73%), yet their evolutionary relationships and taxonomy remain poorly understood, particularly regarding the placement of presumed endemic species on Sumatra and Borneo (Streicher et al., 2012; Hetwig et al., 2013). Using representative species from all Sundaland *Rhacophorus* clades, we generated mitochondrial (mtDNA) and nuclear (nuDNA) sequence data to answer the following questions: 1) When did diversification occur on the Sunda Shelf, and what do divergence dates suggest about how geological and climatological processes have influenced diversification? 2) What has been the rate and tempo of diversification

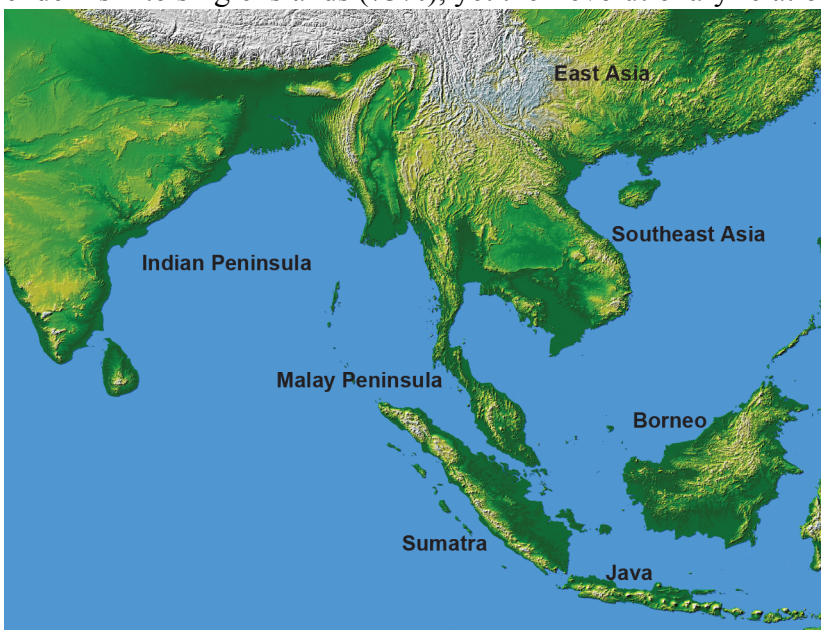

Fig. 4.1. Map of the Sunda Shelf demonstrating biogeographic regions highlighted in this study.

in *Rhacophorus* and how does this compare with other Sundaland taxa? 3) Is *Rhacophorus* species richness and endemism on the Sunda Shelf primarily driven by between-island dispersal or within-island diversification?

MATERIALS AND METHODS

Taxon sampling and molecular sequence data

We collected samples during field surveys in 1996, and between 2013-2016 in primarily highland habitats across Sumatra (Harvey et al., 2015). We extracted DNA from liver and thigh muscle tissue samples from eight *Rhacophorus* species from Sumatra and Java including *R. achantharrhena, R. catamitus*, *R. modestus*, *R. poecilonotus*, *R. bengkuluensis, R. margaritifer*, *R. reinwardtii* and *R. prominanus*, stored in SDS buffer. Extractions were done using a standard salt extraction protocol (Sambrook and Russell, 2001). We checked the quality of our extractions using a 1% Agarose gel and quantified the DNA using QUBIT® 2.0 Fluorometer (Life Technologies, Grand Island, NY, USA). We sequenced a 609 base pair fragment of the 16S ribosomal RNA gene using 16S AR, BR primers (Palumbi *et al.,* 1991). We also sequenced the brain derived neurotrophic factor gene (BDNF; Van der Meijden *et al.,* 2007) for the same eight species. Each PCR reaction occurred in a 25µl reaction. The amplification protocol for all PCR reactions was: initial denaturation at 94˚C for 2 min, 40 cycles of 94˚C for 30 sec, 50˚C for 30 sec, and 72˚C for 30 sec, with a final extension at 72˚C for 10 min. PCR purifications were performed using Sera-Mag Speedbeads (Rohland & Reich, 2012). Cycle sequencing reactions were conducted using PCR primers under the following conditions: initial denaturation at 95˚C for 2 min, 40 cycles of 95˚C for 15 sec, 50˚C for 15 sec, final extension at 60˚C for 4 sec.

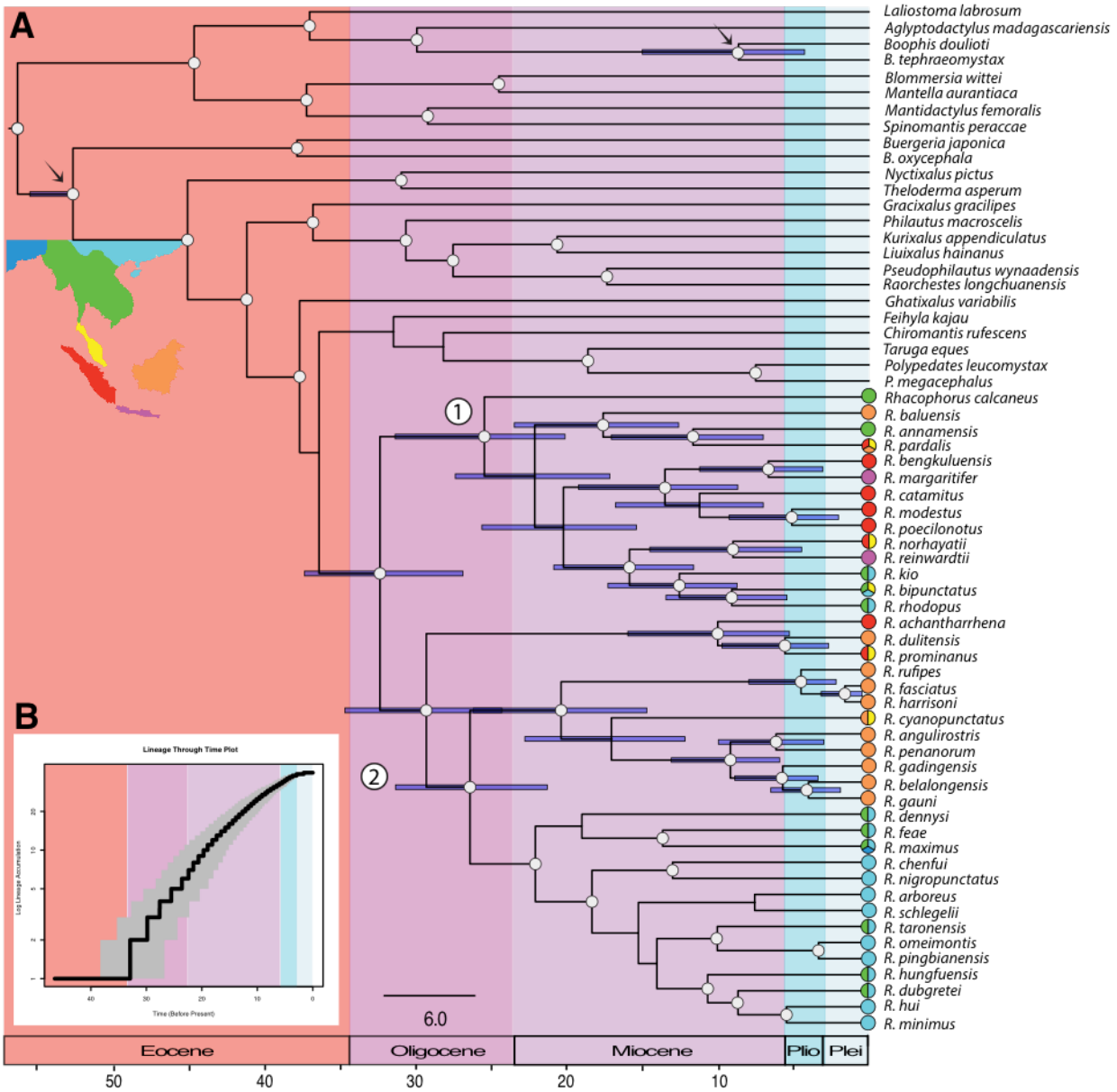

Fig. 4.2. A) Dated Bayesian phylogeny generated using BEAST2. Nodes with ≥ 95% posterior probability are highlighted with gray circles. Blue node bars span the 95% confidence interval for node ages. The two primary clades recovered for Rhacophorus are marked with the large numbers. Black arrows show the location of calibration points. On the right, subclades mentioned in the text are labeled 1–12. B) Lineage through time plot shown in log scale for Rhacophorus showing relatively constant diversification through time.

Sequencing products were resolved on an Applied Biosystems 3130XL at the University of Texas Arlington Genomics Core Facility (gcf.uta.edu; Arlington, TX, USA). Sequences were assembled and edited in Geneious v.7.0 (Kearse et al., 2012).

To expand our taxonomic sampling, we included sequences from Genbank for 33 other species of *Rhacophorus*. We also included at least one species of each genus within the family

Rhacophoridae (n=17), and eight outgroup species of the family Mantellidae based on the relationships recovered by Li et al., (2013). We excluded *Rhacophorus* endemic to the Indian subcontinent as these were not relevant to our research questions. Our expanded dataset included sequences for 12S rRNA (12S; n=25), 16S rRNA (16S; n=67), Cytochrome oxidase c subunit I (COI; n=26), Cytochrome b (CYTB; n=37), brain derived neurotrophic factor gene (BDNF; n=42), pro-opiomelanocortin (POMC; n=29), recombination-activating gene 1 (RAG1; n=37), Rhodopsin (RHOD; n=38) and Tyrosinase (TYR; n=24). All information regarding Genbank IDs can be found in Table A4.1. We aligned non-coding loci individually using the Geneious aligner using default parameters, aligned coding genes using Muscle (Edgar, 2004) using default parameters, and concatenated all loci.

Phylogenetic analysis and divergence dating

We began phylogenetic analyses by selecting the most probable model of nucleotide evolution for Bayesian inference (BI) using Bayesian information criteria implemented in PartitionFinder (Lanfear et al., 2012), partitioning by locus. We estimated the phylogeny and divergence times in BEAST v.2.4.5 (Bouckaert et al., 2014). We defined three gene partitions: 12S and 16S, CYTB and COI, and all five nuclear genes based on PartitionFinder results. PartitionFinder selected the GTR+I+Γ model for all partitions, but due to a lack of run convergence (ESS values < 200), we assigned the HKY model to each partition following Drummond & Bouckaert, (2015). Following Li et al., (2013), we calibrated the origin of Rhacophoridae to 53.2 Ma using the fossil *Indorana prasadi*. We applied a relaxed Log Normal clock model and the Yule tree prior. We assigned a Log Normal calibration to the most recent common ancestor (MRCA) of all Rhacophoridae, with a mean of 1.0, a standard deviation of

1.25, and an offset of 52.3 (the age of the fossil). This produced a Rhacophoridae MRCA distribution 95% confidence interval (CI) of 52.3–57.6 Ma. We constrained to monophyly all members of Rhacophoridae. We assigned a uniform prior distribution on the MRCA of *Boophis doulioti* and *Boophis tephraeomystax*, with a range of 0.0 to 15 Ma based on the oldest estimated age for the Comoro island of Mayotte where *B. tephraeomystax* is endemic

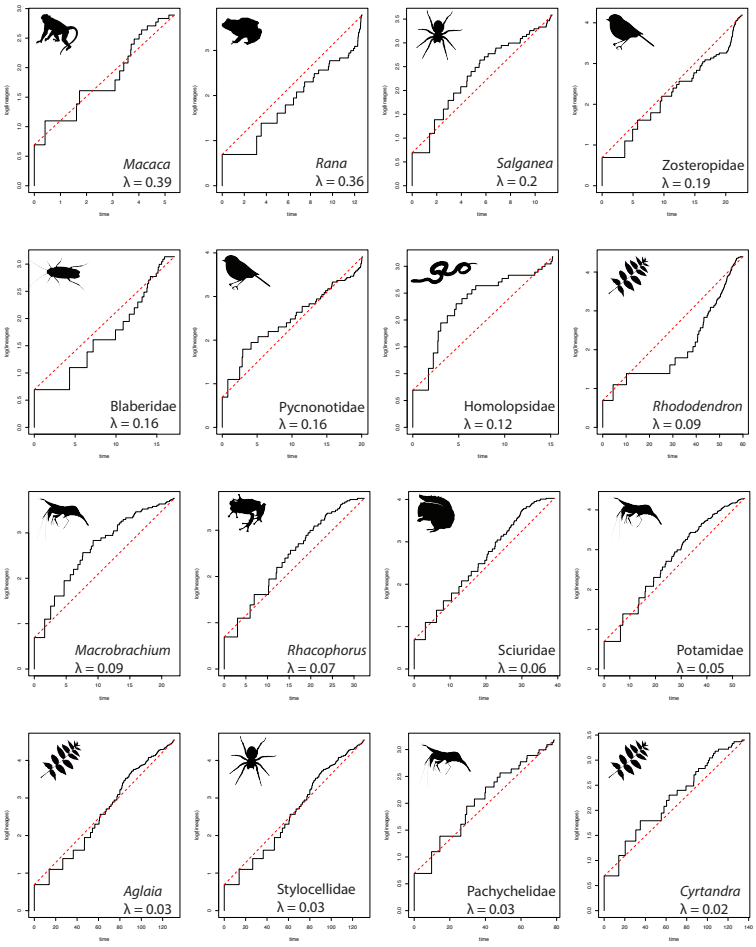

Fig. 4.4. Lineage-through-time plots sorted by speciation rate.

(Vences et al., 2003). We changed the default diffused Uniform prior on the ucldMean to an Exponential distribution and set the mean to 10.0. We left all other priors at default values. We ran our analyses for 1,000,000,000 MCMC generations, sampling every 10,000 generations. We confirmed convergence of runs (ESS values > 200) using Tracerv1.6 (Rambaut et al., 2014). We removed the first 25% of trees as burnin (2,500 trees) using TreeAnnotator (Bouckaert et al., 2014), and combined the remaining 7,500 trees to produce the maximum clade credibility tree with median node heights.

Lineage-through-time plots

We constructed lineage-through-time plots (LTT) to understand two questions regarding the rate and tempo of Sundaland diversification. First, at what stage of the classical island diversification model are Sundaland taxa? For example, accelerating rates would suggest an excess of ecological opportunity, while decelerating rates would indicate species equilibrium. Second, did any specific geological processes leave a signal on diversification rates? For example, do we observe rate shifts across taxa correlated with recent glacial cycles? We began by constructing a LTT for the time-calibrated phylogeny of *Rhacophrous* using the R package Phytools, and visualized results in RStudio 0.99 (Revel, 2012; Racine 2012; R Core Team 2013). We sampled the last 1000 trees from our BEAST analysis and estimated the LTT of the mean and 95% CI using the R package paleotree (Bapst, 2012). We used Phytools to calculate the speciation rate (λ) under a Yule model, and calculated the γ statistic of Pybus and Harvey (2000) as implemented in the R package LASER (Rabosky, 2006). LASER attempts to correct for incomplete lineage sampling, which can bias the γ statistic towards a negative diversification rate (Pybus & Harvey, 2000). To investigate diversification through time more broadly, we constructed an additional 15 plots using time-calibrated phylogenies from de Bruyn et al., (2014) encompassing diverse taxa of varying ages. We only included phylogenies with taxa from the Sunda Shelf, and those with sampling of approximately

Fig. 4.4. Histograms of binned values of lambda and gamma for the 16 lineage-through-time plots shown in Fig. 3. The red line shows where *Rhacophorus* measures for both values.

one individual per species. Our dataset included terrestrial and freshwater invertebrates, plants, amphibians, reptiles, birds, and mammals (Table 4.1). Root ages ranged from 5.4–135.7 Ma (Table 4.1). We calculated λ and the γ statistic for the MCC tree from each dataset. To evaluate how *Rhacophorus* compared with the other 14 taxa, we grouped these values into bins and plotted the number of species within each bin. Bin sizes were 0.052 for λ and 1.062 for the γ statistic.

Table 4.1: Results of lineage-through-time analyses.

Genus	Common Name	Root Age	nTaxa (tips)	nSpecies	Total Species in taxa	λ	γ	Critical value	*p value*
Macaca	Mammals	5.4	17	15	20	0.39	-1.14	-1.91	0.29
Homolopsidae	Reptile	15.2	24	20	34	0.12	-1.45	-2.22	0.25
Zosteropidae	Birds	18.9	57	42	80	0.19	2.08	-2.75	1
Pycnonotidae	Birds	20.1	64	43	130	0.155	1.69	-2.73	1
Rhododendron	Plants	60	46	46	300	0.09	-0.92	-2.3	0.43
Pachychelidae	Freshwater molluscs	78.7	21	21	140	0.03	-0.8	-2.1	0.46
Aglaia	Plants	108.3	42	42	390	0.03	-1.32	-1.91	0.25
Stylocellidae	Insects and Spiders	131	95	95	300	0.03	-2	-3.15	0.35
Cyrtandra	Plants	135.7	30	26	300	0.02	-1.86	-2.26	0.1
Salganea	Insects and Spiders	12.6	36	22	50	0.2	-0.68	-2.33	0.58
Rana	Amphibians	12.7	15	14	14	0.36	4		
Blaberidae	Insects and Spiders	17.2	22	21	22	0.16	-0.53		
Macrobrachium	Freshwater crustaceans	22	43	43	105	0.09	-3.02	-2.63	0.02
Rhacophorus	Amphibians	32.4	40	40	90	0.07	-3.37	-2.43	0.02
Sciuridae	Mammals	39	15	15	15	0.06	-4.3		
Potamidae	Freshwater crustaceans	54.5	65	65	650	0.05	-3.33	-2.4	0.01

Ancestral range evolution

We reconstructed ancestral range evolution for *Rhacophorus* species in a ML framework using the R package BioGeoBEARS (Matzke, 2013). BioGeoBEARS adopts a maximum likelihood framework to employ several popular models of range evolution (i.e. DIVALIKE, LAGRANGE or DEC and BAYAREALIKE), and allows the addition of "founder-event

speciation," a cladogenetic process termed "J" (i.e. DIVALIKE + J, DEC + J, and BAYAREALIKE + J). The four models (plus the addition of "J") vary in parameters describing range expansion, vicariance, and speciation (Matzke, 2013). A model that includes founder event speciation would suggest that long distance dispersal, rather than simple vicariance best describes the evolutionary history of a taxon. We utilized the Akaike Information Criterion (AIC) to compare the fit of different models to the data. We pruned all outgroup taxa from the time-calibrated phylogeny obtained from the BEAST analysis to retain only *Rhacophorus* species. We defined seven geographic areas based on the distribution of extant species following de Bruyn et al., (2014): the Indian peninsula (to the exclusion of the Northeast), East Asia (China, Taiwan, and Japan), Southeast Asia (SE Asia; Northeast India, and the geographical area between southern China and north of the Isthmus of Kra), the MP (Thai peninsula south of the Isthmus of Kra and Malaysia), Borneo, Sumatra, and Java (Fig. 4.5A). We performed an unconstrained analysis and limited the number of areas at each node to a maximum of four as no taxa were distributed across more than four geographic areas.

RESULTS

Phylogenetic analysis and divergence dating

We used BI to estimate phylogeny and divergence times between *Rhacophorus* species (Fig. 4.2A). We recovered a root age of 56.40 Ma (95% HPD 52.38–63.70 Ma) for the MRCA of Mantellidae and Rhacophoridae, of 52.73 Ma (52.31–55.45 Ma) for the root of Rhacophoridae, and of 32.98 Ma (26.95–37.58 Ma) for the root of *Rhacophorus* (Fig. 4.2A; A4.1). We recovered two primary clades (clades 1 and 2) within *Rhacophorus* with high support (posterior probability

≥ 95%; Fig. 4.2A; A4.1). The MRCA of clade 1 diverged at 25.47 Ma (20.08–31.37 Ma), and the MRCA of clade 2 at 29.32 Ma (24.24–34.62 Ma). Sundaland *Rhacophorus* belonged to both clades. Generally, clades exhibited two patterns: a collection of widespread species from several biogeographical areas, implicating dispersal-driven diversification, or a group of endemic species, implicating *in situ* diversification (Fig. 4.2). All but two Sumatran species were placed within clade 1, while all but two Bornean species were placed within clade 2. All Javan species were placed within clade 1, while species from the MP belonged to both clades (Fig. 4.2A).

Lineage-through-time plots

Our LTT plot of *Rhacophorus* suggested that species accumulated at a rate of one species per 16.7 Ma ($\lambda = 0.07$; Fig. 4.2B; 3). We recovered $\gamma = -3.37$, which suggested that *Rhacophorus* diversified at a slower-than-constant rate (critical value = -2.43; $p = 0.02$; Table 4.1). A visual inspection of the LTT plot revealed that diversification began to slow between 20–25 Ma after an initial burst (Fig. 4.3). The LTT plots of the other 14 taxa produced mean values of $\lambda = 0.13$ (stdev = 0.11) and $\gamma = -1.06$ (stdev = 2.16; Table 4.1). The γ statistic was negative in all but two taxa, but after corrected for incomplete taxon sampling, we were only able to reject a constant diversification rate in six taxa, which all diversified at a slower-than-constant rate except *Rana* (Table 4.1). We found that *Rhacophorus* had low λ and γ statistic values when compared with the other taxa (Fig. 4.4).

Ancestral range evolution

Our ancestral range evolution analyses supported BAYAREA+J

Fig. 4.3. Lineage-through-time plots sorted by speciation rate.

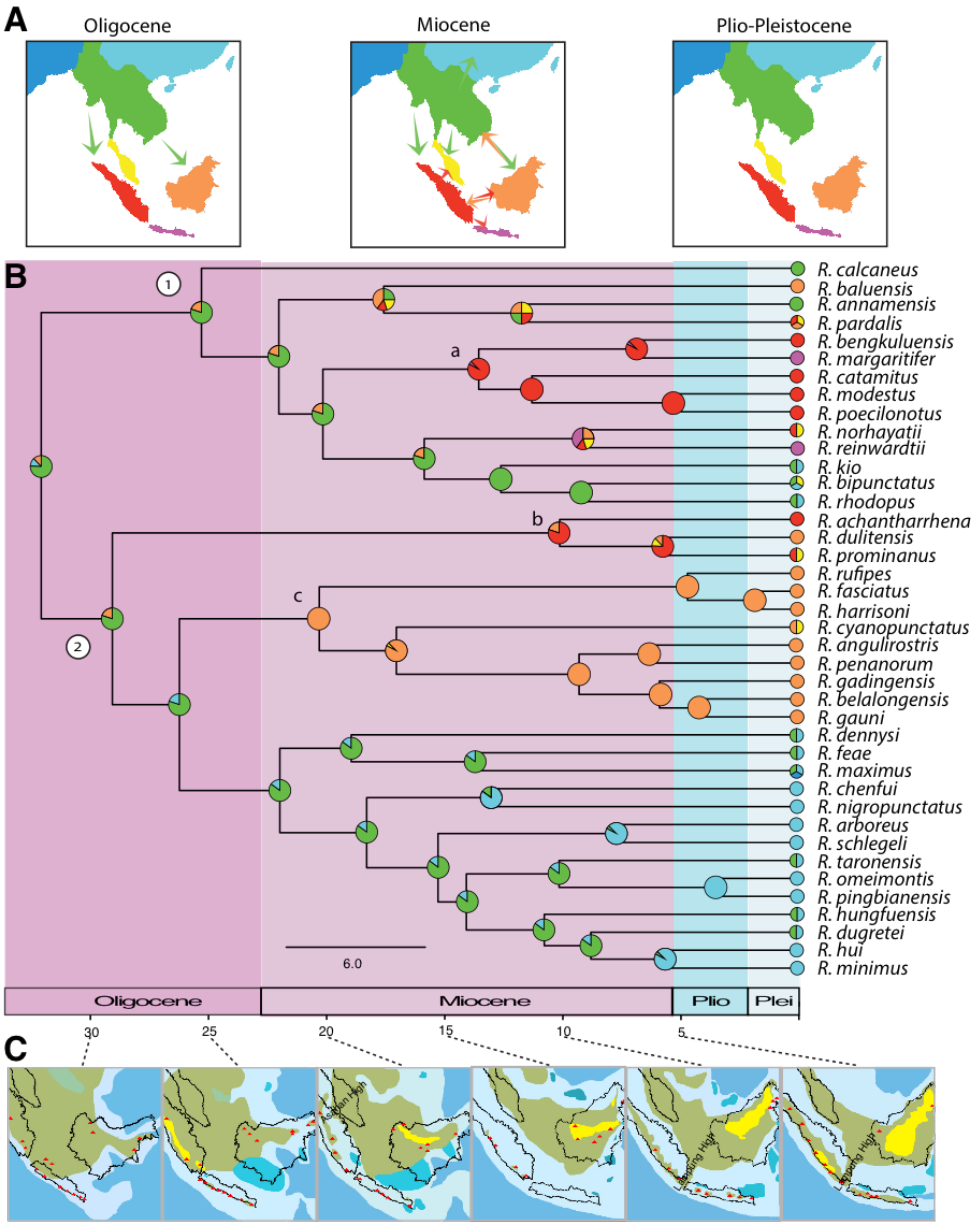

Fig. 4.5. Ancestral range reconstructions estimated in BioGeoBears. A) Maps showing the geographical areas defined in our analysis: Indian Peninsula (dark blue), East Asia (light blue), South East Asia (green), the MP (yellow), Sumatra (red), Borneo (orange), and Java (purple). Arrows signify dispersal events. The three maps correspond to three time periods: Oligocene, Miocene, and Pliocene-Pleistocene. B) Ancestral range reconstructions placed over the dated Bayesian phylogeny. Circles on nodes correspond to estimated ancestral ranges; colors correspond to the areas shown in A. Nodes with probability of more than one ancestral range are shown with pie charts. Clades 1 and 2 are labeled. C) Illustrations of land and sea cover for the Sunda Shelf at five million year intervals modified from Lohman et al., (2011) and Hall (2012b). Colors correspond to land (green), highlands (yellow), volcanoes (red triangles), shallow seas (light blue), deep seas (dark blue), lakes (dark green), and carbonate platforms (turquoise).

as the model with the best fit to our data, suggesting that founder event speciation was an important process in *Rhacophorus* speciation (LnL = -87.68; AIC = 181.37; Table 4.2). SE Asia received the highest probability as the ancestral range of the MRCA of *Rhacophorus* (Fig. 4.5B). In clade 1, we recovered dispersal events from SE Asia to Sumatra,

Borneo, Java, and the MP, as well as from Sumatra to Java (Fig. 4.5A,B). All dispersal events in clade 1 occurred during the Miocene (Fig. 4.5A,B). We also recovered evidence for extensive *in situ* diversification in clade 1 on Sumatra during the Miocene (node a; Fig. 4.5B). In clade 2, we recovered evidence for dispersal from SE Asia to Borneo and Sumatra in the Oligocene (Fig. 4.5A,B). Two subclades within clade 2 contained Sundaland species. One subclade dispersed from Sumatra to Borneo and the MP in the late Miocene (node b; Fig. 4.5A,B). The other subclade diversified *in situ* on Borneo from the early Miocene to the Pleistocene (node c; Fig. 4.5B). Dispersal events within clade 2 occurred during the Oligocene and Miocene (Fig. 4.5A,B). We recovered no dispersal events during the Pleistocene (Fig. 4.5).

DISCUSSION

We used mtDNA and nuDNA to investigate diversification dynamics on the Sunda Shelf. *Rhacophorus* are an ancient taxon that diversified slowly through time. We recovered two primary clades that originated in SE Asia 33.0 Ma (27.0–38.0 Ma; Figs. 2A, A4.1). Oligocene and early Miocene dispersal events initiated high rates of *in situ* diversification on Sumatra (55%) and Borneo (66%) beginning in the late Miocene (Fig. 4.5). All species that diversified *in situ* on Sumatra and Borneo are endemic to their respective islands.

Between-island dispersal has commonly occurred on the Sunda Shelf

We found that ancient dispersal events to Sumatra and Borneo during the Oligocene and early Miocene facilitated the within-island diversification that produced many Sundaland

Rhacophorus. These dispersal events occurred before peak sea levels in the middle Miocene, when Borneo was directly connected to SE Asia via the MP, and the Asahan High connected Sumatra to the mainland until the early Miocene (van Bemmelen, 1949; Meijaard, 2004; Barber, 2005). We propose that these connections facilitated these ancient dispersal events (Fig. 4.5A,B). Likewise, we found support for dispersal from Sumatra to Java in the late Miocene, when west Java was connected to southern Sumatra via the Lampung High (Fig. 4.5A,B; van Bemmelen, 1943; Hall 2012a).

Table 4.2: Complete parameters from our ancestral range reconstruction analysis.

Model	LnL	# Parameters	d	e	j	AIC	AIC wt	Delta AIC
BAYAREALIKE+J	-87.68875	3	0.001682662	7.64E-03	0.027363634	181.3775	0.34513144	–
DIVALIKE+J	-88.68622	3	0.006213599	1.37E-09	0.004347585	183.3724	0.127287473	1.9949
DIVALIKE	-88.80588	2	0.006631346	1.00E-12	0	181.6118	0.306983801	0.2343
DEC+J	-89.01036	3	0.005194147	1.00E-12	0.007778975	184.0207	0.092048369	2.6432
DEC	-89.6765	2	0.005751864	2.00E-09	0	183.353	0.128531006	1.9755
BAYAREALIKE	-98.55556	2	0.003632584	3.37E-02	0	201.1111	1.79E-05	19.7336

Past studies have also highlighted the importance of dispersal to Sundaland diversification, and found that dispersal either occurred during Pliocene-Pleistocene glacial cycles, or across more ancient land bridges during the Miocene and Oligocene. Leonard et al., (2015) compared 28 taxonomic groups across the Sunda Shelf and found that the oldest dispersal event was 3.9 Ma (Pliocene). De Bruyn et al., (2014) compared dispersal events across 61 taxa and found that 59% (80/135) of dispersal events on the Sunda Shelf occurred during the Pliocene or Pleistocene. Likewise, Ruedi (1996) and Gorog et al., (2004) found support among shrew species for Pliocene or Pleistocene dispersal and subsequent vicariance; primates show a similar pattern (Harrison et al., 2006). These dispersal events are similar to patterns observed in Philippine mammals, where recent between-island dispersal is the primary mechanism of

diversification (Esselstyn et al., 2009; Brown et al., 2013). Alternatively, many ancient dispersal events have led to *in situ* diversification in Sundaland. For example, de Bruyn et al., (2014) found that 35% (48/135) of dispersal events (and subsequent diversification) occurred during the Miocene. This disparity in dispersal times across taxa may reflect dispersal abilities of species, or, it may also reflect the increased land area available for dispersal during the Pliocene and Pleistocene. Thus, older dispersal events may have occurred less commonly due to fewer land bridges and higher sea levels, but they may have contributed more to species accumulation by promoting *in situ* diversification.

Within-island diversification was important to species accumulation on Sumatra and Borneo

In situ diversification occurred on Sumatra in the late Miocene, and on Borneo from the middle Miocene to the Pleistocene. We propose that the size of Sumatra and Borneo promoted *in situ* diversification, rather than stability or the level of isolation (the MP was less isolated and more stable than Sumatra**)**. Larger islands provide more opportunities for allopatric divergence, especially when rising seas provided barriers to dispersal on Sumatra (Heaney 1986, 2000; Meijaard, 2004; Hall 2012a). In addition, volcanic uplift during the late Miocene and early Pliocene may have promoted elevational partitioning on Borneo where endemic species are restricted to both high and low elevations, a similar scenario to *in situ* diversification observed in the Gulf of Guinea (Bell et al., 2015). In addition, larger and more complex islands contain greater niche variation, which may allow for ecological speciation (Losos & Ricklefs, 2009).

We also recovered a relationship between *in situ* diversification and endemism, a pattern observed in many other island taxa (Paulay, 1985; Gomez-Diaz et al., 2012; Blackburn et al., 2012; Kubota et al., 2017). Many past studies have identified Borneo as an important source of

endemic species, which may be explained by the high level of *in situ* diversification we recovered there (Harrison et al., 2006; Roberts et al., 2011; Klaus et al., 2013; de Bruyn et al., 2014; Janssens et al., 2016).

Diversification has proceeded slowly on the Sunda Shelf

We found that the pattern and tempo of diversification demonstrated by *Rhacophorus* was not exceptional compared with other Sunda Shelf taxa (Figs. 3, A4.2; Table 4.1). With a root age of 33.0 Ma, *Rhacophorus* were close to the mean root age of 47.7 Ma in our dataset (Table 4.1). The *Rhacophorus* speciation rate was well below the mean of $\lambda = 0.13$ (stdev = 0.11; Table 4.1). Likewise, the *Rhacophorus* γ value of -3.4 was close to the mean value of -1.6 among taxa where we rejected a constant rate. These findings suggest that Sundaland diversification usually pre-dated the Pliocene, and occurred at a constant, or slower-than-constant rate in all but a few taxa (Figs. 3; A4.2; Table 4.1). As an island system, diversification rates on the Sunda Shelf have largely reached equilibrium as extinction has outpaced speciation, a finding that is largely supported by past studies (McPeek, 2008; Roberts et al., 2011; Etienne et al., 2012; Klaus et al., 2013; Janssens et al., 2016). Fig. 4.3 also suggests that while diversification rate shifts are largely idiosyncratic, several taxa show rate increases during the past 2–5 Ma supporting a "glaciation species pump" hypothesis in some taxa. Yet, our data show that Pleistocene process did not contribute significantly to longer-term diversification patterns on the Sunda Shelf. We would also expect that diversification rates would increase upon arrival to a new island. In *Rhacophorus*, the diversification rate began to slow around the time Sumatran and Bornean clades invaded those islands.

DATA ACCESSIBILITY

Mitochondrial sequence data generated in this study has been accessioned with Genbank under the numbers KX398867, KX398877, KX398884, KX398889, KX398904, KX398920, KX398925, and nuclear sequences under KY886351– KY886358.

ACKNOWLEDGMENTS

We are grateful to the Ministry of Research and Technology of the Republic of Indonesia, RISTEK, for coordinating and granting research permission, and past and present representatives of LIPI at the Museum Zoologicum Bogoriense for facilitating export and field research permits. RISTEK and LIPI reviewed and approved our fieldwork in Indonesia and provided export permits for specimens to the United States for study and deposition at UTA. We want to thank D. Portik, E. Wostl, and other members of the Fujita and Smith labs for their valuable feedback throughout this project. We also want to thank the two reviewers for their helpful and constructive feedback.

REFERENCES

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

Barber AJ, Crow MJ, Milsom J (2005) Sumatra: geology, resources and tectonic evolution. Geological Society of London.

Bell RC, Drewes RC, Zamudio KR (2015) Reed frog diversification in the Gulf of Guinea: Overseas dispersal, the progression rule, and in situ speciation. *Evolution*, 69, 904-915.

Blackburn DC, Siler CD, Diesmos AC, McGuire JA, Cannatella DC, Brown RM (2013) An adaptive radiation of frogs in a Southeast Asian island archipelago. *Evolution*, 67, 2631-2646.

Borregaard MK, Amorim IR, Borges PAV, Cabral JS, Fernández-Palacios JM, Field R, Heaney LR, Kreft H, Matthews TJ, Olesen JM, Price J, Rigal F, Steinbauer MJ, Triantis KA, Valente L, Weigelt P, Whittaker RJ (2016) Oceanic island biogeography through the lens of the general dynamic model: assessment and prospect. *Biological Reviews*, 92, 830-853.

Bouckaert R, Heled J, Kühnert D, Vaughan T, Wu CH, Xie D, Suchard MA, Rambaut A, Drummond AJ (2014) BEAST 2: A Software Platform for Bayesian Evolutionary Analysis.

PLoS Computational Biology, 10, e1003537–6.
Brown, RM, Siler CD, Oliveros CH, Esselstyn JA, Diesmos AC, Hosner PA, Linkem CW, Barley AJ, Oaks JR, Sanguila MB (2013) Evolutionary processes of diversification in a model island archipelago. *Annual Review of Ecology, Evolution, and Systematics*, 44, 411–435.
Cornuault J, Warren BH, Bertrand JA, Milá B, Thébaud C, Heeb P (2013) Timing and number of colonizations but not diversification rates affect diversity patterns in hemosporidian lineages on a remote oceanic archipelago. *The American Naturalist*, 182, 820–833.
De Bruyn M, Stelbrink B, Morley RJ, Hall R, Carvalho GR, Cannon CH, van den Bergh G, Meijaard E, Metcalfe I, Boitani L (2014) Borneo and Indochina are major evolutionary hotspots for Southeast Asian biodiversity. *Systematic Biology*, 63, 879–901.
Drummond AJ, Bouckaert RR (2015) *Bayesian evolutionary analysis with BEAST*. Cambridge, UK: Cambridge University Press.
Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research*, 32, 1792-1797.
Esselstyn JA, Timm RM, Brown RM (2009) Do geological or climatic processes drive speciation in dynamic archipelagos? The tempo and mode of diversification in Southeast Asian shrews. *Evolution*, 63, 2595-2610.
Frost DR (2017) Amphibian Species of the World: an Online Reference. Version 6.0 (1/15/17). Electronic Database accessible at http://research.amnh.org/herpetology/amphibia/index.html. American Museum of Natural History, New York, USA.
Gillespie R (2004) Community assembly through adaptive radiation in Hawaiian spiders. *Science*, 303, 356-359.
Gomez-Diaz E, Sindaco R, Pupin F, Fasola M, Carranza S (2012) Origin and in situ diversification in Hemidactylus geckos of the Socotra Archipelago. *Molecular Ecology*, 21, 4074-4092.
Gorog AJ, Sinaga MH, Engstrom MD (2004) Vicariance or dispersal? Historical biogeography of three Sunda shelf murine rodents (Maxomys surifer, Leopoldamys sabanus and Maxomys whiteheadi). *Biological Journal of the Linnean Society*, 81, 91–109.
Hall R (2009). Southeast Asia's changing palaeogeography. *Blumea - Biodiversity, Evolution and Biogeography of Plants*, 54, 148–161.
Hall R (2012a) Late Jurassic–Cenozoic reconstructions of the Indonesian region and the Indian Ocean. *Tectonophysics*, 570, 1–41.
Hall R (2012b). A review of the Cenozoic palaeoclimate history of Southeast Asia. In DJ. Gower, K.G. Johnson, J.E. Richardson, B.R. Rosen, L. Ruber, and S.T. Williams (Eds.), *Biotic Evolution and Environmental Change in Southeast Asia*. Cambridge, UK: Cambridge University Press.
Haq BU, Hardenbol J, Vail PR (1987) Chronology of fluctuating sea levels since the Triassic. *Science*, 235, 1156–1167.
Harrison T, Krigbaum J, Manser J (2006) Primate biogeography and ecology on the Sunda Shelf Islands: a paleontological and zooarchaeological perspective. *Primate biogeography* pp. 331–372. New York, Springer.
Harvey MB, O'Connell KA, Barraza G, Riyanto A, Kurniawan N (2015) Two new species of Cyrtodactylus (Squamata: Gekkonidae) from the Southern Bukit Barisan Range of Sumatra and an estimation of their phylogeny. *Zootaxa*, 4020, 495–23.
Heaney LR (1986) Biogeography of mammals in SE Asia: estimates of rates of colonization,

extinction and speciation. *Biological Journal of the Linnean Society*, 28, 127–165.
Heaney LR (2000) Dynamic disequilibrium: a long-term, large-scale perspective on the equilibrium model of island biogeography. *Global Ecology and Biogeography*, 9, 59–74.
Inger RF, Voris HK (2001) The biogeographical relations of the frogs and snakes of Sundaland. *Journal of Biogeography*, 28, 863–891.
Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A, Markowitz S, Duran C, Thierer T, Ashton B, Meintjes P, Drummond A (2012) Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics*, 28, 1647–1649.
Kisel Y, Barraclough TG (2010) Speciation has a spatial scale that depends on levels of gene flow. *The American Naturalist*, 175, 316–334.
Klaus S, Selvandran S, Goh JW, Wowor D, Brandis D, Koller P. ... Yeo, DC (2013) Out of Borneo: Neogene diversification of Sundaic freshwater crabs (Crustacea: Brachyura: Gecarcinucidae: Parathelphusa). *Journal of Biogeography*, 40, 63-74.
Kubota Y, Kusumoto B, Shiono T, Tanaka T (2017) Phylogenetic properties of Tertiary relict flora in the East Asian continental islands: imprint of climatic niche conservatism and in situ diversification. *Ecography*, 40, 436-447.
Lanfear R, Calcott B, Ho SYW, Guindon S (2012) Partitionfinder: combined selection of partitioning schemes and substitution models for phylogenetic analyses. *Molecular Biology and Evolution*, 29, 1695–1701.
Leonard JA, Tex RJ, Hawkins MT, Muñoz Fuentes V, Thorington R, Maldonado JE (2015) Phylogeography of vertebrates on the Sunda Shelf: A multi-species comparison. *Journal of Biogeography*, 42, 871–879.
Janssens SB, Vandelook F, De Langhe E, Verstraete B, Smets E, Vandenhouwe I, Swennen R (2016) Evolutionary dynamics and biogeography of Musaceae reveal a correlation between the diversification of the banana family and the geological and climatic history of Southeast Asia. *New Phytologist*, 210, 1453-1465.
Li J-T, Li Y, Murphy RW, Rao D-Q, Zhang Y-P (2012) Phylogenetic resolution and systematics of the Asian tree frogs, Rhacophorus (Rhacophoridae, Amphibia). *Zoologica Scripta*, 41, 557–570.
Li J-T, Li Y, Klaus S, Rao D-Q, Hillis DM, Zhang Y-P (2013) Diversification of rhacophorid frogs provides evidence for accelerated faunal exchange between India and Eurasia during the Oligocene. *Proceedings of the National Academy of Sciences*, 110, 3441–3446.
Lohman DJ, De Bruyn M, Page T, Rintelen von K, Hall R, Ng PK, Shih H-T, Carvalho GR, Rintelen von T (2011) Biogeography of the Indo-Australian archipelago. *Annual Review of Ecology, Evolution, and Systematics*, 42, 205–226.
Losos JB, Schluter D (2000) Analysis of an evolutionary species–area relationship. *Nature*, 408, 847–850.
Losos JB, Ricklefs RE (2009) Adaptation and diversification on islands. *Nature*, 457, 830–836.
MacArthur RH, Wilson EO (1963) An equilibrium theory of insular zoogeography. *Evolution*, 17, 373–387.
MacArthur RH, Wilson E0 (1967) The Theory of Island Biogeography. In *Monographs in population biology.* Princeton, NJ: Princeton University Press.
Matzke NJ (2013) BioGeoBEARS: biogeography with Bayesian (and likelihood) evolutionary analysis in R scripts. *R package, version 0.2*, 1.
McPeek MA, Brown JM (2007) Clade age and not diversification rate explains species richness

among animal taxa. *The American Naturalist*, 169, E97–E106.
McPeek MA (2008) The ecological dynamics of clade diversification and community assembly. *The American Naturalist*, 172, E270-E284.
Meijaard E (2004) Solving mammalian riddles: a reconstruction of the Tertiary and Quaternary distribution of mammals and their palaeoenvironments in island South-East Asia. (Australian National University, Canberra) PhD thesis.
Morley R (2012) A review of the Cenozoic palaeoclimate history of Southeast Asia. In D.J. Gower, K.G. Johnson, J.E. Richardson, B.R. Rosen, L. Ruber, and S.T. Williams (Eds.), *Biotic Evolution and Environmental Change in Southeast Asia*. Cambridge, UK: Cambridge University Press.
Palumb SR, Martin AP, Romano S, Mcmilan WO, Stice L, Grabowski G (1991) *The simple fool's guide to PCR*. University of Hawaii.
Patino J, Whittaker RJ, Borges PA, Fernández-Palacios JM, Ah-Peng C, Araújo MB, ... Nascimento L (2017) A roadmap for island biology: 50 fundamental questions after 50 years of The Theory of Island Biogeography. *Journal of Biogeography*, 44, 963-983.
Paulay G (1985) Adaptive radiation on an isolated oceanic island: the Cryptorhynchinae (Curculionidae) of Rapa revisited. *Biological Journal of the Linnean Society*, 26, 95-187.
Pybus OG, Harvey PH (2000) Testing macro-evolutionary models using incomplete molecular phylogenies. *Proceedings of the Royal Society London*, 267: 2267–2272.
R Core Team (2016) R: A language and environment for statistical computing. Foundation for statistical computing, Vienna, Austria. Available from: https://www.R-project.org/ (accessed 25 June 2017).
Rabosky DL (2006) LASER: a maximum likelihood toolkit for detecting temporal shifts in diversification rates from molecular phylogenies. *Evolutionary bioinformatics online*, 2, 247.
Rabosky DL, Glor RE (2010) Equilibrium speciation dynamics in a model adaptive radiation of island lizards. *Proceedings of the National Academy of Sciences*, 107, 22178-22183.
Rabosky DL, Lovette IJ (2008) Explosive evolutionary radiations: decreasing speciation or increasing extinction through time?. *Evolution*, 62, 1866-1875.
Racine JS (2012) RStudio: A platform-independent IDE for R and Sweave. *Journal of Applied Econometrics*, 27, 167–172.
Rambaut A, Suchard MA, Xie D, Drummond AJ (2014) Tracer v1.6, Available from http://beast.bio.ed.ac.uk/Tracer.
Revell LJ (2012) phytools: an R package for phylogenetic comparative biology (and other things). *Methods in Ecology and Evolution*, 3, 217-223.
Roberts TE, Lanier HC, Sargis EJ, Olson LE (2011) Molecular phylogeny of treeshrews (Mammalia: Scandentia) and the timescale of diversification in Southeast Asia. *Molecular Phylogenetics and Evolution*, 60, 358–372.
Rohland N, Reich D (2012) Cost-effective, high-throughput DNA sequencing libraries for multiplexed target capture. *Genome research*, 22, 939-946.
Rubinoff D, Holland BS (2005) Between two extremes: mitochondrial DNA is neither the panacea nor the nemesis of phylogenetic and taxonomic inference. *Systematic Biology*, 54, 952-961.
Ruedi M, Fumagalli L (1996) Genetic structure of Gymnures (genus Hylomys; Erinaceidae) on continental islands of Southeast Asia: historical effects of fragmentation. *Journal of Zoological Systematics and Evolutionary Research*, 34, 153–162.

Sambrook J, Russell D (2001). Molecular Cloning: A Laboratory Manual, 3rd Edition. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.

Streicher JW, Harvey MB, Sheehy CM III, Anders B, Smith EN (2012) Identification and description of the tadpole of the parachuting frog *Rhacophorus catamitus* from southern Sumatra, Indonesia. *Journal of Herpetology*, 46, 503–506.

Van Bemmelen RW (1949) The geology of Indonesia. General geology of Indonesia and adjacent archipelagoes. Martinus Nijhoff, The Hague, The Netherlands.

Van der Meijden A, Vences M, Hoegg S, Boistel R, Channing A, Myer A (2007) Nuclear gene phylogeny of narrow-mouthed toads (Family: Microhylidae) and a discussion of competing hypotheses concerning their biogeographical origins. *Molecular Phylogenetics and Evolution*, 44, 1017-1030.

Vences M, Vieites DR, Glaw F, Brinkmann H, Kosuch J, Veith M, Meyer A (2003) Multiple overseas dispersal in amphibians. *Proceedings of the Royal Society of London B: Biological Sciences*, 270, 2435-2442.

Voris HK (2000) Maps of Pleistocene sea levels in Southeast Asia: shorelines, river systems and time durations. *Journal of Biogeography*, 27, 1153–1167.

Warren BH, Simberloff D, Ricklefs RE, … Thebaud C (2015) Islands as model systems in ecology and evolution: prospects fifty years after MacArthur-Wilson. *Ecology Letters*, 18, 200–217.

Weigelt P, Steinbauer MJ, Cabral JS, Kreft H (2016) Late Quaternary climate change shapes island biodiversity. *Nature*, 532, 99–102.

Whittaker RJ, Fernández-Palacios JM, Matthews TJ, Borregaard MK, Triantis KA (2017) Island biogeography: Taking the long view of nature's laboratories. *Science*, 357, eaam8326.

SYNCHRONOUS DIVERSIFICATION OF PARACHUTING FROGS (GENUS *RHACOPHORUS*) ON SUMATRA AND JAVA

Kyle A. O'Connell[1,2], Amir Hamidy[3], Nia Kurniawan[4], Eric N. Smith[1,2], Matthew K. Fujita[1,2]

[1]Department of Biology, The University of Texas at Arlington, Arlington, Texas 76019, USA.
[2]The Amphibian and Reptile Diversity Research Center, University of Texas at Arlington, Arlington, Texas 76010, USA.
[3]Research and Development Center for Biology, Indonesian Institute of Science (LIPI), Widyasatwaloka Building, Cibinong 16911, West Java, Indonesia.
[4]Department of Biology, Universitas Brawijaya, Jl. Veteran, Malang 65145, East Java, Indonesia.

Submitted, 2017

CHAPTER 5

SYNCHRONOUS DIVERSIFICATION OF PARACHUTING FROGS (GENUS *RHACOPHORUS*) ON SUMATRA AND JAVA

ABSTRACT

Geological and climatological processes can drive the synchronous diversification of co-distributed species. The islands of Sumatra and Java have experienced complex geological and climatological histories, including extensive sea level changes and the formation of valleys between northern, central, and southern components of the Barisan Mountain Range which may have promoted diversification of their resident species. We investigate diversification on these islands using 13 species of the parachuting frog genus *Rhacophorus*. We use both mitochondrial and nuclear sequence data, along with genome-wide SNPs to estimate phylogenetic structure and divergence times, test for synchronous diversification, and test demographic models to elucidate the drivers of diversification on these islands. We find support for synchronous divergence among sister species pairs from Sumatra and Java, as well as of populations of four co-distributed taxa on Sumatra. Our data suggest that divergence in several highland Sumatran species occurred in allopatry in highland refugia, followed by size changes. We conclude that divergence on Sumatra and Java was affected by changing sea levels that isolated populations in allopatry.

INTRODUCTION

Biotic responses to climatological or geological changes often drive diversification on tropical islands (Esselstyn et al., 2009). Climatic fluctuations can accelerate diversification by isolating species into refugia or by expanding suitable habitat, thus promoting dispersal (Nater et al., 2015). Likewise, geological changes can initiate diversification by isolating populations in allopatry. The Sunda Shelf (Sumatra, Java, Borneo, and the Malay Peninsula) has experienced a turbulent geological and climatological history from the Miocene to present (Lohman et al., 2011). Sumatra in particular has experienced dynamic tectonic processes, volcanism, dramatic sea level changes, and extensive connectivity with surrounding landmasses during the Pleistocene (Hall, 2001; 2002; 2009; 2011; 2012a; 2012b; Lohman et al., 2011). For most of the past 25 million years (Ma), highland habitats on Sumatra have remained tropical, while lowland forests were frequently inundated by marine incursions, and also experienced extensive cooling and drying (Hall, 2009; 2012a).

While past studies have largely focused on the role of Pleistocene sea level fluctuations on diversification on Sumatra and Java, few studies have investigated the role of Miocene-Pliocene sea level changes, or of the formation of physical barriers during this time period (Inger & Voris, 2001; Leonard et al., 2015). During much of the Miocene Sumatra was composed of several islands, with marine incursions serving as barriers to dispersal (van Bemmelen, 1979; Meijaard, 2004; Hall 2012a). From the early Miocene to ~15 Ma, a sea level high-stand persisted on Sumatra, transforming volcanic peaks into small islands (Baumann, 1982; Haq et al., 1987, Batchelor, 1979; Anderson et al., 1993; Collins et al., 1995; Lourens & Hilgen, 1997; Barber et al., 2005). From 14–9 Ma sea levels receded, presumably allowing for dispersal between previously isolated volcanic islands (Batchelor, 1979; Baumman 1982; Haq et al., 1987; Morley, 1999). This cycle continued, with sea levels rising from 8.5–6 Ma, receding from 5.8–5.4 My,

and again rising from 5–4 My (Baumann, 1982; Haq, 1987; Krantz, 1991; Anderson, 1993; Van der Bergh et al., 2001).

Furthermore, van Bemmelen (1949) hypothesized the persistence of two transverse inland seaways on Sumatra from the early Miocene onward that divided Sumatra between the northern and central components of the Barisan Mountain Range (just south of the Asahan High in the Padang Sidempuan Valley), and between the Gumai and Garba Mts (in the Pagar Alam Valley, Fig 1). These seaways formed in the early Miocene, and only completely subsided in the middle Pliocene due to Barisan Mountain uplift (van Bemmelen, 1949). As such, Sumatra was composed of at least three large islands for much of its geologic history, and even at times of low sea-levels (when marine incursions subsided), the persistence of the Padang Sidempuan and Pagar Alam Valleys likely maintained allopatric distributions of dispersal-limited species in the northern, central, and southern components of the Barisan Mountain Range (Meijaard, 2004).

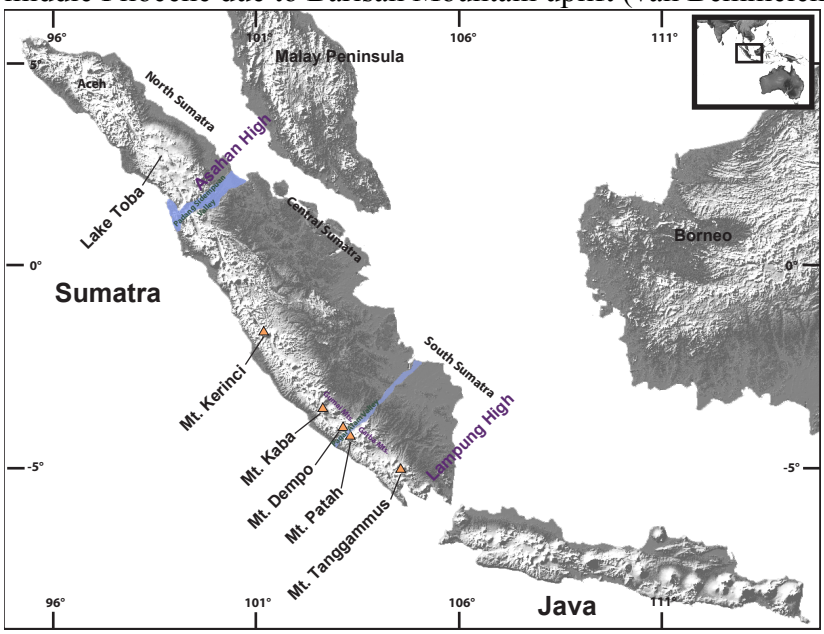

Fig. 5.1. Map of the islands of Sumatra and Java, showing their placement within the Sunda Shelf region. We also label historical and contemporary geological features on Sumatra referenced in this study.

Equivalently, Java was composed of small volcanic islands from 10 Ma onward, and did not completely emerge above sea level until ~5 Ma (Lohman et al., 2011). West Java may have been periodically connected to southern Sumatra via the Lampung High as early as the Mid-Miocene, allowing for early dispersal from southern Sumatra (van Bemmelen, 1949; Meijaard,

2004). Signals of these historical processes may be detected in the diversification histories, population structure and distribution of genetic diversity of extant biota (Weigelt et al., 2016, Portik et al., 2017; Xu & Hickerson, 2017). Under a comparative phylogeographic framework, shared diversification patterns between species can indicate synchronous responses to geological or climatological events (Hickerson et al., 2010; Bagley & Johnson, 2014; Smith et al., 2014; Prates et al., 2016).

We explore diversification processes on Sumatra and Java using species from the parachuting-frog genus *Rhacophorus*. *Rhacophorus* includes ~90 species distributed from the Indian peninsula to East and Southeast Asia (Frost, 2017). Sumatra and Java contain 16 described species of *Rhacophorus*, including *R. achantharrhena*, *R. barisani*, *R. bengkuluensis*, *R. bifasciatus*, *R. catamitus*, *R. cyanopunctatus*, *R. indonesiensis*, *R. margaritifer*, *R. modestus*, *R. nigropalmatus*, *R. norhayatii*, *R. pardalis*, *R. poecilonotus*, *R. prominanus*, *R. pseudacutirostris*, and *R. reinwardtii*. (Harvey et al., 2002; Streicher et al., 2012; 2014; Hamidy & Kurniati, 2015; O'Connell et al., *In Revision* (b)). On Sumatra, some species distributions span the length of the island, while others are restricted to small geographic areas (Harvey et al., 2002; Streicher et al., 2012; Hamidy & Kurniati, 2015). *Rhacophorus* occupy a variety of niche spaces, and most species' ranges are partitioned by elevation and island region (Harvey et al., 2002). On Sumatra, up to four highland endemic species occur in sympatry across the Barisan mountain range (KAO, personal observation). Java contains two species: *R. margaritifer* and *R. reinwardtii* (Streicher et al., 2012; Frost 2017, Fig. 4.2).

This study uses both mitochondrial and nuclear DNA sequence data, along with genome-wide SNPs, to pursue the following questions: 1) do species with similar geographic distributions

respond synchronously to geological and climatological events on islands? 2) What historical processes promoted these diversification events?

MATERIALS AND METHODS

Sampling and molecular sequence generation

Taxonomic sampling

The taxonomy of several *Rhacophorus* species is currently under review, thus we focused this study on 13 species. We extracted DNA from liver and thigh muscle tissue from 12 species from Sumatra and Java stored in SDS buffer or 70% ethanol. Our sampling included: *R. achantharrhena* (n = 8), *R. bengkuluensis* (n = 4), *R. catamitus* (n = 27), *Rhacophorus* sp. (n = 9), *R. cyanopunctatus* (n = 3), *R. margaritifer* (n = 5), *R. modestus* (n = 23), *R. nigropalmatus* (n = 1), *R. pardalis* (n = 3), *R. poecilonotus* (n = 25), *R. prominanus* (n = 5), *R. reinwardtii* (n = 4).

Molecular sequence data generation and alignments

We sequenced a 609 base pair fragment of the 16S ribosomal RNA gene following O'Connell et al., (2017a). To create a multi-locus concatenated alignment, we used BDNF data from O'Connell et al., (*In Review* (a)), and downloaded sequences from Genbank of all other available *Rhacophorus* (n = 56), at least one species of each genus within the family Rhacophoridae (n = 17), eight species of Mantellidae, and two outgroups (*Rana kukunoris*, and *Occidozyga lima*) following O'Connell et al., (*In Revision* (a)) and Li et al., (2013). Our dataset included sequences for 12S rRNA (n = 17), 16S rRNA (n = 180), Cytochrome oxidase c subunit

I (COI, n = 23), Cytochrome b (CYTB, n = 29), brain derived neurotrophic factor gene (BDNF, n = 30), pro-opiomelanocortin (POMC, n = 27), recombination-activating gene 1 (RAG1, n = 18), Rhodopsin (RHOD, n = 15) and Tyrosinase (TYR, n = 7). All information regarding sequence information and Genbank ID is in Table A4.1. We aligned each locus individually using the Geneious aligner using default parameters.

To place Sumatran and Javan species within a broad phylogenetic context, we created an alignment that included all available loci for one individual from each *Rhacophorus* species, and included all outgroups (phylogenetic dataset, n = 91). This dataset included one to nine loci for each sample (19 *Rhacophorus* species had only 16S data). For species distributed on Sumatra and another landmass (ex. *R. pardalis*), we included a Sumatran sequence as well as a sequence from the other landmass when available. We also created an alignment that would allow us to conduct comparative phylogeographic analyses across Sumatra and Java. This alignment included all individuals from the phylogenetic dataset, as well as all 16S sequences for Sumatran and Javan species that we generated (phylogeographic dataset, n = 181).

To form a more robust understanding of mitochondrial relationships within the genus, we generated partial mitochondrial genomes (mtgenome) for four Sumatran species. Using the method described by Fujita et al., (*In Preparation*), we generated ~13,000 bp of mitochondrial sequence data for *R. achantharrhena*, *R. modestus*, *R. poecilonotus*, and *R. catamitus*. Briefly, we digested the nuclear genome using plasmid safe DNAase, amplified the isolated mitochondrial template with whole genome amplification, and then prepared Illumina® genomic shotgun libraries for each sample. We sequenced our libraries on a partial lane of the Illumina® MiSeq at the University of Texas Arlington Genomics Core Facility (gcf.uta.edu, Arlington, TX,

USA). We filtered our raw reads using FASTX Toolkit (Gordon & Hannon, 2010), and assembled filtered contigs in CLC Genomics WorkBench (https://www.qiagenbioinformatics.com). We annotated each mtgenome in Mitos (Bernt et al., 2013). We downloaded additional mtgenomes of 15 species from Genbank: *Mantella madagascariensis*, *Buergeria buergeria*, *B. oxycephala*, *Gracixalus jinxiuensis*, *Raorchestes longchuanensis*, *Kurixalus odontotarsus*, *K. verrucosus,* *Chiromantis vittatus*, *Polypedates braueri*, *P. megacephalus*, *Rhacophorus dennysi*, *R. schlegelii*, *R. bipunctatus*, and *R. kio* (Table A4.1. We extracted coding sequences and the two ribosomal RNA sequences to create a concatenated mitochondrial alignment for phylogenetic analysis. Our alignment measured 10,837 base pairs (bp) in length, and was comprised of 12S (926 bp), 16S (1,544 bp), NADH dehydrogenase subunit 1 (ND1, 746 bp), NADH dehydrogenase subunit 2 (ND2, 622 bp), COI (1,488 bp), Cytochrome oxidase c subunit II (COII, 1,089 bp), ATP synthase subunit 8 (ATP8, 969 bp), ATP synthase subunit 6 (ATP6, 142 bp), Cytochrome oxidase c subunit III (COIII, 858 bp), NADH dehydrogenase subunit 3 (ND3, 1,315 bp), NADH dehydrogenase subunit 4L (ND4L, 250 bp), NADH dehydrogenase subunit 4 (ND4, 597 bp), and CYTB (279 bp). We refer to this as the mtgenome dataset.

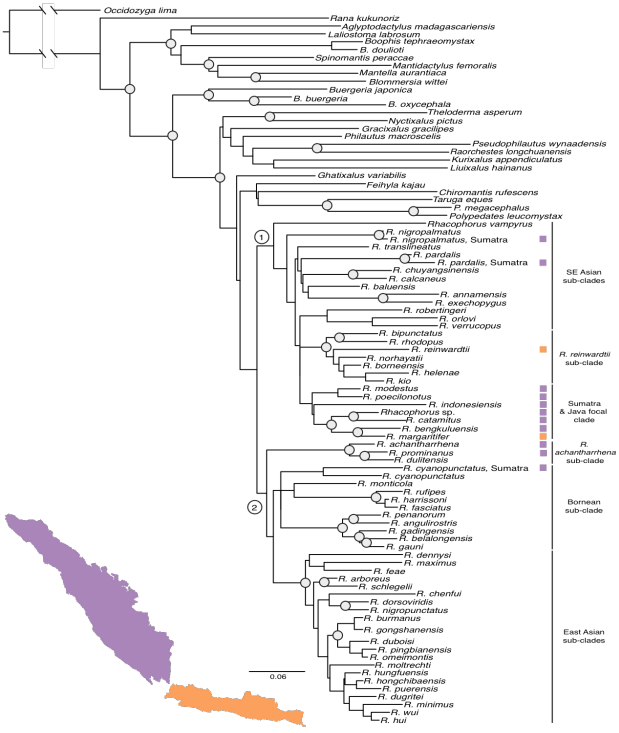

Fig. 5.2. Maximum likelihood phylogeny of 63 *Rhacophorus* species. Species found on Sumatra are designated with purple squares; species on Java are designated with orange squares. Subclades referenced in the study are labeled on the right. The two primary clades recovered in this study are labeled at the MRCA of each clade. Nodes with $\geq$ 70% bootstrap support are denoted with gray circles.

Genomic data generation and processing

We prepared double-digest restriction site associated DNA sequencing (ddRADseq) libraries for 60 individuals following Streicher et al., (2014a). We sequenced our final library (100 bp fragments, paired end run) on a partial lane of the Illumina® HISEQ 2500 at the University of Texas Southwestern Genomics Core facility (genomics.swmed.edu).

Double-digest RAD data were analyzed using the STACKS v1.37 pipeline (Catchen et al., 2013). After an initial round of data exploration that recovered very few SNPs, we removed all individuals with less than 500,000 reads, leaving 29 individuals. We followed the recommended workflow which implemented the following scripts and programs: (i) process_radtags, which filtered out reads below 90% quality score threshold, (ii) ustacks, which set a maximum distance of 4 between 'stacks', (iii) cstacks, which creates a catalogue of all of the loci within all individuals (-n flag, setting of 0) (iv) sstacks, which searches the stacks created in ustacks against the catalogue from cstacks, and (v) populations, which genotypes each individual according to the matched loci from sstacks (–r = 0.7). We further filtered our data with custom python scripts following O'Connell et al., (2017b) to remove loci with more than two haplotypes and invariant sites, and to remove individuals with more than 55% missing data. This allowed us to control the amount of missing data at the locus and individual level. We used these filtered data to create input files for downstream analyses. We analyzed each species group separately to produce four data sets which increased the number of shared loci and minimized missing data caused by allelic dropout (Arnold et al., 2013). Our filtering retained 17 individuals and 1,355 SNPs for *Rhacophorus* sp. and *R. catamitus*, 8 individuals and 2,939 SNPs for *R. catamitus*, 11 individuals and 2,387 SNPs for *R. poecilonotus*, and 4 individuals and 1,994 SNPs for *R. modestus*.

Phylogenetics and comparative phylogeography

Phylogenetic and divergence dating analyses

We selected the most probable model of nucleotide evolution for Bayesian inference (BI) and maximum likelihood (ML) analyses for all alignments using Bayesian information criteria implemented in PartitionFinder v.1.1.1 (Lanfear et al., 2012) partitioning by gene. The ML phylogeny for the phylogenetic dataset was estimated using raxmlGUI v1.3 (Silvestro & Michalak, 2012). Four gene partitions were defined: 12S and 16S, COI and CYTB, BDNF and RHOD, and POMC, RAG1, and TYR. We assigned a GTR + Γ rate to each partition and sampled 1,000 rapid bootstrap iterations.

Phylogeny and divergence times were estimated for Sumatran and Javan clades using the phylogeographic dataset in BEAST v.2.4.5 (Bouckaert *et al.,* 2014). We defined three gene partitions: 12S and 16S, CYTB and COI, and all five nuclear genes. Due to a lack of run convergence using a GTR model of nucleotide evolution (ESS values < 200), we assigned the HKY model to each partition (after Drummond & Bouckaert, 2015). Following Li et al., (2013), we calibrated the origin of Rhacophoridae to 53.2 Ma based on the fossil *Indorana prasadi.* We assigned a relaxed Log Normal clock model and the constant-growth coalescent tree prior to best estimate divergence times within species. A Log Normal calibration was assigned to the most recent common ancestor (MRCA) of all Rhacophoridae, with a mean of 1.0, a standard deviation of 1.25, and an offset of 52.3 (the age of the fossil). This produced a Rhacophoridae MRCA distribution 95% confidence interval (CI) of 52.3–57.6 Ma. All members of Rhacophoridae were constrained to monophyly. A uniform prior distribution was placed on the MRCA of *Boophis doulioti* and *Boophis tephraeomystax*, with a range of 0.0 to 15 Ma (the oldest estimated age for

the Comoro island of Mayotte where *B. tephraeomystax* is endemic, Vences et al., 2003). An exponential distribution with a mean of 10 was assigned to the ucldMean, and all other priors were left at default values. The analysis was run for 200,000,000 MCMC generations, sampling every 20,000 generations. We checked convergence of runs (ESS values > 200), and mixing, using Tracerv1.6 (Rambaut et al., 2014). We removed the first 25% of trees as burnin (2,500 trees) using TreeAnnotator (Bouckaert et al., 2014), and combined the remaining 7,500 trees to produce the maximum clade credibility tree with median node heights. BEAST and TreeAnnotator were run on Cipres web portal (Miller et al., 2010).

All steps of the divergence dating analysis were repeated using the mtGenome dataset with a few modifications. Four partitions were assigned: 12S and 16S; ND2, ATP6, ATP8, and NAD4L; COI; ND4, ND1, COII, COIII, CYTB, and NAD3, with a GTR + Γ model of nucleotide evolution on all partitions. The same fossil calibration was assigned to the MRCA of Rhacophoridae, but we were unable to include the *Boophis* calibration because mtgenomes were not available for those species. We assigned a Yule tree prior.

Genomic clustering

The program STRUCTURE (Pritchard et al., 2000) was used to explore how genomic variation is partitioned across Sumatra. The three species groups were analyzed separately (*Rhacophorus* sp. and *R. catamitus* analyzed together) using a range of *K* values (1–10), with five iterations per *K* value. Each analysis was run for 1,000,000 generations with a burn-in of 100,000 MCMC generations using the independent allele frequency and the admixture ancestry model. Results were summarized using the Evanno method (Evanno et al., 2005) implemented in STRUCTURE HARVESTER (Earl, 2012). We chose the highest DeltaK value and visually

inspected results files at each value of *K*. We used CLUMPP (Jakobsson & Rosenberg, 2007) to summarize population assignments across runs and created graphical summaries using DISTRUCT (Rosenberg, 2004).

Estimating effective migration and genetic diversity

We visualized patterns of historical migration and the spatial distribution of genetic diversity across Sumatra using the program EEMS (Petkova et al., 2015). We focused on *Rhacophorus* sp. and *R. catamitus* because our spatial sampling was limited in the other two species, and these two species are closely related and distributed across Sumatra (O'Connell et al., 2017a). EEMS estimates effective migration across the landscape by visualizing regions where genetic dissimilarity decays more quickly than expected under a model of isolation by distance. It relates effective migration rates to expected genetic dissimilarities to identify barriers to migration between populations, and estimates genetic diversity for each locality. We ran three independent chains using a deme size of 500 for 8,000,000 MCMC iterations, with 3,200,000 iterations of burnin and 9,999 thinning iterations. We checked for convergence and mixing, and visualized migration and diversity surfaces using rEEMSplots in Rstudiov3.1.1 (Racine, 2012; Petkova et al., 2015).

Evaluating diversification hypotheses

Testing for synchronous divergence

The hierarchical Approximate Bayesian Computation (hABC) program msBayes v.20140305 (Hickerson et al., 2007) was used to test for synchronous divergence on Sumatra and

Java. We ran two msBayes analyses using 16S sequence data; in both analyses, 1,000,000 simulations were drawn from the hyper-prior, and the local multinomial logistic regression was used to determine the number of divergence events. All other values were left as default settings. The first analysis tested for the synchronicity of divergence between four sister-species pairs on Sumatra and Java: *R. catamitus* and *Rhacophorus* sp., *R. modestus* and *R. poecilonotus*, *R. bengkuluensis* and *R. margaritifer*, and *R. achantharrhena* and *R. prominanus*. The analysis was also run without *R. achantharrhena* and *R. prominanus* because *R. achantharrhena* is sister to both *R. prominanus* and *R. dulitensis* (we lacked adequate sampling of *R. dulitensis* to test this pair directly), and because *R. prominanus* is not endemic to Sumatra (Malkmus, 2003; Frost, 2017). Our second analysis tested for synchronous divergence within co-distributed high and middle elevation Sumatran species at the oldest cladogenetic event within each species. Population pairs included central and southern *R. catamitus*, northern and central/southern *R. modestus*, central and northern/southern *R. poecilonotus*, and northern and southern *R. bengkuluensis*.

Demographic model-testing using diffusion approximation

Population genetic models of divergence were compared by using δ*a*δ*i* to analyze two-dimensional joint site frequency spectra (2D-JSFS; Gutenkunst et al., 2009). We assigned individuals to populations based on the results of STRUCTURE analyses (Fig. 5A–C). For each dataset, the folded 2D-SFS was generated from the SNP data, and to account for missing data we down-projected all datasets: *Rhacophorus* sp./*catamitus* (*Rhacophorus* sp.: 12 alleles, *R. catamitus*: 10 alleles, 1,223 segregating sites), *R. catamitus* (central: 3 alleles, south: 3 alleles, 2,109 segregating sites), *R. modestus* (northern: 6 alleles, central: 2 alleles, 1,960 segregating

sites), and *R. poecilonotus* (northern: 10 alleles, southern: 4 alleles, 1,446 segregating sites). We performed pairwise comparisons by comparing 15 models that varied in their magnitude and direction of size change and migration, and period of isolation (Portik et al., 2017; Table A5.2). We generated 20 sets of randomly perturbed parameters for each model, and optimized each parameter using the Nelder-Mead method for a maximum of 50 iterations. We used each optimized parameter to simulate the 2D-JSFS, and estimated the log-likelihood of the 2D-JSFS given the model using a multinomial approach. Using the best scoring replicate for each parameter, we conducted a second round of perturbation with 50 optimization iterations, followed by a final round of 100 optimization replicates. The final 2D-SFS was simulated from each parameter set, and extrapolation was performed for all analyses with grid sizes of 50, 60, and 70. Log-likelihoods were estimated using the multinomial approach, and models were evaluated using the Akaike information criterion (AIC) based on the replicate with the highest log-likelihood score. AIC model weights (Burnham & Anderson, 2003) were calculated for each model using R (R Core Team, 2013). We identified the best-supported model(s) based on a cutoff value of delta AIC value greater than three, and a model weight greater than 10%, which represented a natural point of differentiation between models (Table A5.2).

RESULTS

Phylogenetic, phylogeographic, and divergence dating analyses

Sumatran and Javan *Rhacophorus* are not monophyletic

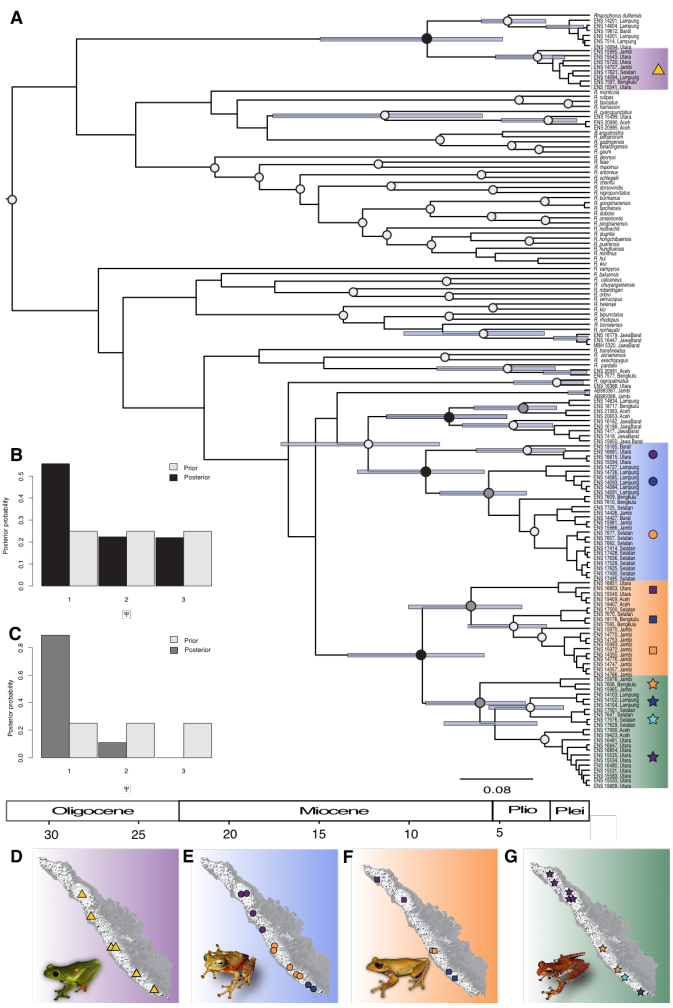

Fig. 5.3. A) Bayesian phylogeny of all sampled *Rhacophorus* species with a focus on population level sampling for species from Sumatra and Java. Nodes with ≥ 95% posterior probability are marked by light-gray circles. Species from Sumatra and Java are labeled on the right. The colored shapes correspond to the population assignment of each clade in focal Sumatran species shown in the maps below: purple (northern), orange (central), and blue (southern). Yellow signifies no phylogenetic structure. Shapes correspond to species identities: triangle = *R. achantharrhena*, circle = *Rhacophorus sp.* and *R. catamitus*, square = *R. modestus*, and star = *R. poecilonotus*. B–C) Histogram showing the probability of the number of divergence events inferred from hABC analysis. The color of the posterior estimates B = black and C = dark gray, correspond to the colored and magnified node circles on the phylogeny in 3A. 3B shows the analysis with sister species pairs, including *R. achantharrhena* and *R. prominanus*, *R. bengkuluensis* and *R. margaritifer*, *Rhacophorus sp.* and *R. catamitus*, and *R. modestus* and *R. poecilonotus*. 3C shows the histogram testing the synchrony of the oldest divergence event within co-distributed Sumatran species, including *R. bengkuluensis*, *R. catamitus*, *R. modestus*, and *R. poecilonotus*. D–G) Maps of sampling localities and population assignments for individuals. Maps correspond to different focal species: D = *R. achantharrhena,* E = *Rhacophorus* sp. and *R. catamitus*, F = *R. modestus*, and G = *R. poecilonotus*.

Our ML analysis of the phylogenetic dataset recovered two primary clades within *Rhacophorus* (Fig. 5.2). Most of the deeper nodes within *Rhacophorus* were recovered with low bootstrap support (≤ 70%), while relationships within subclades were generally well supported (Fig. 5.2). Sumatran and Javan species belonged to both primary clades. Clade 1 included species from South East (SE) Asia, Borneo, Sumatra, and the Malay Peninsula grouped into four sub-clades, and included 9/14 species from Sumatra and Java (Fig. 5.2). Seven species endemic to Sumatra and Java belonged to a single subclade: *R.*

indonesiensis, *R. poecilonotus*, *R. modestus*, *R. catamitus*, *Rhacophorus* sp*., R. bengkuluensis*, and *R. margaritifer*. In Clade 2, we recovered five sub-clades composed primarily of species from Borneo and SE and East Asia (Fig. 5.2). Sumatran species belonged to two subclades. *Rhacophorus achantharrhena*, *R. prominanus*, and *R. dulitensis* formed their own subclade (*R. achantharrhena* subclade), and *R. cyanopunctatus* was most closely related to Bornean species (Fig. 5.2).

Our BI analysis using mtgenomes was fully resolved and recovered the two clades described above, with Clade 1 containing *R. catamitus*, *R. poecilonotus*, and *R. modestus*, and Clade 2 containing *R. achantharrhena* (Fig. A1). *Rhacophorus achantharrhena* was more closely related to the East Asian *Rhacophorus* than to the other Sumatran species as found in the analysis above.

Congruent phylogeographic structure on Sumatra

Our Bayesian analysis of the phylogeographic dataset recovered a similar topology to our genus-wide analysis, recovering Clades 1 and 2 with low support at deeper nodes (Fig. 5.3A). Within Sumatra and Java, we recovered seven species that demonstrated negligible within-island population divergence (*R. cyanopunctatus*, *R. prominanus*, *R. achantharrhena*, *R. nigropalmatus*, *R. pardalis*, *R. reinwardtii*, and *R. indonesiensis*), and six species with substantial population structure (*R. bengkuluensis*, *R. margaritifer*, *R. poecilonotus*, *R. modestus*, *Rhacophorus* sp., and *R. catamitus*, Fig. 5.3A). We included five species with multi-landmass distributions within the Sunda Shelf, but only *R. cyanopunctatus* demonstrated deep phylogenetic structure between landmasses (note that *R. reinwardtii* and *R. prominanus* were only sampled from a single landmass). Among species endemic to Sumatra or Java, all exhibited

phylogenetic structure except *R. achantharrhena* (Fig. 5.3A,D). Several Sumatran species shared two congruent phylogenetic breaks: a northern and central break between Lake Toba and Mount Kerinci, and a central and southern break between the Gumai and Garba Mountains (Fig. 5.1). For example, the boundary between *Rhacophorus* sp. and *R. catamitus* lies between Lake Toba and Mt. Kerinci, and *R. catamitus* clades exhibit a phylogenetic break at the Pagar Alam Valley (Fig. 5.3A,E). Likewise, in *R. modestus* northern and southern clades are divided between Lake Toba and Mt. Kerinci, and central and southern clades are divided between Mt. Kerinci and the Gumai Mts. (Fig. 5.3A,F). In *R. poecilonotus*, the northern and central clades are divided between Lake Toba and Mt. Kerinci, and the Pagar Alam Valley divides the central and southern clades (Fig. 5.3A,G). Finally, *R. bengkuluensis* exhibits a phylogenetic break between mountains in Aceh province and Mts. Kaba and Tangammus (Fig. 5.3A).

Divergence dating suggests synchrony of divergence

We used divergence dating to estimate the temporal congruence between divergence times (Figs. 3A, 4). We recovered substantial overlap in divergence times between species pairs on Sumatra and Java, including *R. catamitus* and *Rhacophorus* sp. at 9.07 Ma (5.83–12.85 Ma), *R. modestus* and *R. poecilonotus* at 9.26 Ma (5.84–13.34 Ma), *R. bengkuluensis* and *R. margaritifer* at 7.77 Ma (4.59–11.30 Ma), *R. cyanopunctatus* on Sumatra and Borneo at 11.34 Ma (5.93–17.57 Ma), and *R. achantharrhena* and *R. prominanus*/*R. dulitensis* at 9.0 Ma (4.83–14.92 Ma, Figs. 3A, 4). The oldest cladogenetic events within co-distributed highland and middle elevation species were also largely congruent (Figs. 3A, 4B). In *R. catamitus* this corresponded to the divergence between central and southern clades at 5.58 Ma (3.46–8.31 Ma). In *R. modestus* the oldest divergence event occurred between the northern and central/southern clades

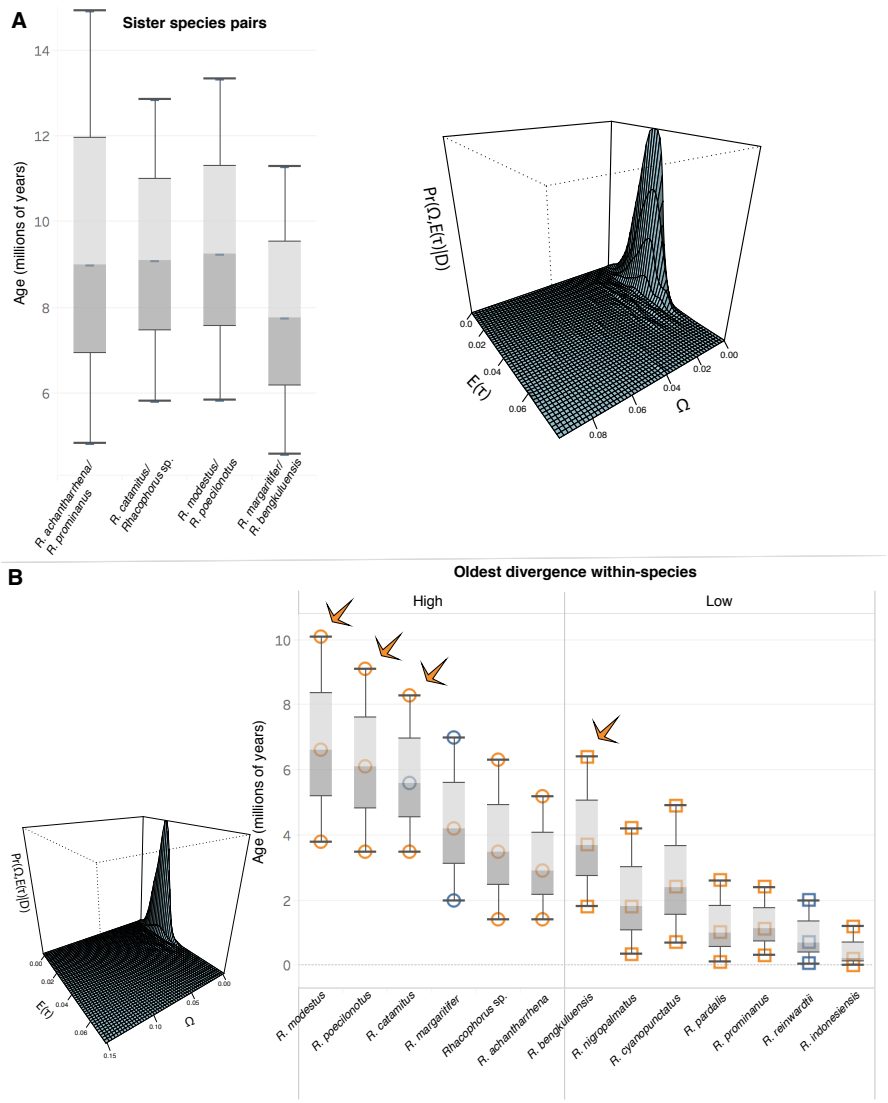

Fig. 5.4. Box and whisker plots showing the mean and 95% CI of divergence date estimates from BEAST output. A) Divergence date estimates of selected sister pairs used to test for synchronous divergence and the joint posterior probability of the average divergence time and the variance in divergence times/average divergence time. This plot shows that while more than one divergence event is supported, they occurred in close succession. B) Divergence date estimates of the oldest cladogenetic event within each Sumatran and Javan species. Arrows show the four species used in the hABC analysis. The plot shows that one divergence event was best-supported for these four Sumatran species.

at 6.56 Ma (3.77–10.10 Ma), while in *R. poecilonotus* it corresponded to the divergence between the central and northern/southern clades at 6.10 Ma (3.55–9.10 Ma). Finally, in *R. bengkuluensis* the oldest divergence event was between the northern and southern clades at 3.70 Ma (1.81–6.38 Ma, Figs. 3A, 4B). In other Sumatran and Javan species, the oldest divergence event within the species was much younger, including *R. achantharrhena* at 2.90 Ma (1.39–5.20 Ma), *R. prominanus* at 1.12 Ma (0.31–2.41 Ma), and *R. reinwardtii* at 0.71 Ma (0.06–1.97 Ma, Figs, 3A, 4B). Lowland species were generally younger than highland species (Fig. 5.4B).

Divergence dating of the mtgenomes revealed older divergence dates compared to the phylogeographic dataset (Fig. A1). We recovered a divergence time of 18.45 Ma (14.00–23.81 Ma) between *R. catamitus* and *R. modestus*/*R. poecilonotus*, and of 10.70 Ma (7.12–14.80 Ma) between *R. modestus* and *R. poecilonotus*. Within Clade 2, we found that *R. achantharrhena* diverged from the East Asian sub-clade 25.19 Ma (19.26–31.55 Ma).

Genomic clustering analyses support congruent population structure on Sumatra

Our Bayesian clustering results from STRUCTURE supported one population within *Rhacophorus* sp., and two populations within *R. catamitus*, *R. poecilonotus* and *R. modestus* (Fig. 5.5A–F). We recovered a single northern population of *Rhacophorus* sp., and central and southern populations within *R. catamitus* (Fig. 5.5A,D). We observed mixed assignment probabilities in two central individuals of *R. catamitus* (Fig. 5.5A,D). In *R. poecilonotus* we observed differentiation between individuals in the north and south, with mixed assignment probabilities in three of the northern individuals (Fig. 5.5B,E). In *R. modestus* we recovered one northern and one central population (Fig. 5.5C,F).

Estimates of gene flow and genetic diversity reveal two barriers to gene flow

We used the program EEMS to estimate effective migration surfaces and levels of genetic diversity for the three *Rhacophorus* sp. and *R. catamitus* populations (Fig. 5.5G,H). EEMS recovered less gene flow than expected between three populations under a model of isolation by distance (Fig. 5.5G). The greatest barrier to gene flow was at the species boundary between *Rhacophorus* sp. and central *R. catamitus* (Fig. 5.5G). We found evidence for low gene flow between demes in the central population, and moderate gene flow within the southern

population. We found that genetic diversity was highest in *R. catamitus* in the central population and in *Rhacophorus* sp. at the northern end of the northern population, while it was lowest around Lake Toba (Fig. 5.5H).

Testing diversification hypotheses

Divergence is largely synchronous on Sumatra and Java

We found evidence for synchronous diversification among sister species pairs and between co-distributed populations (Figs. 3B–C, 4). In our first analysis of sister species pairs, we recovered the highest posterior support for one divergence event (PP = 0.56), but also received support for two (PP = 0.22) or three divergence events (PP = 0.22, Figs. 3B, 4A). The HDP interval of the dispersion index of divergence times, Ω was 0–0.01, indicating that minimal time elapsed between divergence events. When we included only species endemic to Sumatra and Java (excluded *R. achantharrhena* and *R. prominanus*), we recovered strong support for only one divergence event (PP = 0.98), with an Ω HDP interval of 0–0.003. When investigating the oldest cladogenetic event within co-distributed species, we found support for a single divergence event (PP = 0.89), but some support for a second divergence event (PP = 0.11), with an Ω HDP interval of 0–0.002 (Figs. 3C, 4).

Demographic model testing

Among our four comparisons, we were unable to identify a single best model. However, we identified trends in each population comparison that helped differentiate between species-specific and community-wide responses to environmental changes (Table A5.2). The analysis of *Rhacophorus* sp. and *R. catamitus* found support for two models: divergence with size change and secondary contact with symmetrical gene flow (model weight = 0.42), and divergence with symmetrical gene flow but no size change (ΔAIC = 1.88, model weight = 0.18). Both models supported a much larger N_e in

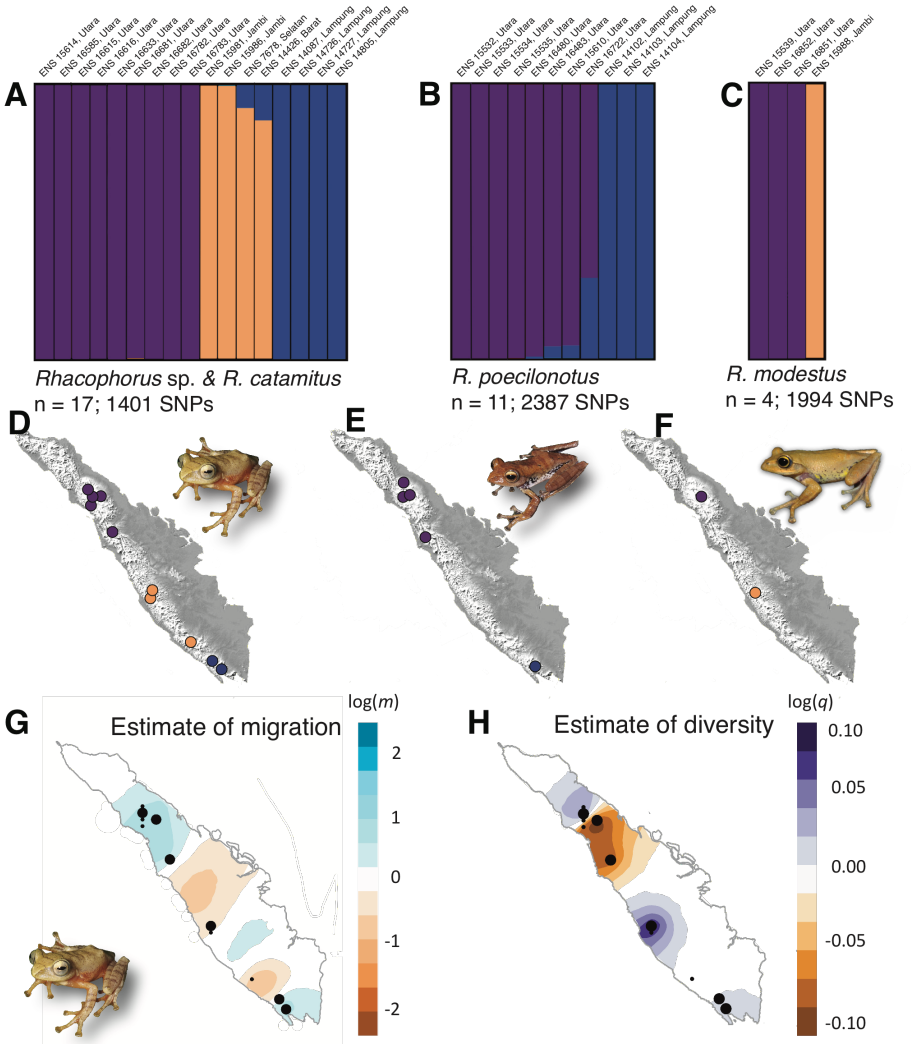

Fig. 5.5. A-C) STRUCTURE plots for *Rhacophorus* sp. and *R. catamitus*, *R. modestus*, and *R. poecilonotus*. Colors correspond to population assignments, where purple = northern, orange = central, and blue = southern. D-F) Maps of population assignments for SNP data as inferred by STRUCTURE. G) Estimates of effective migration for *Rhacophorus* sp. and *R. catamitus* inferred by EEMS. Blue colors indicate high levels of migration, orange colors indicate low levels of migration. Black circles indicate sampling localities, larger circles represent multiple samples from a single locality. H) Estimates of effective genetic diversity for *Rhacophorus* sp. and *R. catamitus*. Purple colors indicate high levels of genetic diversity, orange colors indicate low levels of genetic diversity. The highest genetic diversity is recovered in the central population of *R. catamitus*.

R. catamitus and symmetrical gene flow between the two species. In the model that included size change, we found that both species expanded, although *Rhacophorus* sp. expansion was minor and *R. catamitus* N_e remained substantially larger (Table A5.2).

The analysis of central and southern *R. catamitus* found support for three models: divergence with size change, but not migration (model weight = 0.47), divergence with ancient symmetrical gene flow followed by size change (ΔAIC = 2.56, model weight = 0.13), and divergence with symmetrical gene flow and size change (ΔAIC = 2.96, model weight = 0.11). In each model, the southern N_e was larger than the central population in the past, but after a central expansion and southern contraction, the central N_e was larger in the present. Second, in both models with gene flow, it was symmetrical, although the best-supported model did not support gene flow. Finally, we found no evidence of secondary contact.

In *R. poecilonotus*, we found support for two models: divergence with size change and secondary contact with asymmetrical gene flow (model weight = 0.46), and divergence with size change and secondary contact with symmetrical gene flow (ΔAIC = 0.38, model weight = 0.38). Both models support divergence without gene flow, and suggest that gene flow resumed after initial isolation in conjunction with population expansion. Across both models, the southern effective population size (N_e) was larger than the northern N_e in the past, but after size change, the northern population was larger.

Model testing of *R. modestus* populations found support for two simple models: divergence with no migration or size change (model weight = 0.43), and divergence with symmetrical migration (ΔAIC = 1.34, model weight = 0.22). In both models, central N_e was much larger than northern N_e. We found no evidence of secondary contact in our top two models.

DISCUSSION

We used a combination of genetic and genomic data to investigate the patterns and processes of diversification on Sumatra and Java. We found evidence of synchronous diversification between three Sumatran and Javan sister species pairs, as well as between populations of four species on Sumatra. SNP-based demographic model testing suggested that *R. poecilonotus*, *Rhacophorus* sp., and *R. catamitus* lineages diversified in isolation and later experienced population size changes. *Rhacophorus* sp. and *R. modestus* originated in northern Sumatra, while *R. poecilonotus* and *R. catamitus* originated in central or southern Sumatra, providing evidence of synchronous allopatric diversification of two sister pairs. We discuss the implications of these findings and suggest some future directions for Sumatran and Javan phylogeography.

Diversification on Sumatra and Java was largely synchronous

We recovered a strong signal of synchronous diversification across three species pairs on Sumatra and Java (Figs. 3B–C, 4, A2). With the inclusion of a fourth species pair, *R. achantharrhena* and *R. prominanus*, we recovered support for up to three divergence events (Fig. 5.3B). Although we found evidence of synchronous divergence in this subset of species, multiple cycles of diversification likely produced the full number of *Rhacophorus* species on Sumatra and Java (Fig. 5.3), possibly corresponding to cycles of marine incursion during the Miocene and Pliocene. The mean divergence date estimate for sister species pairs was 9.0 Ma (Fig. 5.4A). During this time in the late Miocene, sea levels were low, indicating that the two inland seaways were likely not present during this time, but the underlying valleys (Padang Sidempuan and

Pagar Alam Valleys) may have served as barriers to dispersal of highland *Rhacophorus* (van Bemmelen, 1949, Haq et al., 1987, Meijaard, 2004, Lohman et al., 2011, Hall, 2012b).

We also found support for synchrony of the oldest cladogenetic event within species from the Sumatran/Javan focal clade, with a mean age of 5.6 Ma (Figs. 3A, 4B). This was a time of high sea levels on Sumatra, which may have isolated the three components of the island into northern, central, and southern units (Meijaard, 2004). This was also a time of increased mountain building and subsequent volcanism (Barber et al., 2005). Demographic modeling suggested that populations of *R. catamitus* and *R. poecilonotus* diverged in allopatry, but later experienced size change. Thus, these populations may have been isolated by high sea levels, but afterward dispersed across the island and expanded (Table A1). The correspondence of population boundaries across Sumatra with the two hypothesized marine incursions in the north and south suggest that these barriers drove Sumatran diversification more than volcanic activity. We propose that marine incursions, and their underlying valleys were the primary barriers to dispersal on Sumatra during the Miocene and Pliocene, and promoted the synchronous divergence we observed. This also supported by the finding that two sister pairs (*Rhacophorus* sp. and *R. catamitus*, and *R. modestus* and *R. poecilonotus*) diversified in allopatry in northern and central/southern Sumatra.

Rhacophorus comparative phylogeography

Few studies have investigated phylogeographic patterns across Sumatra, instead focusing on geographically restricted species (Brandon-Jones, 1996; Nater et al., 2011; 2012), or on lowland taxa with little within-island phylogenetic structure (Leonard et al., 2015). In fact, most studies that have included Sumatran and Javan taxa have focused on regional Sundaland patterns,

rather than on within-island patterns (Inger & Voris, 2001; De Bruyn et al., 2014; Leonard et al., 2015). Island-wide divergence patterns have also been complicated by high levels of endemism to single mountains of many Sumatran and Javan taxa (Esselstyn et al., 2013; Harvey et al., 2014; Demos et al., 2016; Harvey et al., 2017). This is one of the first studies to investigate phylogeographic patterns of multiple species across Sumatra.

Among Sumatran and Javan *Rhacophorus*, some species exhibited extensive phylogeographic structure, while others exhibited very little. We propose that this discrepancy in the level of phylogenetic structure on Sumatra or Java could be due to three reasons. The first is the age of the species. We found that younger species generally exhibited less phylogenetic structure (Fig. 5.3A, 4B). Younger species have had less time to accumulate genetic variation, and may have diverged after the cessation of historical processes that shaped older highland taxa (i.e. marine incursions). Second, elevation may influence the level of phylogenetic structure. When we plotted the age of the oldest divergence within each species in Fig. 5.4B, a clear pattern emerged of older divergences (and usually more phylogenetic structure; Fig. 5.3A) in highland taxa. This was likely biased by our reduced sampling of lowland species, but may explain why we observed very little differentiation in *R. nigropalmatus* and *R. pardalis* between Sumatra and adjacent landmasses. This may also explain the reduced structure observed in *R. prominanus* and *R. reinwardtii*. With most of Sumatran lowland forest inundated by inland seaways, and Java only emerging from the ocean ~5–10 Ma, there may have been few opportunities for lowland species to colonize Sumatra or Java until the recent past. In addition, Plio-Pleistocene climate conditions may have presented fewer barriers to gene flow in lowland species compared with highland species as sea levels receded. Nonetheless, this does not explain why *R. achantharrhena* exhibits shallow phylogenetic structure across Sumatra despite being a high

elevation species. Also, *R. prominanus* and *R. bengkuluensis* are both middle elevation species, yet they exhibit very different phylogeographic patterns across Sumatra (high verses low levels of phylogeographic structure). Additionally, *R. cyanopunctatus* is a low to middle elevation species that shows deep divergence from Malaysia (Fig. 5.3A). As such, we propose that life history differences, and specifically reproductive histories, also explain why some species exhibit deep phylogenetic structure while others do not. *Rhacophorus* are foam nesters, and are thought to breed either in streams, or in ephemeral pools and wetlands (Streicher et al., 2014b). Species from the focal clade (and their larvae) were usually collected near streams, and it is likely that they are stream breeders (Harvey et al., 2002; KAO Personal Comment). On the other hand, species such as *R. reinwardtii* and *R. dulitensis* (from Borneo) are thought to breed in pools on the forest floor (Malkmus et al., 2002). While the reproductive strategy of *R. achantharrhena* is unknown, we infer that it likely shares a similar reproductive strategy to its close relative, *R. dulitensis,* which breeds in ephemeral pools or wetlands. Different *Rhacophorus* also inhabit different niche spaces within the forest. Species from the *R. reinwardtii* subclade inhabit the canopy (Onn & Grismer, 2010), while species from the focal clade largely inhabit low vegetation near streams (Harvey et al., 2002; Streicher et al., 2012, 2014b). This difference in niche, as well as dispersal ability of canopy verses shrub species may also greatly influence phylogenetic structure in *Rhacophorus* as seen in other frog species (Bell et al., 2017).

Conclusions and future directions

Using 13 species of parachuting frogs, we described patterns of diversification on Sumatra and Java. We found that marine barriers and their underlying valleys likely drove

divergence between and within species. Following divergence in allopatry, we found signals of expansion and contraction in two highland species, suggesting that cyclical marine incursions isolated species during the Mio-Pliocene.

Parachuting frogs of the Sunda Shelf are an ideal system for studying diversification on islands. However, additional data are needed to more fully elucidate the processes driving diversification. First, the life history characteristics of Sumatran and Javan species need to be more fully understood to connect ecology with diversification. For example, the reproductive strategy of most Sumatran species is uncertain. Second, we need to collect more loci, perhaps using a target capture approach, to better resolve phylogenetic relationships and uncover phylogeographic patterns. Most of the deeper nodes in our mitochondrial phylogeny remained unresolved, and the branching order of clades in *R. poecilonotus* was also uncertain. Finally, exciting advancements in demographic modeling will allow us to test for not only co-expansion or contraction using SNP data, but also to conduct robust tests of synchronous diversification using thousands of SNPs (Xu & Hickerson, 2017).

ACKNOWLEDGMENTS

We are grateful to the Ministry of Research and Technology of the Republic of Indonesia, RISTEK, for coordinating and granting research permission. We are grateful to past and present representatives of LIPI at the Museum Zoologicum Bogoriense for facilitating export and field research permits, namely Boadi, M. Amir, R. Ubaidillah, I Sidik, and Ir. R. M. Marwoto. RISTEK and LIPI reviewed and approved our fieldwork in Indonesia and provided export permits for specimens to the United States for study and deposition at UTA. We want to thank E. Wostl and D. Portik, as well as other members of the Fujita and Smith labs for their feedback on

this manuscript. This work was supported by a National Science Foundation grant (DEB-1146324) awarded to ENS and MBH, and the University of Texas at Arlington.

DATA ACCESSIBILITY

Mitochondrial data generated for this study can be located under the Genbank accession numbers KX139178–KX671728. Fastq files from ddRADseq can be found under SAMN05426771–SAMN05426803. Table A5.1 contains all information regarding Genbank IDs and sample localities.

REFERENCES

Anderson BL, Bon J, Wahono HE (1993 Reassessment of the Miocene stratigraphy, paleogeography and petroleum geochemistry of the Langsa Block in the offshore north Sumatra basin. Proceedings Indonesian Petroleum Association, 22nd Annual Convention, 169–189.

Arnold B, Corbett Detig RB, Hartl D, Bomblies K (2013) RADseq underestimates diversity and introduces genealogical biases due to nonrandom haplotype sampling. *Molecular Ecology,* 22, 3179–3190. doi:10.1111/mec.12276.

Bagley JC, Johnson JB (2014) Testing for shared biogeographic history in the lower Central American freshwater fish assemblage using comparative phylogeography: concerted, independent, or multiple evolutionary responses?. *Ecology and Evolution,* 4, 1686–1705.

Barber AJ, Crow MJ, Milsom J (2005) Sumatra: geology, resources and tectonic evolution. Geological Society of London.

Batchelor BC (1979) Discontinuously rising late Cainozoic eustatic sea–levels, with special reference to Sundaland, SE Asia. *Geologie en Mijnbouw*, 58, 1–20.

Baumann P (1982) Depositional cycles on magmatic and back arcs; an example from western Indonesia. *Revue de l'Institut Français du Petrole*, 37, 3–17.

Bernt M, Donath A, Jühling F, Externbrink F, Florentz C, Fritzsch G, Pütz J, Middendorf M, Stadler PF (2013) MITOS: Improved de novo metazoan mitochondrial genome annotation. *Molecular Phylogenetics and Evolution,* 69, 313–319.

Bell RC, Parra JL, Badjedjea G, Barej MF, Blackburn DC, Burger M, ... & Kielgast J. (2017) Idiosyncratic responses to climate-driven forest fragmentation and marine incursions in reed frogs from Central Africa and the Gulf of Guinea Islands. *Molecular Ecology*, 26, 5223–5244.

Bouckaert R, Heled J, Kühnert D, Vaughan T, Wu CH, Xie D, Suchard MA, Rambaut A, Drummond AJ (2014) BEAST 2: A Software Platform for Bayesian Evolutionary Analysis. *PLoS Computational Biology,* 10, e1003537–6.

Brandon-Jones D (1996) The Asian Colobinae Mammalia: Cercopithecidae. as indicators of quaternary climatic change. *Biological Journal of the Linnean Society,* 59, 327–350.
Burnham KP, Anderson D (2003) Model selection and multi–model inference. A Pratical informatio–theoric approch. Springer, 1229.
Catchen J, Hohenlohe PA, Bassham S, Amores A, Cresko WA (2013) Stacks: an analysis tool set for population genomics. *Molecular Ecology*, 22, 3124–3140.
Chan LM, Brown JL, Yoder AD (2011) Integrating statistical genetic and geospatial methods brings new power to phylogeography. *Molecular Phylogenetics and Evolution,* 59, 523–537.
Collins JF, Kristanto AS, Bon J, Caughey CA (1995) Sequence stratigraphic framework of Oligocene and Miocene car– bonates, North Sumatra Basin, Indonesia. In: International Symposium on Sequence Stratigraphy in S.E. Asia. Indonesian Petroleum Association, 267–279.
De Bruyn M, Stelbrink B, Morley RJ, Hall R, Carvalho GR, Cannon CH, van den Bergh, G, Meijaard E, Metcalfe I, Boitani L (2014) Borneo and Indochina are major evolutionary hotspots for Southeast Asian biodiversity. *Systematic Biology,* 63, 879–901.
Demos T, Giarla TC, Handika H, Rowe KC (2016) Local endemism and within-island diversification of shrews illustrate the importance of speciation in building Sundaland mammal diversity. *Molecular Ecology,* 25, 5158–5173.
Drummond AJ, Bouckaert RR (2015) Bayesian evolutionary analysis with BEAST. Cambridge University Press. Cambridge, UK.
Earl DA (2012) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources,* 4, 359–361.
Esselstyn JA, Timm RM, Brown RM (2009) Do geological or climatic processes drive speciation in dynamic archipelagos? The tempo and mode of diversification in Southeast Asian shrews. *Evolution,* 63, 2595–2610.
Esselstyn JA, Achmadi AS, Siler CD, Evans BJ (2013) Carving out turf in a biodiversity hotspot: multiple, previously unrecognized shrew species co-occur on Java Island, Indonesia. *Molecular Ecology,* 22, 4972–4987.
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.
Frost DR (2017) Amphibian Species of the World: an Online Reference. Version 6.0 (1/15/17). Electronic Database accessible at http://research.amnh.org/herpetology/amphibia/index.html. American Museum of Natural History, New York, USA.
Fujita MK, Maldonado J, Streicher J, High–throughput sequencing and molecular evolution of squamate mitochondrial genomes. *In Preparation*.
Gutenkunst RN, Hernandez RD, Williamson SH, Bustamante CD (2009) Inferring the joint demographic history of multiple populations from multidimensional SNP frequency data. *PLoS Genet*ics, 5, e1000695.
Hall R (2001) Cenozoic reconstructions of SE Asia and the SW Pacific: changing patterns of land and sea. Faunal and floral migrations and evolution in SE Asia–Australasia, 35–56.
Hall R (2002) Cenozoic geological and plate tectonic evolution of SE Asia and the SW Pacific: computer–based reconstructions, model and animations. *Journal of Asian Earth Sciences*, 20, 353–431.
Hall R (2009) Southeast Asia's changing palaeogeography. *Blumea,* 54, 148–161.
Hall R (2011) Australia–SE Asia collision: plate tectonics and crustal flow. *Geological Society,*

London, Special Publications, 355, 75–109.
Hall R (2012a) Late Jurassic–Cenozoic reconstructions of the Indonesian region and the Indian Ocean. *Tectonophysics*, 570, 1–41.
Hall R (2012b) A review of the Cenozoic palaeoclimate history of Southeast Asia. In D.J. Gower, K.G. Johnson, J.E. Richardson, B.R. Rosen, L. Ruber, and S.T. Williams Eds., Biotic Evolution and Environmental Change in Southeast Asia. Cambridge, UK: Cambridge University Press.
Hamidy A, Kurniati H (2015) A new species of tree frog genus Rhacophorus from Sumatra, Indonesia Amphibia, Anura. *Zootaxa*, 3947, 049–066.
Haq BU, Hardenbol J, Vail PR (1987) Chronology of fluctuating sea levels since the Triassic. *Science*, 235, 1156–1167.
Harvey MB, Pemberton AJ, Smith EN (2002) New and poorly known parachuting frogs Rhacophoridae: Rhacophorus. from Sumatra and Java. *Herpetological Monographs,* 161, 46–92.
Harvey MB, Hamidy A, Kurniawan N, Shaney K, Smith EN (2014) Three new species of Pseudocalotes Squamata: Agamidae. from southern Sumatra, Indonesia. *Zootaxa*, 38412, 211–238.
Harvey MB, O'Connell KA, Barraza G, Riyanto A, Kurniawan N (2015) Two new species of Cyrtodactylus Squamata: Gekkonidae from the Southern Bukit Barisan Range of Sumatra and an estimation of their phylogeny. *Zootaxa*, 4020, 495–23.
Harvey MB, Shaney K, Sidik I, Kurniawan N, Smith EN (2017) Endemic Dragons of Sumatra's Volcanoes: New Species of Dendragama Squamata: Agamidae and Status of Salea rosaceum Thominot. *Herpetological Monographs,* 31, 69–97.
Hickerson MJ, Stahl E, Takebayashi N (2007) msBayes: pipeline for testing comparative phylogeographic histories using hierarchical approximate Bayesian computation. *BMC Bioinformatics*, 8, 268.
Hickerson MJ, Carstens BC, Cavender–Bares J, Crandall KA, Graham CH, Johnson JB,… Yoder AD (2010) Phylogeography's past, present, and future: 10 years after. *Molecular Phylogenetics and Evolution,* 54, 291–301.
Inger RF, Voris HK (2001) The biogeographical relations of the frogs and snakes of Sundaland. *Journal of Biogeography,* 28, 863–891.
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.
Lanfear R, Calcott B, Ho SYW, Guindon S (2012) Partitionfinder: combined selection of partitioning schemes and substitution models for phylogenetic analyses. *Molecular Biology and Evolution,* 29, 1695–1701.
Leonard JA, Tex RJ, Hawkins MT, Muñoz Fuentes V, Thorington R, Maldonado JE (2015) Phylogeography of vertebrates on the Sunda Shelf: A multi-species comparison. *Journal of Biogeography,* 42, 871–879.
Li J–T, Li Y, Klaus S, Rao D–Q, Hillis DM, Zhang Y–P (2013) Diversification of rhacophorid frogs provides evidence for accelerated faunal exchange between India and Eurasia during the Oligocene. *Proceedings of the National Academy of Science,* 110, 3441–3446.
Lohman DJ, De Bruyn M, Page T, Rintelen von K, Hall R, Ng PK, Shih H–T, Carvalho GR, Rintelen von T (2011) Biogeography of the Indo–Australian archipelago. *Annual Reviews in Ecology and Evolution,* 42, 205–226.

Lourens LJ, Hilgen FJ (1997) Long-periodic variations in the Earth's obliquity and their relation to third-order eustatic cycles and Late Neogene glaciations. *Quaternary International*, 40, 43–52.

Palynological evidence for Tertiary plant dispersals in the SE Asian region in relation to plate tectonics and climate. Pages 211 - 234 in R. Hall, and J. D. Holloway, editors. Biogeography and Geological Evolution of SE Asia. Backhuys Publishers, Leiden, The Netherlands.

Meijaard E (2004) Solving mammalian riddles: a reconstruction of the Tertiary and Quaternary distribution of mammals and their palaeoenvironments in island South-East Asia. *(*Australian National University*,* Canberra*)* PhD thesis.

Malkmus R, Brühl C (2002) Amphibians Reptiles of Mount Kinabalu North Borneo. Ruggell: ARG Ganter.

Miller MA, Pfeiffer W, Schwartz T (2010) Creating the CIPRES Science Gateway for inference of large phylogenetic trees" in Proceedings of the Gateway Computing Environments Workshop (GCE), 14 Nov. 2010, New Orleans, LA pp 1 – 8.

Nater A, Nietlisbach P, Arora N, van Schaik CP, van Noordwijk MA, Willems EP, Verschoor EJ (2011) Sex–biased dispersal and volcanic activities shaped phylogeographic patterns of extant orangutans genus: Pongo. *Molecular Biology and Evolution,* 288, 2275–2288.

Nater A, Arora N, Greminger MP, van Schaik CP, Singleton I, Wich SA, Krützen M (2012) Marked population structure and recent migration in the critically endangered Sumatran orangutan Pongo abelii. *Journal of Heredity,* 1041, 2–13.

Nater A, Greminger MP, Arora N, Schaik CP, Goossens B, Singleton I, Krützen M (2015) Reconstructing the demographic history of orang-utans using Approximate Bayesian Computation. *Molecular Ecology,* 242, 310–327.

O'Connell, KA, Smith EN, Shaney KJ, Arifin U, Kurniawan N, Sidik I, Fujita MK (2017a) Coalescent species delimitation of a parachuting frog from Sumatra. Zool. Scr. *In Press.* Doi: 10.1111/zsc.12248

O'Connell KA, Streicher JW, Smith EN, Fujita MK (2017b) Geographical features are the predominant driver of molecular diversification in widely distributed North American whipsnakes. *Molecular Ecology,* 26: 5729–5751.

O'Connell KA, Smith EN, Hamidy A, Kurniawan N, Fujita MK. Within–island diversification underlies parachuting frog (*Rhacophorus*) species accumulation on the Sunda Shelf. *In Review* (a).

O'Connell KA, Smith EN, Shaney KJ, Arifin U, Kurniawan N, Sidik I, Fujita MK. Description of a new parachuting frog from Sumatra using micro CT scans. *In Review* (b).

Onn CK, Grismer LL (2010) Re–assessment of the Reinwardt's Gliding Frog, Rhacophorus reinwardtii (Schlegel 1840)(Anura: Rhacophoridae) in southern Thailand and Peninsular Malaysia with it re=description as a new species. *Zootaxa,* 2505, 40–50.

Papadopoulou A, Knowles LL (2015) Species–specific responses to island connectivity cycles: refined models for testing phylogeographic concordance across a Mediterranean Pleistocene Aggregate Island Complex. *Molecular Ecology,* 24, 4252–4268.

Peterson BK, Weber JN, Kay EH, Fisher HS, Hoekstra HE (2012) Double Digest RADseq: An Inexpensive Method for De Novo SNP Discovery and Genotyping in Model and Non–Model Species. *PLoS ONE*, 7, e37135.

Petkova D, Novembre J, Stephens M (2015) Visualizing spatial population structure with estimated effective migration surfaces. *Nature Genetics,* 48, 94–100.

Portik DM, Leaché AD, Rivera D, Barej MF, Burger M, Hirschfeld M, Rödel MO, Blackburn

DC, Fujita MK (2017) Evaluating mechanisms of diversification in a Guineo-Congolian tropical forest frog using demographic model selection. *Molecular Ecology,* 26, 5245–5263.

Prates I, Xue AT, Brown JL, Alvarado–Serrano DF, Rodrigues MT, Hickerson MJ, Carnaval AC (2016) Inferring responses to climate dynamics from historical demography in neotropical forest lizards. *Proceeding of the National Academy of Science,* 11329, 7978–7985.

Pritchard JK, Stephens M, Donnelly P (2000) Inference of Population Structure Using Multilocus Genotype Data. *Genetics,* 155, 945–959.

R Core Team., 2016. R: A language and environment for statistical computing. Foundation for statistical computing, Vienna, Austria. URL https://www.R–project.org/.

Racine JS (2012) RStudio: A Platform–Independent IDE for R and Sweave. *Journal of Applied Econometrics,* 27, 167–172.

Rambaut A, Suchard, MA, Xie D, Drummond, AJ (2014) Tracer v1.6, Available from http://beast.bio.ed.ac.uk/Tracer.

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

Silvestro D, Michalak I (2012) RaxmlGUI: a graphical front–end for RAxML. *Organismal Diversification and Evolution,* 12, 335–337.

Smith BT, McCormack JE, Cuervo AM, Hickerson MJ, Aleixo A, Cadena CD … Faircloth, BC (2014) The drivers of tropical speciation. *Nature*, 515, 406–409.

Streicher JW, Harvey MB, Sheehy CM III, Anders B, Smith EN (2012) Identification and description of the tadpole of the parachuting frog Rhacophorus catamitus from southern Sumatra, Indonesia. *Journal of Herpetology,* 46, 503–506.

Streicher JW, Devitt TJ, Goldberg CS, Malone JH, Blackmon H, Fujita MK (2014a) Diversification and asymmetrical gene flow across time and space: lineage sorting and hybridization in polytypic barking frogs. *Molecular Ecology,* 23, 3273–3291.

Streicher JW, Hamidy A, Harvey MB, Anders B, Shaney KJ, Kurniawan N, Smith EN (2014b) Mitochondrial DNA reveals a new species of parachuting frog (Rhacophoridae: *Rhacophorus*) from Sumatra. *Zootaxa*, 3878, 351–4.

van Bemmelen RW (1949) The geology of Indonesia. General geology of Indonesia and adjacent archipelagoes. Martinus Nijhoff, The Hague, The Netherlands.

Voris HK (2000) Maps of Pleistocene sea levels in Southeast Asia: shorelines, river systems and time durations. *Journal of Biogeography,* 27 1153–1167.

Weigelt P, Steinbauer MJ, Cabral JS, Kreft H (2016) Late Quaternary climate change shapes island biodiversity. *Nature* 532, 99–102.

Xue AT, Hickerson MJ (2017) Multi-DICE: R package for comparative population genomic inference under hierarchical co-demographic models of independent single-population size changes. *Molecular Ecology Resources,* DOI :10.1111/1755–0998.12686

CHAPTER 6

CONCLUSION

This dissertation explored two biological systems within two very different geological contexts. Yet, despite these differences, the processes that promoted divergence in each system were very similar. In *Masticophis*, divergence was promoted in allopatry across geographical features, with each species restricted to opposite sides of geographica features. At some features, such as the Cochise Filter Barrier, multiple species exhibit phylogeographic structure in the same geographical region. In *Rhacophorus,* divergence occurred in allopatry within the same island within several highland species, where multiple species exhibit cladogenetic breaks at the same two geographical features.

This dissertation contributes to evolutionary theory in multiple ways. Chapter 1 reinforced the role of geographical features to diversification in North America, as found in many past studies. However, this dissertation was one of the first to quantify gene flow across multiple barriers in North America. My work showed that genetic diversity was proportional to the level of migrants entering a population, and that migration in North America generally moved from east to west. My second chapter tested the effect of missing data on species delimitation and conducted population genetic analyses, and found that more missing data led to stronger support for species models. This suggests that in SNP datasets, too stringent of filtering regimes remove lineage-specific loci that are useful for population genetic or species delimitation inferences. In other words, the potential errors posed by missing data are outweighted by the benefits of including more loci, and more importantly, lineage specific loci.

In addition, Chapter 2 suggested several taxonomic changes that helped to resolve long-standing questions regarding whipsnake taxonomy. In Chapter 3, I explored island biogeography theory in Sundaland. I found that within-island diversification was responsible for a majority of *Rhacophorus* species accumulation on Sumatra and Borneo, but not on the Malay Peninsula or on Java. I also found that islands with high levels of *in situ* diversification also had high species richness and endemism. This suggests that the size of the island, rather than the colonization rate has a stronger influence on species richness, despite the prediction of traditional island biogeography theory that the distance from the mainland most strongly influenced richness. Finally, in Chapter 4 I explored the processes that promoted divergence of *Rhacophorus* on Sumatra. I found that although several species diversified *in situ* on Sumatra, this diversification occurred in allopatry across two geographical features. These congruent phylogeographic breaks correspond to two marine incursions (and underlying lowland valleys) that isolated populations into northern, central, and southern lineages.

APPENDIX 1

Table A2.1: Locality and online repository information on all samples used in this study.

Specimen Number	Species	Country	State: Locality	Latitude	Longitude	GenBank ID	SRA	Source
JAC 30654	*flagellum*	Mexico	Sonora	27.12109	-109.517115	KT713652		This study
AMNHR 139223	*flagellum*	USA	Arizona: Cochise	31.95	-109.97	KT713629		This study
AMNHR 146742	*flagellum*	USA	New Mexico: HW 9: Between Animas and Windmill	31.95	-108.73	KT713630	SRS1047343	This study
ANF I5	*flagellum*	USA	Texas: Angelia	31.145794	-94.709484		SRS1047342	This study
ANF I7	*flagellum*	USA	Texas: Angelia	31.145794	-94.709484		SRS1047341	This study
ASH 177	*flagellum*	USA	Texas: Big Bend Ranch State Park	29.49495	-103.8988833		SRS1047339	This study
ASH 188	*Salvadora deserticola*	USA	Texas: Sanderson	30.00105	-102.4428333		SRS1047338	This study
ASH 222	*flagellum*	Mexico	Coahuila: Rancho El Salado	29.33978333	-102.6546667		SRS1047337	This study
CAS 195954	*flagellum*	USA	Florida: HW 41 S Florida Ave	28.69	-82.33	KT713631	SRS1047336	This study
CAS 199521	*lateralis*	USA	California: Ishi Wilderness: Ponderosa Way	40.183372	-121.719286	KT713686	SRS1047335	This study
CAS 200366	*lateralis*	USA	California: Perris: Lukens Ln	33.800556	-117.255278	KT713687	SRS1047334	This study
CAS 208508	*lateralis*	USA	California: Callender: HW 1	35.054933	-120.594833	KT713688	SRS1047333	This study
CAS 210354	*lateralis*	USA	California: Highlands: E Highland Ave	34.135817	-117.212264	KT713689	SRS1047332	This study
CAS 214850	*flagellum*	USA	Florida: Avon Park: HW 27	27.541194	-81.48825	KT713632	SRS1047331	This study
CAS 214877	*lateralis*	USA	California: Mines Rd: North of Del Valle Regional Park	37.606969	-121.671319	KT713690	SRS1047330	This study
CAS 223420	*taeniatus*	USA	Utah: HW 21	38.927233	-114.035583	KT713729	SRS1047329	This study
CAS 227889	*taeniatus*	USA	Nevada: Snake Creek	38.911728	-114.162406	KT713730	SRS1047328	This study
CAS 227922	*taeniatus*	USA	Nevada: HW 447	40.713167	-119.468833	KT713731	SRS1047327	This study
CAS 229232	*flagellum*	USA	New Mexico: South of Conchas: HW 104	35.344667	-104.219167	KT713633	SRS1047326	This study
CAS 229237	*flagellum*	USA	New Mexico: HW 11 south of Deming	32.124	-107.7515	KT713634	SRS1047325	This study
CAS 229248	*taeniatus*	USA	Utah: south of Tooele	40.42815	-112.271689	KT713732	SRS1047323	This study
CAS 231705	*flagellum*	USA	Florida: North of Woodville	30.34	-84.25	KT713635	SRS1047324	This study
CAS 253122	*taeniatus*	USA	Idaho: Wildhorse Creek Road	42.236	-118.2322		SRS1047321	This study

CAS 253173	*taeniatus*	USA	Idaho: HW 95	42.8384	-117.6195		SRS1047322	This study
CJF 5011	*taeniatus*	USA	Texas: Menard	30.9348	-99.768208	KT713733		This study
CJF 5736	*flagellum*	USA	Texas: East of Fort McKavett: South of Hw 190	30.863233	-100.030272	KT713636	SRS1047320	This study
CJF 5816	*flagellum*	USA	Arizona: South west of Phoenix: HW 8	32.701	-113.75992		SRS1047319	This study
CLC 531	*flagellum*	USA	Texas	27.60302	-98.41345	KT713638	SRS1047318	This study
CLC 559	*schotti*	USA	Texas: Three Rivers	28.48086	-98.17325	KT713723		This study
CLC 618	*flagellum*	USA	Texas: Old Decatur Rd	33.307	-97.60629	KT713639	SRS1047317	This study
CLC 620	*schotti*	USA	Texas: Old Decatur Rd	33.307	-97.60629	KT713724	SRS1047316	This study
CLC 63	*flagellum*	USA	Texas: North of Alice: HW 3376	27.79479	-98.04662	KT713637		This study
CLC 711	*schotti*	USA	Texas	27.93954	-97.59	KT713725		This study
CLC 759	*flagellum*	USA	Texas: HW 352 and HW 147	32.6166	-99.4666	KT713640	SRS1047315	This study
CLC 849	*flagellum*	USA	Texas: Beeville: HW 351	28.38935	-97.77164	KT713641	SRS1047314	This study
CLC 889	*flagellum*	USA	Texas: South of Alice: West of HW 281	27.60302	-98.14345	KT713642		This study
CLP 137	*taeniatus*	USA	Texas: El Paso	31.92355	-105.95148	KT713734	SRS1047312	This study
CLP 138	*taeniatus*	USA	New Mexico: Gila National Park	34.1152	-108.4264	KT713735	SRS1047313	This study
DRS 011	*flagellum*	USA	Texas: North west of Kileen: Blakely Rd	31.33527	-98.09553		SRS1047308	This study
EACP152	*flagellum*	USA	Texas: Dallas Co.	32.078	-96.930922	KY007696		This study
ENS 10456	*mentovarius*	Guatemala	Huehuetango	15.2985	-91.50205		SRS1047306	This study
ENS 8558	*mentovarius*	Guatemala	Suchitepequez: HW 2 Interamerican HWY	14.538056	-90.472447	KT713698	SRS1047304	This study
FTB 1142	*flagellum*	USA	Georgia: East of Eastman: HW 46	32.22	-82.99	KT713645	SRS1047303	This study
FTB 2451	*flagellum*	USA	Georgia: South of Atlanta: HW 75	33.6	-84.28	KT713646	SRS1047302	This study
FTB 840	*flagellum*	USA	Georgia: South of Atlanta: HW 75	33.6	-84.28	KT713644		This study
JAC 21970	*mentovarius*	Mexico	Oaxaca: San Juan Lagunas	17.01	-97.93	KT713696		This study
JAC 24853 (MX 20-30)	*flagellum*	Mexico	Sonora: Hwy between Hornos and San Nicolas	27.74358	-109.76166	KY007698		This study
JAC 27528	*mentovarius*	Mexico	Guerrero: HW 51 Teloloapan- Arcelia	18.41563	-99.98695	KT713702		This study
JAC 27587	*flagellum*	Mexico	Michoachan	19.6986	-102.185	KT713648		This study
JAC 27804	*mentovarius*	Mexico	Guerrero: HW 134	17.86013	-101.388816	KT713703	SRS1047299	This study
JAC 28106	*mentovarius*	Mexico	Colima: Colima: Acatitan	19.150743	-103.71067	KT713706	SRS1047298	This study
JAC 28128	sp.	Mexico	Colima:Road from Colima to Minatitlan and HWY 98 from Minatitlan to Manzanillo	19.21555	-104.21592	KY007699		This study

JAC 28602	*mentovarius*	Mexico	Colima: Colima: Acatitan	19.150743	-103.71067	KT713707		This study
JAC 29104	*flagellum*	Mexico	Chihuahua: HW 24	26.242652	-106.518098	KY007697	SRS1047297	This study
JAC 30152	*taeniatus*	Mexico	Jalisco: Taquila	20.90493	-103.83241	KT713737		This study
JAC 30342	*mentovarius*	Mexico	Jalisco: HW 15 Guadalajara-Morelia	20.31195	-103.491635	KT713708	SRS1047294	This study
JAC 30568	*flagellum*	Mexico	Sonora: Navojoa: HW 15	27.12109	-109.517115	KT713651	SRS1047292	This study
JMM 778	*flagellum*	Mexico	Baja California: HW 5 San Felipe- Mexicali: South of Colonia la Puerta	32.308442	-115.324732	KT713653		This study
JWS 698	*flagellum*	USA	Texas: East of Pecos: HW 20	31.4	-103.45	KT713656	SRS1047287	This study
JWS 699	*flagellum*	USA	Texas: East of Pecos: HW 20	31.4	-103.45	KT713657		This study
JWS 706	*flagellum*	USA	Texas: East of Pecos: HW 20	31.4	-103.45	KT713658	SRS1047286	This study
JWS 707	*flagellum*	USA	Texas: South of llano: South west of Horseshoe Bay	30.45	-98.5	KT713659	SRS1047285	This study
JWS 709	*flagellum*	USA	Texas: East of Pecos: HW 20	31.4	-103.45		SRS1047288	This study
JWS233	*schotti*	USA	Texas: Hidalgo Co.: Old Military Road	26.25839	98.57032	KY007701		This study
KAO 002	*flagellum*	USA	Texas: South West of San Angelo: East of 163	30.997	-101.054	KT713660	SRS1047284	This study
LSUMZ H-14708	*flagellum*	USA	Louisiana: Kisatchie National Park: HW 167	31.736	-92.577	KT713647		This study
LSUMZ H-15951	*flagellum*	USA	Louisiana: Lime Kiln Rd	31.6799	-93.1779	KT713664	SRS1047283	This study
LSUMZ H-18425	*flagellum*	USA	New Mexico: HW 9 and 338	31.949675	-108.806706	KT713666		This study
LSUMZ H-19717	*flagellum*	USA	Florida: East of HW 98, West of Citrus Wildlife Management Area	28.789	-82.495	KT713667	SRS1047282	This study
LSUMZ H-20959	*flagellum*	USA	Louisiana: North west of Bogalusa, south of Dean Lee State Forest	30.8375	-89.95		SRS1047281	This study
LSUMZ H-21180	*flagellum*	USA	Louisiana: East of Kepler Creek Lake on Piney Woods Rd	32.335	-93.1233	KT713668	SRS1047280	This study
LSUMZ H-21205	*flagellum*	USA	Louisiana: West of Bienville: on Bp 699	32.408	-93.022	KT713669	SRS1047279	This study
LSUMZ H-2451	*flagellum*	USA	Florida: Munson Recreation Area: HW 4	30.8559	-86.853	KT713661	SRS1047277	This study
LSUMZ H-8153	*flagellum*	USA	Alabama: East of Oyster Bay	30.261	-87.709	KT713662		This study
LSUMZ H18262	*flagellum*	USA	Louisiana: Sabine Parish: East of Sabine National Forest	31.42	-93.582	KT713665		This study
LSUMZ H21274	*flagellum*	USA	Louisiana: Kepler Creek Lake	32.32	-93.116	KT713670	SRS1047278	This study
MAFL SL5	*flagellum*	USA	Louisiana: Bienville Parish	321.25	-93.265		SRS1047276	This study
MEX 23720	sp.	Mexico	Jalisco: Ambrosio: Close to HW 80	20.29284	-103.99	KT713693	SRS1047267	This study
MEX 23884	*mentovarius*	Mexico	Jalisco: HW 80 La Huerta - Casimiro Castillo	19.58153	-104.48836	KT713710	SRS1047264	This study
MEX 23955	*mentovarius*	Mexico	Colima: HW 110 Colima- Tecoman	19.12926	-103.77389	KT713711	SRS1047263	This study
MEX 24405	*mentovarius*	Mexico	Yucatan: HW 295 Tizimin- Rio Lagartos	21.57226	-88.16544	KT713713	SRS1047257	This study

MEX 24527	*mentovarius*	Mexico	Veracruz: Abasolo del Valle	17.777	-95.49	KT713714	SRS1047256	This study
MVZ 161425	*flagellum*	Mexico	Baja California: HW 1 Lazaro Cardenas- Guerrero Negro	28.515833	-114.03	KT713671	SRS1047273	This study
MVZ 182251	*lateralis*	Mexico	Baja Sur: San Juan: HW 1 Guerrero Negro- Sta Rosalia	27.325556	-113.016111	KT713691	SRS1047272	This study
MVZ 204113	*mentovarius*	Costa Rica	Alajuela: Montenegro: HW 1 Panamerican HWY	10.48333	-85.21667		SRS1047247	This study
MVZ 233302	*schotti*	Mexico	Quetetaro: El Paraiso	20.576	-100.1868	KT713727	SRS1047246	This study
MX 24271	*mentovarius*	Mexico	Oaxaca: HW 200 Acapulco - Salina Cruz	15.72556	-96.66186	KT713712	SRS1047259	This study
ROM 14197	*flagellum*	MX	Sonora: HW 8 Puerto Penasco- Sonoyta	31.39	-113.5	KT713674	SRS1047245	This study
ROM 14948	*flagellum*	MX	Chihuahua: North of El Bachivo: South of HW 13 Carr Navojoa - Alamos	27.07	-109.32	KT713675	SRS1047244	This study
ROM 14965	*bilineatus*	Mexico	Sinaloa: North west of Fresnilla: West of HW 45	23.37	-103.36	KT713628		This study
ROM 15050	*flagellum*	MX	Sonora: West of Hermosillo: HW 100	28.8	-111.91	KT713676	SRS1047243	This study
ROM 15326	*taeniatus*	MX	Durango: HW 49	26.2	-103.85	KT713728		This study
TCC 37RL16	*constrictor*	USA	Texas: South of Weatherford: HW 20	32.71	-97.878		SRS1047242	This study
TCWC 95170	*flagellum*	USA	Texas: West of College Station	30.7	-96.2011		SRS1047240	This study
UTA 22057	*mentovarius*	Mexico	Guerrero: HW 200 Coyuca de Benitez- Acapulco	16.96	-99.97	KT713716	SRS1047241	This study
UTA 22140	*mentovarius*	Mexico	Guerrero: Iguala	18.375	-99.52	KT713717	SRS1047271	This study
UTA 22543	*mentovarius*	Mexico	Oaxaca: Santiago Niltepec	16.57	-94.6	KT713718	SRS1047270	This study
UTA 23306	*bilineatus*	Mexico	Jalisco: HW 23 Tlaltenagno de Sanchez Roman-Colotlan	22.03587	-103.26819	KT713726	SRS1047269	This study
UTA 23619	sp.	Mexico	Jalisco: Ambrosio: Close to HW 80	20.29284	-103.99	KT713694	SRS1047268	This study
UTA 23750	*mentovarius*	Mexico	Jalisco: Ambrosio: Close to HW 80	20.24336	-103.99	KT713719	SRS1047266	This study
UTA 24067	*mentovarius*	Mexico	Colima: HW Zihuatanejo- Manzanillo	18.72513	-103.72293	KT713720		This study
UTA 24067	*flagellum*	Mexico	Colima: North of La Boca de Apiza: HW 200	18.72513	-103.72293		SRS1047262	This study
UTA 24074	*mentovarius*	Mexico	Michoachan: HW 120 Tepalcatepec- Apatzingan	19.15038	-102.47448		SRS1047261	This study
UTA 24191	*mentovarius*	Mexico	Guerrero: HW 200 Marquelia	16.60751	-98.7306		SRS1047254	This study
UTA 24216	*mentovarius*	Mexico	Oaxaca: HW 200 Acapulco - Salina Cruz	16.40998	-98.29134	KT713721	SRS1047260	This study
UTA 24305	*mentovarius*	Mexico	Oaxaca: HW 200 Acapulco - Salina Cruz	15.73917	-96.80701	KT713722	SRS1047258	This study
UTA 29887 (ENS 2669)	*mentovarius*	Guatemala	Escuintla: HW 14	14.333	-90.841939	KT713697	SRS1047305	This study
UTA 43612 (MSM 484)	*mentovarius*	Honduras	Comayagua: Lo de reina, Ajuterique, Comayagua	14.4236	-87.72416	KT713715	SRS1047274	This study
UTA 46670 (MEA 1244)	*mentovarius*	Guatemala	Huehuetenango: Nenton	15.794	-91.745	KT713709	SRS1047275	This study
UTA 46723 (JAC 20794)	*mentovarius*	Guatemala	Zacapa: Zacapa North East of the city	14.978242	-89.515497		SRS1047301	This study

UTA 520101 (ENS 10132)	*mentovarius*	Guatemala	Zacapa: HW 3: Communidad Santiago	15.075086	-89.438964	KT713699	SRS1047307	This study
UTA 52286 (ENS 10465)	*mentovarius*	Guatemala	Huehuetenango: Nenton	15.794533	-91.745436	KT713700		This study
UTA 53441 (JAC 23753)	sp.	Mexico	Nayarit: Valle Dorado: Blvrd Niviera Nayarit	20.714	-105.27647	KT713695	SRS1047265	This study
UTA 57591 (JAC 27879)	*mentovarius*	Mexico	Guerrero: Zihuatanejo	17.625543	-101.469369	KY007700	SRS1047300	This study
UTA 57592 (JAC 27880)	*mentovarius*	Mexico	Guerrero: Zihuatanejo	17.625543	-101.469369	KT713704		This study
UTA 57600 (JAC 26985)	*mentovarius*	Mexico	Michoachan: HW 120 Apatzingan- Patzcuaro	18.99167	-102.147	KT713701		This study
UTA 57601 (JAC 27951)	*mentovarius*	Mexico	Michoachan: San Juan de Alima	18.57893	-103.601759	KT713705		This study
UTA 57751 (JAC 30567)	*flagellum*	Mexico	Sonora: Navojoa: HW 15	27.12109	-109.517115	KT713650	SRS1047293	This study
UTA 57967 (JAC 29377)	sp.	Mexico	Durango: East of la Reserva de la Biosfera de La Michilia: HW 241	23.422366	-104.175217	KT713692	SRS1047296	This study
UTA 57968 (JAC 29855)	*flagellum*	Mexico	Tamaulipas: HW 80: North of Cuauhtemoc	22.587148	-98.21237	KT713649	SRS1047295	This study
UTA 57969 (JAC 30222)	*bilineatus*	Mexico	Nayarit: Road W of Mesquites	20.92790	-104.54950	KY007695		This study
UTA 58495 (JWS 233)	*schotti*	USA	Texas: Cuevitas: Military Rd	26.25839	-98.57032		SRS1047290	This study
UTA 58701 (JWS 261)	*flagellum*	USA	Texas: North of Abilene on HW 351	32.6547	-99.45863	KT713655	SRS1047289	This study
UTA 58992 (JWS 025)	*taeniatus*	USA	Texas: Fort Davis	30.6638	-104.01667	KT713738		This study
UTA 59001 (JWS 041)	*flagellum*	USA	Texas: Fort Davis: HW 17	30.605	-103.87533	KT713654	SRS1047291	This study
UTA 60490	*flagellum*	USA	Texas			KT713685		This study
UTA 62912 (JAC 29106)	*taeniatus*	Mexico	Jalisco: Vaquerias	21.76642	-101.62672	KT713736		This study
UTAR 26244	*flagellum*	USA	Texas: North of Decatur	33.391	-97.55	KT713677	SRS1047195	This study
UTAR 55431	*flagellum*	USA	Texas: South of Lake Whitney State Park	31.79	-97.42	KT713679		This study
UTAR 60140	*flagellum*	USA	Texas: Crescent Heights: West of 753	32.169	-95.947	KT713673	SRS1047255	This study
UTAR 60458	*flagellum*	USA	Texas: East of Laredo: West of 16	27.253	-98.94	KT713683	SRS1047253	This study
UTAT 40786	*flagellum*	Mexico	Tamaulipas: HW 97 East of Monterrey	25.3	-98.305655	KT713678	SRS1047252	This study
UTAT 55873	*flagellum*	USA	Texas: South of Antelope: Prideaux Rd	33.40475	-98.40544	KT713684	SRS1047251	This study
UTAT 59083	*flagellum*	Mexico	Tamaulipas: HW 97 East of Monterrey	25.3	-98.305655	KT713680	SRS1047250	This study
UTAT 60441	*flagellum*	USA	Texas: North of Comanche: HW 16	31.93427	-98.575658	KT713681	SRS1047249	This study
UTAT 60442	*flagellum*	USA	Texas: West of Kileen: South of Colorado Bend State Park	30.987421	-98.575658	KT713682	SRS1047248	This study
UTEP 18542 (CSL 9506)	*flagellum*	USA	New Mexico: Lincoln National Forest: Cloudcroft	32.96032	-105.685018		SRS1047310	This study
UTEP 20749 (CSL 9485)	*flagellum*	USA	Texas: North of Fort Davis: HW 118	30.726954	-104.131703		SRS1047311	This study
UTEP 20772 (CSL 9511)	*flagellum*	USA	Texas: North of Fort Davis: HW 118	30.726954	-104.131703		SRS1047309	This study

AMNH 502345	*flagellum*	USA	Texas	32.76217	-99.0825	KX835748	Myers et al., 2016
AMNH 502347	*flagellum*	USA	Texas: west Texas	32.71121	-102.58082	KX835749	Myers et al., 2016
AMNH 502349	*flagellum*	USA	New Mexico: Near Malaga	32.1014	-104.07269	KX835750	Myers et al., 2016
AMNH 502351	*flagellum*	USA	Texas: near Pecos	31.6463	-103.66731	KX835751	Myers et al., 2016
AMNH 502352	*flagellum*	USA	New Mexico: Near Carlsbad	32.18395	-103.40491	KX835752	Myers et al., 2016
AMNH 502358	*flagellum*	USA	New Mexico: Near Lake Arthur	32.95682	-104.38033	KX835753	Myers et al., 2016
AMNH 502359	*flagellum*	USA	New Mexico: Near Carlsbad	32.19728	-104.33687	KX835754	Myers et al., 2016
AMNH 502363	*flagellum*	USA	New Mexico	33.5808	-105.96581	KX835755	Myers et al., 2016
AMNH 502370	*flagellum*	USA	New Mexico: Near Gila National Forest	32.60183	-107.32455	KX835756	Myers et al., 2016
AMNH 502378	*flagellum*	USA	New Mexico: Hachita	31.875	-108.33517	KX835757	Myers et al., 2016
AMNH 502409	*flagellum*	USA	Arizona: Sonora Desert National Monument	33.059965	-112.254944	KX835758	Myers et al., 2016
AMNH 502410	*flagellum*	USA	Arizona: Sonora Desert National Monument	33.043203	-112.322021	KX835759	Myers et al., 2016
AMNH 502504	*flagellum*	USA	Arizona: far west	33.91423	-114.02694	KX835760	Myers et al., 2016
AMNH 502420	*flagellum*	USA	Arizona: Near Wickenburg	33.93536	-112.6926	KX835761	Myers et al., 2016
AMNH 502423	*flagellum*	USA	Arizona: Dudleyville	32.93947	-110.73831	KX835762	Myers et al., 2016
AMNH 502426	*flagellum*	USA	Arizona: Hw 79	32.74914	-111.13174	KX835763	Myers et al., 2016
AMNH 502445	*flagellum*	USA	Arizona: Apache	31.65849	-109.1551	KX835764	Myers et al., 2016
AMNH 502450	*flagellum*	USA	New Mexico: Columbus	31.82879	-107.6392	KX835765	Myers et al., 2016

AMNH 502451	*flagellum*	USA	New Mexico: Alamogordo, near wt sands nat. monument	32.797448	-106.134896	KX835766	Myers et al., 2016
AMNH 500882	*flagellum*	USA	Arizona: W of Phoenix	33.2577000000	-113.1387500000	KX835773	Myers et al., 2016
AMNH 500883	*flagellum*	USA	Arizona: Near Chiricahua national monument	31.936722	-109.135809	KX835774	Myers et al., 2016
AMNH 500884	*flagellum*	USA	Arizona: Near Chiricahua national monument	31.9604500000	-109.1441800000	KX835771	Myers et al., 2016
AMNH 500885	*flagellum*	USA	New Mexico, South of Albuquerque	34.3376300000	-106.8786900000	KX835775	Myers et al., 2016
AMNH 500886	*flagellum*	USA	New Mexico, South of Albuquerque	34.6502400000	-106.8120500000	KX835776	Myers et al., 2016
AMNH 500887	*flagellum*	USA	New Mexico, South of Albuquerque	34.4195700000	-106.7651000000	KX835777	Myers et al., 2016
AMNH 500889	*flagellum*	USA	New Mexico: South of Deming	32.0477400000	-107.7097300000	KX835769	Myers et al., 2016
AMNH 500890	*flagellum*	USA	New Mexico: Columbus	31.8227600000	-107.6561700000	KX835770	Myers et al., 2016
AMNH 500891	*flagellum*	USA	New Mexico: near Apache	31.5972400000	-109.2357400000	KX835772	Myers et al., 2016
UF 150082	*flagellum*	USA	Florida	30.568611	-85.812222	KT447218	Myers et al., 2016
MVZ 245881	*flagellum*	USA	California: Kern County	35.1245	-118.16728	KX835787	Myers et al., 2016
MVZ 234614	*flagellum*	USA	California: San Bernardino County	34.90243	-115.74564	KX835786	Myers et al., 2016
MVZ 229145	*flagellum*	USA	California: Kern County	35.611423	-118.2426654	KX835785	Myers et al., 2016
CAS 223614	*flagellum*	USA	California: San Diego County	33.1532475	-116.1584318	KX835783	Myers et al., 2016
CAS 200662	*flagellum*	USA	California: Riverside County	33.675	-117.0115	KX835782	Myers et al., 2016
CAS 200381	*flagellum*	USA	California: San Diego County	33.0083	-116.8678167	KX835781	Myers et al., 2016
CAS 200375	*flagellum*	USA	California: Riverside County	33.80055556	-117.2552778	KX835780	Myers et al., 2016
JAC 30652	*flagellum*	MX	Sonora	27.12109	-109.517115	KX835768	Myers et al., 2016
CAS 219734	*flagellum*	USA	California: Kern County	35.6485	-118.3580278	AY486928	Myers et al., 2016
AMNH 502362	*flagellum*	USA	New Mexico	33.5808	-105.96581	KX835755	Myers et al., 2016
CAS 218707	*Coluber constrictor*	USA	Florida	29.03526111	-82.46094	EU180430	Burbrink et al., 2008

CAS 218699	*Coluber constrictor*	USA	Florida	30.16586111	-82.2003333	EU180433	Burbrink et al., 2009
CAS 219499	*Coluber constrictor*	USA	California	39.60305556	-122.9013611	EU180466	Burbrink et al., 2010
CAS 200845	*Tantilla relicta*	USA	Florida	28.69115	-82.33772	AF471045	Lawson et al., 2005
CAS 206503	*Sonora semiannulata*	USA	California	36.2452	-117.45315	AF471048	Lawson et al., 2005
CAS 212760	*Salvadora mexicana*	USA	California	39.16058333	-122.6680833	AY486914	Nagy et al., 2004
no voucher	*Spilotes pullatus*					AF471041	Lawson et al., 2005
no voucher	*Phyllorhynchus decurtatus*					AF471083	Lawson et al., 2005
CAS 198327	*Drymarchon corais*					AF471064	Lawson et al., 2005
CAS 175557	*Oxybelis aeneus*					AF471056	Lawson et al., 2005
CAS 172661	*Opheodrys aestivus*					AF471057	Lawson et al., 2005

Table A2.2: Information on number of loci used in each analysis

Species	Analysis	SNPS	%Missing per Locus		n
M. flagellum	STRUCTURE/EEMS	2504		50	36
M. flagellum	STRUCTURE/EEMS	499		20	36
M. flagellum	STRUCTURE/EEMS	80		10	36
M. mentovarius	STRUCTURE/EEMS	2169		50	24
M. mentovarius	STRUCTURE/EEMS	1000		20	24
M. mentovarius	STRUCTURE/EEMS	629		10	24
M. lateralis	STRUCTURE	1555		20	5
M. taeniatus	STRUCTURE	958		20	4
M. flagellum West to East	Migrate-n	2313		46	13
M. flagellum West to Chihuahua	Migrate-n	3006		43	14
M. flagellum Sonora to Chihuahua	Migrate-n	1413		0	7
M. mentovarius	Migrate-n	1680		16	48
M. taeniatus	Migrate-n	1622		13	8

Table A2.3: Complete results of Migrate-n analysis

Parameter	*flagellum* West-East	
NeW		0.00679
NeE		0.00589
M_E_W		632.4
M_W_E		597.8
Migr/Gen_E_W		1.073499
Migr/Gen_W_E		0.8802605
Loci		1430
pyRADClustering TX	6 indiv	
pyRADClustering FL	6 indiv	

Parameter	*flagellum* Chihuahua-West	
NeW		0.00657
NeCh		0.00677
M_Ch_W		535
M_W_Ch		706.6
Migr/Gen_Ch_W		0.8787375
Migr/Gen_W_Ch		1.1959205
Loci		715
pyRADClustering TX	6 individuals	
pyRADClustering NM	3 individuals	

Parameter	*flagellum* Sonora-Chihuahua	
NeS		0.00608
NeCh		0.0057
M_S_Ch		561.6
M_Ch_S		524.2
Migr/Gen_S_Ch		0.80028

Migr/Gen_Ch_S		0.796784
Loci		360
pyRADClustering S	4 individuals	
pyRADClustering Ch	3 individuals	

Parameter	*taeniatus* North-South	
NeW		0.00591
NeE		0.00683
MSW		556.8
MNE		612.5
Migr/Gen_E_W		0.822672
Migr/Gen_W_E		1.04584375
Loci		1622
pyRADClustering N	1 individual	
pyRADClustering S	1 individual	

Parameter	*mentovarius* North-South	
NeW		0.0073
NeE		0.0053
M_E_W		656
M_W_E		492.8
Migr/Gen_E_W		1.1972
Migr/Gen_W_E		0.65296
Loci		1680
pyRADClustering N	8 individuals	
pyRADClustering S	8 individuals	

Table A2.4: Primers used for Mitochondrial Analysis

Primer Name	Sequence
CA_LATERALIS F1	5'-GAGGTATGGGTGAATGGTAT 3'
CA_LATERALIS R1	5'-GGCACAACATTAACTACCTG 3'
GUAT-F1	5'-GTAATGAATGTAGCGATTAGGG 3'
GUAT-R1	5'-CTTCTTCCTAGCAATCCACTA 3'
MX GUE_OAX F1	5'-CGATGAGGGTTCAAAATACTAG 3'
MX GUE_OAX R1	5'-GTTCCATACGGATGAATCATAC 3'
MX_OAX F1	5'-GAAGGCTATGGATCGGATGT 3'
MX_OAX R1	5'-GATGTTCCATACGGATGAATCA 3'
MX_JAL F1	5'-CGATGAGGGTTCAAATACTAG 3'
MX_JAL R1	5'-GTTCCATACGGATGAATCATAC 3'
TX_FLAG F1	5'-TTTGTATGAATGGTCGGAAGG 3'
TX_FLAG R1	5'-CCTAGCCTTCTCATCTATTGTT 3'

Table A3.1: Locality and online repository information on all samples used in this study.

Specimen Number	**Species**	**Latitude**	**Longitude**	**GenBank ID**	**SRA**	**Source**
UTA 57969 (JAC 30222)	*bilineatus*	20.92790	-104.54950	KY007695		O'Connell et al. 2017
ROM 14965	*bilineatus*	23.37	-103.36	KT713628		O'Connell et al. 2017
UTA 23306	*bilineatus*	22.03587	-103.26819	KT713726	SRS1047269	O'Connell et al. 2017
MVZ 225550	*bilineatus*	31.9136	-109.144	KP765657		

JAC 30654	*flagellum*	27.12109	-109.517115	KT713652		O'Connell et al. 2017
AMNHR 139223	*flagellum*	31.95	-109.97	KT713629		O'Connell et al. 2017
AMNHR 146742	*flagellum*	31.95	-108.73	KT713630		O'Connell et al. 2017
CAS 195954	*flagellum*	28.69	-82.33	KT713631		O'Connell et al. 2017
CAS 214850	*flagellum*	27.541194	-81.48825	KT713632	SRS1047331	O'Connell et al. 2017
CAS 229232	*flagellum*	35.344667	-104.219167	KT713633	SRS1047326	O'Connell et al. 2017
CAS 229237	*flagellum*	32.124	-107.7515	KT713634	SRS1047325	O'Connell et al. 2017
CAS 231705	*flagellum*	30.34	-84.25	KT713635	SRS1047324	O'Connell et al. 2017
CJF 5736	*flagellum*	30.863233	-100.030272	KT713636		O'Connell et al. 2017
UTA 58400 (CLC 63)	*flagellum*	27.79479	-98.04662	KT713637		O'Connell et al. 2017
UTA 58885 (CLC 531)	*flagellum*	27.60302	-98.41345	KT713638		O'Connell et al. 2017
CLC 618	*flagellum*	33.307	-97.60629	KT713639	SRS1047317	O'Connell et al. 2017
CLC 759	*flagellum*	32.6166	-99.4666	KT713640		O'Connell et al. 2017
CLC 849	*flagellum*	28.38935	-97.77164	KT713641		O'Connell et al. 2017
CLC 889	*flagellum*	27.60302	-98.14345	KT713642		O'Connell et al. 2017
FTB 840	*flagellum*	33.6	-84.28	KT713644		O'Connell et al. 2017
FTB 1142	*flagellum*	32.22	-82.99	KT713645	SRS1047303	O'Connell et al. 2017
FTB 2451	*flagellum*	33.6	-84.28	KT713646		O'Connell et al. 2017
LSUMZ H-14708	*flagellum*	31.736	-92.577	KT713647		O'Connell et al. 2017
JAC 27587	*flagellum*	19.6986	-102.185	KT713648		O'Connell et al. 2017
UTA 57968 (JAC 29855)	*flagellum*	22.587148	-98.21237	KT713649	SRS1047295	O'Connell et al. 2017
UTA 57751 (JAC 30567)	*flagellum*	27.12109	-109.517115	KT713650	SRS1047293	O'Connell et al. 2017
UTA 57752 (JAC 30568)	*flagellum*	27.12109	-109.517115	KT713651	SRS1047292	O'Connell et al. 2017
JMM 778	*flagellum*	32.308442	-115.324732	KT713653		O'Connell et al. 2017
UTA 59001 (JWS 041)	*flagellum*	30.605	-103.87533	KT713654	SRS1047291	O'Connell et al. 2017
UTA 58701 (JWS 261)	*flagellum*	32.6547	-99.45863	KT713655		O'Connell et al. 2017
JWS 698	*flagellum*	31.4	-103.45	KT713656		O'Connell et al. 2017
JWS 699	*flagellum*	31.4	-103.45	KT713657		O'Connell et al. 2017
JWS 706	*flagellum*	31.4	-103.45	KT713658		O'Connell et al. 2017

JWS 707	*flagellum*	30.45	-98.5	KT713659		O'Connell et al. 2017
KAO 002	*flagellum*	30.997	-101.054	KT713660		O'Connell et al. 2017
LSUMZ H-2451	*flagellum*	30.8559	-86.853	KT713661	SRS1047277	O'Connell et al. 2017
LSUMZ H-8153	*flagellum*	30.261	-87.709	KT713662		O'Connell et al. 2017
LSUMZ H-15951	*flagellum*	31.6799	-93.1779	KT713664		O'Connell et al. 2017
LSUMZ H18262	*flagellum*	31.42	-93.582	KT713665		O'Connell et al. 2017
LSUMZ H-18425	*flagellum*	31.949675	-108.806706	KT713666		O'Connell et al. 2017
LSUMZ H-19717	*flagellum*	28.789	-82.495	KT713667		O'Connell et al. 2017
LSUMZ H-21180	*flagellum*	32.335	-93.1233	KT713668	SRS1047280	O'Connell et al. 2017
LSUMZ H-21205	*flagellum*	32.408	-93.022	KT713669		O'Connell et al. 2017
LSUMZ H21274	*flagellum*	32.32	-93.116	KT713670	SRS1047278	O'Connell et al. 2017
MVZ 161425	*flagellum*	28.515833	-114.03	KT713671	SRS1047273	O'Connell et al. 2017
UTAR 60140	*flagellum*	32.169	-95.947	KT713673		O'Connell et al. 2017
ROM 14197	*flagellum*	31.39	-113.5	KT713674		O'Connell et al. 2017
ROM 14948	*flagellum*	27.07	-109.32	KT713675		O'Connell et al. 2017
ROM 15050	*flagellum*	28.8	-111.91	KT713676	SRS1047243	O'Connell et al. 2017
UTAR 26244	*flagellum*	33.391	-97.55	KT713677		O'Connell et al. 2017
UTAT 40786	*flagellum*	25.3	-98.305655	KT713678		O'Connell et al. 2017
UTAR 55431	*flagellum*	31.79	-97.42	KT713679		O'Connell et al. 2017
UTAT 59083	*flagellum*	25.3	-98.305655	KT713680		O'Connell et al. 2017
UTAT 60441	*flagellum*	31.93427	-98.575658	KT713681	SRS1047249	O'Connell et al. 2017
UTAT 60442	*flagellum*	30.987421	-98.575658	KT713682		O'Connell et al. 2017
UTAR 60458	*flagellum*	27.253	-98.94	KT713683		O'Connell et al. 2017
UTAT 55873	*flagellum*	33.40475	-98.40544	KT713684		O'Connell et al. 2017
UTAR 60490	*flagellum*			KT713685		O'Connell et al. 2017
ANF I5	*flagellum*	31.145794	-94.709484			O'Connell et al. 2017
ANF I7	*flagellum*	31.145794	-94.709484			O'Connell et al. 2017
ASH 177	*flagellum*	29.49495	-103.8988833			O'Connell et al. 2017
ASH 222	*flagellum*	29.33978333	-102.6546667			O'Connell et al. 2017

CJF 5816	*flagellum*	32.701	-113.75992			O'Connell et al. 2017
UTEP 20749 (CSL 9485)	*flagellum*	30.726954	-104.131703			O'Connell et al. 2017
UTEP 18542 (CSL 9506)	*flagellum*	32.96032	-105.685018		SRS1047310	O'Connell et al. 2017
UTEP 20772 (CSL 9511)	*flagellum*	30.726954	-104.131703			O'Connell et al. 2017
DRS 011	*flagellum*	31.33527	-98.09553			O'Connell et al. 2017
JAC 29104	*flagellum*	26.242652	-106.518098	KY007697	SRS1047297	O'Connell et al. 2017
JWS 709	*flagellum*	31.4	-103.45		SRS1047288	O'Connell et al. 2017
LSUMZ H-20959	*flagellum*	30.8375	-89.95		SRS1047281	O'Connell et al. 2017
MAFL SL5	*flagellum*	321.25	-93.265			O'Connell et al. 2017
UTA 53316 (24067)	*flagellum*	18.72513	-103.72293			O'Connell et al. 2017
TCWC 95170	*flagellum*	30.7	-96.2011			O'Connell et al. 2017
JAC 24853 (MX 20-30)	*flagellum*	27.74358	-109.76166	KY007698		Myers et al., 2016
EACP152	*flagellum*	32.078	-96.930922	KY007696		Myers et al., 2016
AMNH 502345	*flagellum*	32.76217	-99.0825	KX835748		Myers et al., 2016
AMNH 502347	*flagellum*	32.71121	-102.58082	KX835749		Myers et al., 2016
AMNH 502349	*flagellum*	32.1014	-104.07269	KX835750		Myers et al., 2016
AMNH 502351	*flagellum*	31.6463	-103.66731	KX835751		Myers et al., 2016
AMNH 502352	*flagellum*	32.18395	-103.40491	KX835752		Myers et al., 2016
AMNH 502358	*flagellum*	32.95682	-104.38033	KX835753		Myers et al., 2016
AMNH 502359	*flagellum*	32.19728	-104.33687	KX835754		Myers et al., 2016
AMNH 502363	*flagellum*	33.5808	-105.96581	KX835755		Myers et al., 2016
AMNH 502370	*flagellum*	32.60183	-107.32455	KX835756		Myers et al., 2016
AMNH 502378	*flagellum*	31.875	-108.33517	KX835757		Myers et al., 2016

AMNH 502409	*flagellum*	33.059965	-112.254944	KX835758	Myers et al., 2016
AMNH 502410	*flagellum*	33.043203	-112.322021	KX835759	Myers et al., 2016
AMNH 502504	*flagellum*	33.91423	-114.02694	KX835760	Myers et al., 2016
AMNH 502420	*flagellum*	33.93536	-112.6926	KX835761	Myers et al., 2016
AMNH 502423	*flagellum*	32.93947	-110.73831	KX835762	Myers et al., 2016
AMNH 502426	*flagellum*	32.74914	-111.13174	KX835763	Myers et al., 2016
AMNH 502445	*flagellum*	31.65849	-109.1551	KX835764	Myers et al., 2016
AMNH 502450	*flagellum*	31.82879	-107.6392	KX835765	Myers et al., 2016
AMNH 502451	*flagellum*	32.797448	-106.134896	KX835766	Myers et al., 2016
AMNH 500882	*flagellum*	33.2577000000	-113.1387500000	KX835773	Myers et al., 2016
AMNH 500883	*flagellum*	31.936722	-109.135809	KX835774	Myers et al., 2016
AMNH 500884	*flagellum*	31.9604500000	-109.1441800000	KX835771	Myers et al., 2016
AMNH 500885	*flagellum*	34.3376300000	-106.8786900000	KX835775	Myers et al., 2016
AMNH 500886	*flagellum*	34.6502400000	-106.8120500000	KX835776	Myers et al., 2016
AMNH 500887	*flagellum*	34.4195700000	-106.7651000000	KX835777	Myers et al., 2016
AMNH 500889	*flagellum*	32.0477400000	-107.7097300000	KX835769	Myers et al., 2016
AMNH 500890	*flagellum*	31.8227600000	-107.6561700000	KX835770	Myers et al., 2016
AMNH 500891	*flagellum*	31.5972400000	-109.2357400000	KX835772	Myers et al., 2016

UF 150082	*flagellum*	30.568611	-85.812222	KT447218		Myers et al., 2016
MVZ 245881	*flagellum*	35.1245	-118.16728	KX835787		Myers et al., 2016
MVZ 234614	*flagellum*	34.90243	-115.74564	KX835786		Myers et al., 2016
MVZ 229145	*flagellum*	35.611423	-118.2426654	KX835785		Myers et al., 2016
CAS 223614	*flagellum*	33.1532475	-116.1584318	KX835783		Myers et al., 2016
CAS 200662	*flagellum*	33.675	-117.0115	KX835782		Myers et al., 2016
CAS 200381	*flagellum*	33.0083	-116.8678167	KX835781		Myers et al., 2016
CAS 200375	*flagellum*	33.80055556	-117.2552778	KX835780		Myers et al., 2016
JAC 30652	*flagellum*	27.12109	-109.517115	KX835768		Myers et al., 2016
CAS 219734	*flagellum*	35.6485	-118.3580278	AY486928		Myers et al., 2016
AMNH 502362	*flagellum*	33.5808	-105.96581	KX835755		Myers et al., 2016
UTA 29887 (ENS 2669)	*mentovarius*	14.333	-90.841939	KT713697	SRS1047305	O'Connell et al. 2017
MVZ 204113	*mentovarius*	10.48333	-85.21667		SRS1047247	O'Connell et al. 2017
UTA 23750	*mentovarius*	20.24336	-103.99	KT713719	SRS1047266	O'Connell et al. 2017
UTA 52010 (ENS 10132)	*mentovarius*	15.075086	-89.438964	KT713699	SRS1047307	O'Connell et al. 2017
MX 24271	*mentovarius*	15.72556	-96.66186	KT713712	SRS1047259	O'Connell et al. 2017
UTAR 53341 (JAC 24305)	*mentovarius*	15.73917	-96.80701	KT713722	SRS1047258	O'Connell et al. 2017
MEX 23720	sp.	20.29284	-103.99	KT713693		O'Connell et al. 2017
JAC 28128	sp.	19.21555	-104.21592	KY007699		Myers et al., 2016
UTA 57967 (JAC 29377)	sp.	23.422366	-104.175217	KT713692	SRS1047296	O'Connell et al. 2017
UTA 23619	sp.	20.29284	-103.99	KT713694	SRS1047268	O'Connell et al. 2017
UTA 53441 (JAC 23753)	sp.	20.714	-105.27647	KT713695	SRS1047265	O'Connell et al. 2017
CAS 218707	Coluber constrictor	29.03526111	-82.46094	EU180430		Burbrink et al., 2008
CAS 218699	Coluber constrictor	30.16586111	-82.2003333	EU180433		Burbrink et al., 2009
CAS 219499	Coluber constrictor	39.60305556	-122.9013611	EU180466		Burbrink et al., 2010
CAS 200845	Tantilla relicta	28.69115	-82.33772	AF471045		Lawson et al., 2005
CAS 206503	Sonora semiannulata	36.2452	-117.45315	AF471048		Lawson et al., 2006
CAS 212760	Salvadora mexicana	39.16058333	-122.6680833	AY486914		Nagy et al., 2004

Table A3.2: The number of loci used in each species delimitation analysis used in Chapter 3.

Dataset	n	%missing at locus	mean % missing individual	# Loci	Analysis Used
A	15	30	20	365	SPLITSTREE
B	26	50	35.3	2077	SNAPP
C	26	20	13.3	325	SNAPP
D	10	50	38.6	1464	SNAPP
E	10	20	16.2	216	SNAPP

Table A4.1 Locality and molecular sequence information.

Museum Number	*Species*	Latitude	Longitude	16S	12S	Cox 1	Cytb	BDNF	POMC	RAG1	Rhodopsin	Tyrosinase
ZSM 203/2004	*Aglyptodactylus madagascariensis*					JN133054	JN132845					
	Aglyptodactylus madagascariensis										AY880664	AF249166
ZCMV8046	*Aglyptodactylus madagascariensis*			KT159891						DQ347233		
ZSM 405/2000	*Blommersia wittei*			AB612030				EF396018		AY323774		
	Blommersia wittei					GU983132	AY723696				AY880667	AY341751
ZSM 29/2006	*Boophis doulioti*					JN133081	JN132870			AY571643	AY341792	
ZCMV 14166	*Boophis doulioti*			KR025902								
ZCMV 5519	*Boophis doulioti*				AY341608			JN664265				
	Boophis tephraeomystax			AJ312117								
UADBA 24183	*Boophis tephraeomystax*				AF026344	JN133129	JN132922			DQ347234	AY880665	AF249168
KUHE 13260	*Buergeria buergeri*								AB728249			
	Buergeria buergeri						AB530012					
SCUM 061101	*Buergeria japonica*								GQ285722			
SCUM 611	*Buergeria japonica*									GQ285754		GQ285801

	Buergeria	AB99								
URE421	*japonica*	8813								
KUHE:505	*Buergeria*				AB99					
34	*japonica*				8812					
SCUM	*Buergeria*						GQ28	GQ28		EU21
050267YJ	*oxycephala*						5726	5758		5585
	Buergeria				EU92				EU92	
SN0300	*oxycephala*				4592				4536	
	Buergeria	EU21								
	oxycephala	5524								
	Buergeria			KP99						
	oxycephala			6758						
	Buergeria					GQ20				
	oxycephala					4482				
MVZ23416	*Chiromantis*	GQ20								
8	*petersii*	4733								
	Chiromantis							GQ20	DQ34	AY34
	rufescens							4605	7356	1748
CAS	*Chiromantis*	KX67			GQ20	GQ20				
254025	*rufescens*	1728			4541	4476				
NMBE							KC96			KC96
1057090	*Feihyla kajau*						1180			1234
	Ghatixalus variabilis									
	Ghatixalus	KT35	KT35					KT35	KT35	
	variabilis	9627	9621					9637	9633	
	Gracixalus	KR82		KR08		GQ28		GQ28	GQ28	GQ28
	gracilipes	7764		7671		5701		5764	5789	5807
	Kurixalus				KF93					
ACD 3857	*appendiculatus*				3134					
	Kurixalus								KF93	
ACD 6325	*appendiculatus*								3179	
FMNH	*Kurixalus*							JQ060		
267904	*appendiculatus*							911		
NMBE	*Kurixalus*		KC96			KC96	KC96			KC96
1056476	*appendiculatus*		1248			1139	1189			1232
FMNH	*Kurixalus*	JQ060								
267904	*appendiculatus*	938								
FMNH	*Kurixalus*		JQ060							
267896	*appendiculatus*		948							
	Laliostoma	AF21	AF21							
	labrosum	5278	5178							
ZCMV	*Laliostoma*						KR80			
5770	*labrosum*						9706			
ZCMV	*Laliostoma*							KR47		
5716	*labrosum*							9849		
	Laliostoma					KR47				
HER 2680	*labrosum*					9744				
	Laliostoma				KR47					
DRV 0659	*labrosum*				9583					
	Laliostoma								AY88	AF24
	labrosum								0666	9169
		KU84		KP99				KU84	KU84	KU84
	Liuixalus hainanus	0552		6848				0689	0671	0757
ZCMV	*Mantella*			JN133						
2048	*aurantiaca*			190						
	Mantella	AF21	X862		AY72			AY72	AY26	DQ28
	aurantiaca	5299	43		3611			3516	3281	2901

ZCMV 12228	*Mantidactylus femoralis*	KF426691	KF426711	KF426675	KF426668					
ZCMV 11251	*Mantidactylus femoralis*	KF426706								
MVZ 239460	*Nyctixalus pictus*				GQ204549	GQ204483				
NMBE 1056413	*Nyctixalus pictus*	JN377342	JN705355							
R081203	*Nyctixalus pictus*						GQ285729			
FMNH231094	*Nyctixalus pictus*							GQ204607		
	Nyctixalus pictus			KR087871					AY880634	GQ285805
TS814	*Occidozyga lima*	KR827960		KR087832						
	Occidozyga lima		AY880465					DQ019503	DQ019564	DQ282951
UNIMAS 8158	*Philautus macroscelis*	KX440534					KC961169			KC961203
NMBE 1056486	*Philautus macroscelis*		JN705346			KC961112				
	Polypedates eques								AY880647	
WHT 2741	*Polypedates eques*				GQ204505		GQ204447	GQ204571		
USNM GZ33880	*Polypedates leucomystax*									AB728314
WHT2564 51	*Polypedates leucomystax*							GQ204583		
SCUM_0607116L	*Polypedates leucomystax*								EU215580	
2000.8285	*Polypedates leucomystax*			KR087871						
NMBE 1057524	*Polypedates leucomystax*						KC961183			
	Polypedates leucomystax		JN541325							
KUHE:38476	*Polypedates leucomystax*				AB451716					
0045Y	*Polypedates megacephalus*	KR828026								
621Rao	*Polypedates megacephalus*						GQ285737			
KIZ 060821016	*Polypedates megacephalus*		EF564477							
SCUM 050508C	*Polypedates megacephalus*							GQ285771	EU215582	
6212 Rao	*Polypedates megacephalus*									GQ285809
KUHE:12971	*Polypedates megacephalus*				AB451722					
	Polypedates megacephalus					KC180112				
	Pseudophilautus wynaadensis	KM052243		KJ631369	GQ204502			GQ204630		

		-1.7449	101.25845	KX398867				KY886352			
	R. achantharhena										
UNIMAS 8681	*R. angulirostris*							KC961099			
ZMH A13090	*R. angulirostris*			JN377347							
FMNH 253934	*R. annamensis*						GQ204534				
FMNH 253934	*R. annamensis*						GQ204536	GQ204470			
VNMN 4092	*R. annamensis*			LC010568							
KUHE 24248	*R. arboreaus*			AY880523							
	R. arboreus				AF118476						AY880653
FMNH 235958	*R. baluensis*			KC961089				KC961093	KC961153		
ZMB 70378	*R. belalongensis*			JN377352							
ZMB 70378	*R. belalongensis*							KC961101			
ZMB 70378	*R. belalongensis*								KC961144		
UTA A 62770	*R. bengkuluensis*	-3.67348	102.535	KM212948				KY886353			
RMB 103	*R. bimaculatus*				KF933273						KF933204
	R. bimaculatus						KF933135				
FMNH 253114	*R. bipunctatus*						GQ204533	GQ204469			
KUHE:53375	*R. bipunctatus*			LC010569							
SN030035	*R. bipunctatus*									EU924518	
SN030035	*R. bipunctatus*										EU924546
THNHM18248	*R. bisacculus*				GU227280				KC961142		
22411	*R. borneensis*			AB781694							
Rao6239	*R. burmanus*			JX219422							
	R. burmanus					KP996807					
FMNH 256465	*R. calcadensis*						GQ204536			GQ204600	GQ204655
SDB.2011.291	*R. calcadensis*			KC571276							
NAP 2649	*R. calcaneus*			KX139181							
VNMN 4097	*R. calcaneus*			LC010574							

		-4.901	104.1	KX398				KY886			
	R. catamitus	49	3401	877				357			
FMNH							GQ204	GQ204			
232964	*R. chenfui*						529	467			
KIZ										EU924	EU924
060821073	*R. chenfui*									519	547
KIZ					EF564						
060821280	*R. chenfui*				466						
				KU840							
KIZ 746	*R. chenfui*			563							
						KP996					
	R. chenfui					749					
	R.			JX219							
KIZ 746	*chuyangsinensis*			451							
		4.574	96.10	KY886							
	R. cyanopunctatus	07	376	349							
		4.575	96101	MF004							
	R. cyanopunctatus	39	28	471							
		3.336	98.58	KX398							
	R. cyanopunctatus	69	391	884							
NMBE				KC961				KC961	KC961		
1056480	*R. cyanopunctatus*			084				098	152		
KIZ							EU924			EU924	EU924
060821050	*R. dennysi*						604			520	548
VNMN				LC010							
4098	*R. dennysi*			576							
YPX						KP996					
12448	*R. dennysi*					757					
									HM998		
ZCMV 110	*R. dennysi*								972		
					AF215						
	R. dennysi				185						
VNMN				LC010							
4103	*R. duboisi*			581							
						KP996					
	R. duboisi					764					
KIZ							EU924				EU924
060821003	*R. dugritei*						605				549
SCUM								GQ285	GQ285	GQ285	
051001L	*R. dugritei*							705	736	768	
				KU840							
	R. dugritei			565							
						KP996					
	R. dugritei					819					
				AB847							
9087	*R. dulitensis*			123							
FMNH							GQ204				
2357	*R. dulitensis*						532				
NMBE								KC961	KC961		
1056482	*R. dulitensis*							122	158		
					AF215						
	R. dulitensis				187						
						KR087					
	R. dulitensis					911					
NMBE								KC961	KC961		
1056492	*R. fasciatus*							104	148		

NMBE 1057405	*R. fasciatus*			KC96 1085					KC96 1147		
KIZ 060821197	*R. feae*				EF56 4474		EU92 4606			EU92 4522	EU92 4550
VNMN 3462	*R. feae*			LC01 0588							
	R. feae					JN700 897					
NMBE 1057173	*R. gadingensis*			KC96 1087				KC96 1102	KC96 1145		
FMNH 235047	*R. gauni*						GQ20 4531				
FMNH 273928	*R. gauni*			JX219 456							
NMBE 1056493	*R. gauni*							KC96 1103	KC96 1146		
NMBE 1056497	*R. harrissoni*			JN377 359				KC96 1107	KC96 1149		
CIB 097696	*R. hongchibaensis*			JN688 882							
KIZ 070521	*R. hui*						EU92 4607			EU92 4523	EU92 4551
SCUM 060425	*R. hungfuensis*			JN688 879	JN688 879	KP99 6813					EU21 5568
	R. hungfuensis					KP99 6779					
SCUM 37941C	*R. kio*							GQ28 5703	GQ28 5734	GQ28 5766	
SCUM043 7979CJ	*R. kio*					KP99 6759					
VNMN 4112	*R. kio*			LC01 0591							
	R. margaritifer	-7.144 57	107.5 183	KX39 8889				KY88 6354			
KIZ 060821140	*R. maximus*				EF56 4476		EU92 4608			EU92 4524	EU92 4552
VNMN 4113	*R. maximus*			LC01 0593							
	R. maximus					KP99 6734					
KIZ 060821020	*R. minimus*				EF56 4489		EU92 4609			EU92 4525	EU92 4553
	R. minimus					KP99 6691					
CIB GZ2009.05 .11	*R. minimus*			LC01 0594							
	R. modestus	2.619 84	98.80 542	KX39 8904				KY88 6356			
A538	*R. moltrechti*			DQ46 8676							
	R. moltrechti				AF11 8477						
	R. moltrechti					KP99 6738					
RMB 1236	*R. monticola*			AY32 6060							

FMNH 230902	*R. nigropalmatus*			GQ204710			GQ204527					
ZMH A10414	*R. nigropalmatus*							KC961094	KC961154			
	R. nigropalmatus				AF215188							
	R. nigropalmatus					KP996740						
GZ 070658	*R. nigropunctatus*			JX219430								
KIZ GZ070604	*R. nigropunctatus*					KP996794						
KIZ 060821199	*R. nigropunctatus*						EU924610					
KIZ 060821287	*R. nigropunctatus*				EF564491							
KIZ 070610	*R. nigropunctatus*										EU924555	
SCUM 070657L	*R. nigropunctatus*							GQ285704	GQ285735	GQ285767		
	R. norhayatii								AB728248			
	R. norhayatii			AB728191								
	R. norhayatii							AB728214				
KIZ 060821282	*R. omeimontis*						EU924612			EU924528	EU924556	EU924584
	R. omeimontis			KU840564								
	R. omeimontis					KP996742						
VNMN 4115	*R. orlovi*			LC010600								
FMNH 231366	*R. pardalis*							GQ204466				
ZMH A10834	*R. pardalis*								KC961157			
ZMH A13091	*R. pardalis*			JN377370								
	R. pardalis				AF215189							
ZMH A10168	*R. penanorum*			JN377349				KC961100	KC961143			
KIZ 060821289	*R. pingbianensis*						EU924613			EU924529	EU924557	EU924585
RaoYN080492	*R. pingbianensis*			JX219412								
	R. poecilonotus	3.21462	98.49793	KX398920				KY886358				
	R. promianus	-5.28364	104.5567	KX398925				KY886351				
AMNH 106539	*R. puerensis*			JN688894								
	R. puerensis					KP996728						

	R. reinwardtii	-7.24749	107.35796	MF004479					KY886355							
CIB 1306010	*R. rhodopus*			LC010612												
KIZ 060821037	*R. rhodopus*						EU924616				EU924532	EU924560	EU924588			
	R. rhodopus					JN700899										
NMBE 1056519	*R. rufipes*									KC961150						
NMBE 1057529	*R. rufipes*			KC961086					KC961108	KC961151						
KUHE 26251	*R. schlegelli*			AY880528								AY880658				
KIZ 1039	*R. taronensis*						EU924617				EU924533	EU924561	EU924589			
CIB 097685	*R. wui*			JN688881												
	R. wui					KP996733										
KIZ 0152	*Rana kukunoris*			KX269185	KX269185					GQ285748	GQ285780	GQ285798	GQ285816			
	Raorchestes longchuanensis				KC465839				GQ285713	GQ285744	GQ285776	GQ285813	GQ285813			
MRSN 6737	*Spinomantis peraccae*			FJ559307												
SMA 202	*Spinomantis peraccae*						JN133048									
	Spinomantis peraccae											AY341777				
ZRC 1.1.93	*Theloderma asperum*				GQ204776		GQ204542				GQ204606	GQ204661				
KIZ060821 2	*Theloderma asperum*			KT461929									EU924590			
	Theloderma asperum					KR087944										

Table A5.1: Locality information for all molecular sequence data used in this study.

Specimen Number	**Museum Number**	***Species***	**Country**	**Region**	**Province**	**Locality**	**Latitude**	**Longitude**	**SRA**	**16S**	**12S**	**Cox 1**	**Cytb**	**BDNF**	**POMC**	**RAG1**	**Rhodopsin**	**Tyrosinase**	**mtGenome**
	ZSM 203/2004	*Aglyptodactylus madagascariensis*										JN133054	JN132845						
		Aglyptodactylus madagascariensis															AY880664	AF249166	
	ZCMV 8046	*Aglyptodactylus madagascariensis*								KT159891						DQ347233			

ZSM 405/2000	*Blommersia wittei*	AB612030				EF396018		AY323774			
	Blommersia wittei			GU983132	AY723696				AY880667	AY341751	
ZSM 29/2006	*Boophis doulioti*			JN133081	JN132870			AY571643	AY341792		
ZCMV 14166	*Boophis doulioti*	KR025902									
ZCMV 5519	*Boophis doulioti*		AY341608			JN664265					
	Boophis tephraeomystax	AJ312117									
UADBA 24183	*Boophis tephraeomystax*		AF026344	JN133129	JN132922			DQ347234	AY880665	AF249168	
KUHE 13260	*Buergeria buergeri*						AB728249				
	Buergeria buergeri				AB530012						NC008975
SCUM 061101	*Buergeria japonica*						GQ285722				
SCUM 611	*Buergeria japonica*							GQ285754		GQ285801	
URE421	*Buergeria japonica*	AB998813									
KUHE:50534	*Buergeria japonica*				AB998812						
SCUM 050267 YJ	*Buergeria oxycephala*						GQ285726	GQ285758		EU215585	
SN0300	*Buergeria oxycephala*				EU924592				EU924536		
	Buergeria oxycephala	EU215524									NC032342
	Buergeria oxycephala			KP996758							
	Buergeria oxycephala					GQ204482					
MVZ234168	*Chiromantis petersii*	GQ204733									
	Chiromantis rufescens							GQ204605	DQ347356	AY341748	
CAS 254025	*Chiromantis rufescens*	KX671728			GQ204541	GQ204476					
	Chiromantis vittatus										KX021995
NMBE 1057090	*Feihyla kajau*						KC961180			KC961234	
	Ghatixalus variabilis										
	Ghatixalus variabilis	KT359627	KT359621					KT359637	KT359633		
	Gracixalus gracilipes	KR827764		KR087671		GQ285701		GQ285764	GQ285789	GQ285807	

	Gracixalus jinxiuensis										KX021970
ACD3857	*Kurixalus appendiculatus*				KF933134						
ACD6325	*Kurixalus appendiculatus*								KF933179		
FMNH267904	*Kurixalus appendiculatus*							JQ060911			
NMBE1056476	*Kurixalus appendiculatus*		KC961248			KC961139	KC961189			KC961232	
FMNH267904	*Kurixalus appendiculatus*	JQ060938									
FMNH267896	*Kurixalus appendiculatus*		JQ060948								
	Kurixalus odontotarus										NC032346
	Kurixalus verrucosus										NC032355
	Laliostoma labrosum	AF215278	AF215178								
ZCMV5770	*Laliostoma labrosum*						KR809706				
ZCMV5716	*Laliostoma labrosum*							KR479849			
HER2680	*Laliostoma labrosum*					KR479744					
DRV0659	*Laliostoma labrosum*				KR479583						
	Laliostoma labrosum								AY880666	AF249169	
	Liuixalus hainanus	KU840552		KP996848				KU840689	KU840671	KU840757	
ZCMV2048	*Mantella aurantiaca*			JN133190							
	Mantella aurantiaca	AF215299	X86243		AY723611			AY723516	AY263281	DQ282901	
	Mantella madagarcariensis										AB212225
ZCMV12228	*Mantidactylus femoralis*	KF426691	KF426711	KF426675	KF426668						
ZCMV11251	*Mantidactylus femoralis*	KF426706									
MVZ239460	*Nyctixalus pictus*				GQ204549	GQ204483					
NMBE1056413	*Nyctixalus pictus*	JN377342	JN705355								
R081203	*Nyctixalus pictus*						GQ285729				

	FMNH 231094	*Nyctixalus pictus*														GQ204607			
		Nyctixalus pictus										KR087871					AY880634	GQ285805	
	TS814	*Occidozyga lima*							KR827960		KR087832								
		Occidozyga lima								AY880465						DQ019503	DQ019564	DQ282951	
		Philautus macroscelis							KX440534										
	UNIMAS 8158	*Philautus macroscelis*												KC961169				KC961203	
	NMBE 1056486	*Philautus macroscelis*								JN705346			KC961112						
		Polypedates braueri																	KT921226
		Polypedates eques															AY880647		
	WHT2741	*Polypedates eques*										GQ204505		GQ204447		GQ204571			
ENS 18001		*Polypedates leucomystax*	Indonesia	Sumatra	Aceh	Rain forest lodge, Kedah	3.98021	97.25383	KY886329										
	USNMGZ33880	*Polypedates leucomystax*																AB728314	
	WHT256451	*Polypedates leucomystax*														GQ204583			
	SCUM_0607116L	*Polypedates leucomystax*															EU215580		
	2000.8285	*Polypedates leucomystax*										KR087871							
	NMBE 1057524	*Polypedates leucomystax*												KC961183					
		Polypedates leucomystax								JN541325									
	KUHE: 38476	*Polypedates leucomystax*										AB451716							
	0045Y	*Polypedates megacephalus*							KR828026										

	621Rao	*Polypedates megacephalus*											GQ285737					
	KIZ 060821 016	*Polypedates megacephalus*								EF564477								
	SCUM 050508 C	*Polypedates megacephalus*												GQ285771	EU215582			
	6212 Rao	*Polypedates megacephalus*														GQ285809		
	KUHE: 12971	*Polypedates megacephalus*										AB451722						
		Polypedates megacephalus											KC180112					NC032344
		Pseudophilautus wynaadensis							KM052243		KJ631369	GQ204502				GQ204630		
ENS 7597		*R. achantharrhena*	Indonesia	Sumatra	Bengkulu		-3.5005	102.633833	KX398866									MF066239
ENS 15995		*R. achantharrhena*	Indonesia	Sumatra	Jambi	Gunung Tujuh Road Between Panyabungan and Natal upper elevations above	-1.71674	101.36169	KX398872									
ENS 15720		*R. achantharrhena*	Indonesia	Sumatra	Sumatera Utara		0.70987	99.5287	KX398871									
ENS 15543		*R. achantharrhena*	Indonesia	Sumatra	Sumatera Utara	Pangur uran upper elevations above	2.61875	98.80412	KX398870									
ENS 15541		*R. achantharrhena*	Indonesia	Sumatra	Sumatera Utara	Pangur uran Gunung	2.61875	98.80412	KX398869									
ENS 14757		*R. achantharrhena*	Indonesia	Sumatra	Jambi	Kirinci above	-1.7449	101.25845	KX398867				KY886352					

ENS 14094	MZB 22209	*R. achantharrhena*	Indonesia	Sumatra	Lampung	Kayu Aro Hill Above Ngarip	-5.28218	104.55773	KX398868							
ENS 17621		*R. achantharrhena*	Indonesia	Sumatra	Selatan	Dempo	-4.0001	103.15877	KY886343							
	UNIMAS 8681	*R. angulirostris*											KC961099			
	ZMHA13090	*R. angulirostris*	Malaysia	Borneo	Sabah				JN377347							
	FMNH 253934	*R. annamensis*									GQ204534					
	FMNH 253934	*R. annamensis*									GQ204536	GQ204470				
	VNMN 4092	*R. annamensis*	Vietnam		GiaLai				LC010568							
	KUHE 24248	*R. arboreaus*	Japan						AY880523							
		R. arboreus								AF118476						AY880653
	FMNH 235958	*R. baluensis*	Malaysia	Borneo	Sabah				KC961089				KC961093	KC961153		
	ZMB 70378	*R. belalongensis*	Brunei	Borneo					JN377352							
	ZMB 70378	*R. belalongensis*											KC961101			
	ZMB 70378	*R. belalongensis*												KC961144		
ENS 21063		*R. bengkuluensis*	Indonesia	Sumatra	Aceh	Pidie Jaya	4.86024	96.21368	KY886350							
ENS 20653		*R. bengkuluensis*	Indonesia	Sumatra	Aceh	Nagan Raya	3.95645	96.85304	MF004472							
ENS 18717	UTA A 62770	*R. bengkuluensis*	Indonesia	Sumatra	Bengkulu	Bengkulu Tenggah	-3.67348	102.535	KM212948				KY886353			
ENS 14834		*R. bengkuluensis*	Indonesia	Sumatra	Lampung	Kubuperahu	-5.06011	104.03222	KY886332							
	RMB 103	*R. bimaculatus*								KF933273						KF933204
		R. bimaculatus									KF933135					
	FMNH 253114	*R. bipunctatus*									GQ204533	GQ204469				
	KUHE: 53375	*R. bipunctatus*	Malaysia	Malay	Pahang				LC010569							

				Peninsula													
	SN030035	*R. bipunctatus*												EU924518			KX022003
	SN030035	*R. bipunctatus*													EU924546		
	THNHM18248	*R. bisacculus*								GU227280			KC961142				
	22411	*R. borneensis*	Malaysia	Borneo	Sabah				AB781694								
	Rao6239	*R. burmanus*	China		Xizang				JX219422								
		R. burmanus									KP996807						
	FMNH256465	*R. calcadensis*										GQ204536		GQ204600	GQ204655		
	SDB.2011.291	*R. calcadensis*	India		Kerala				KC571276								
	NAP2649	*R. calcaneus*							KX139181								
	VNMN4097	*R. calcaneus*	Vietnam		DakLak				LC010574								
ENS7725		*R. catamitus*	Indonesia	Sumatra	Bengkulu	Dempo	-2.0405	101.305167		JF748390							
ENS7678		*R. catamitus*	Indonesia	Sumatera	Selatan	Gunung Dempo, SE side	-4.041667	103.152833	SAMN05426776								
ENS7677		*R. catamitus*	Indonesia	Sumatera	Selatan	Gunung Dempo, SE side	-4.041667	103.152833		KX398882							
ENS7662		*R. catamitus*	Indonesia	Sumatera	Selatan	Dempo	-4.0505	103.14		JF748388							
ENS7657		*R. catamitus*	Indonesia	Sumatera	Selatan	Dempo	-4.0505	103.14		JF748387							
ENS7610		*R. catamitus*	Indonesia	Sumatera	Bengkulu	Kaba	-3.500333	102.634667		JF748392							
ENS7609		*R. catamitus*	Indonesia	Sumatera	Bengkulu	Kaba	-3.500333	102.634667		JF748391							
ENS19165		*Rhacophorus* sp.	Indonesia	Sumatera	Agam	Gunung Singgalang above Desa	-0.37528	100.36335		MF004476							

						Beringin				
ENS 17636		*R. catamitus*	Indonesia	Sumatra	Selatan	Dempo	-4.00074	103.15851		KY886346
ENS 17625		*R. catamitus*	Indonesia	Sumatra	Selatan	Dempo	-4.00046	103.15868		KY886344
ENS 17529		*R. catamitus*	Indonesia	Sumatra	Selatan	Gunung Dempo near Kampung Empat	-4.04591	103.14983		KY886341
ENS 17495		*R. catamitus*	Indonesia	Sumatra	Selatan	Dempo	-4.04591	103.14983		MF004478
ENS 17428		*R. catamitus*	Indonesia	Sumatra	Selatan	Gunung Patah near Desa Segamit	-4.21869	103.46786		KY886338
ENS 17414		*R. catamitus*	Indonesia	Sumatra	Selatan	Gunung Patah near Desa Segamit	-4.21742	103.46823		KY886337
ENS 16783	MZB 26279	*Rhacophorus* sp.	Indonesia	Sumatra	Sumatera Utara	Gunung Sibuatan, Above Kampung Naga Linga	2.91076	98.46313	SAMN05426788	KX398913
ENS 16782	UTA A 63576	*Rhacophorus* sp.	Indonesia	Sumatra	Sumatera Utara	Gunung Sibuatan, Above Kampung Naga Linga	2.91076	98.46313	SAMN05426796	
ENS 16682	UTA A 63574	*Rhacophorus* sp.	Indonesia	Sumatra	Sumatera Utara	Slope of Dolok k Malea above Kampung	0.975	99.57959	SAMN05426801	

						Mompang Slope of Dolok Malea above Kampung				
ENS 16681	MZB 26290	*Rhacophorus* sp.	Indonesia	Sumatra	Sumatera Utara	Mompang	0.975	99.57959	SAMN05426794	KX398873
ENS 16633		*Rhacophorus* sp.	Indonesia	Sumatra	Sumatera Utara	Hambung Hasundutan	2.18325	98.60513	SAMN05426783	
ENS 16616		*Rhacophorus* sp.	Indonesia	Sumatra	Sumatera Utara	Taman Eden. Gunung Pangulubao	2.61587	99.05087	SAMN05426792	
ENS 16615		*Rhacophorus* sp.	Indonesia	Sumatra	Sumatera Utara	Taman Eden. Gunung Pangulubao	2.61587	99.05087	SAMN05426785	KX398876
ENS 16585		*Rhacophorus* sp.	Indonesia	Sumatra	Sumatera Utara	Gunung Pangulubao	2.60514	99.04629	SAMN05426777	
ENS 15986		*R. catamitus*	Indonesia	Sumatra	Jambi	Gunung Tujuh	-1.71676	101.36171	SAMN05426803	KY886334
ENS 15981		*R. catamitus*	Indonesia	Sumatra	Jambi	Gunung Tujuh	-1.71677	101.36172	SAMN05426774	KY886333
ENS 15614		*R. catamitus*	Indonesia	Sumatra	Sumatera Utara	Vicinity of Tele	2.55397	98.59806	SAMN05426772	
ENS 15594		*R. catamitus*	Indonesia	Sumatra	Sumatera Utara	Vicinity of Tele	2.55397	98.59806		KX398880
ENS 14805		*R. catamitus*	Indonesia	Sumatra	Lampung	Mountain Above Ngarip Mauradua, Reman Jaya, Gunung	-5.2818	104.55767	SAMN05426779	
ENS 14727		*R. catamitus*	Indonesia	Sumatra	Sumatera Selatan		-4.90149	104.13401	SAMN05426800	KX398878

ENS 14726		*R. catamitus*	Indonesia	Sumatra	Sumatera Selatan	Pesagi (localy known as Masagi) Maura dua, Remanan Jaya, Gunung Pesagi (localy known as Masagi)	-4.90149	104.13401	SAMN05426798	KX398877			KY886357	
ENS 14427		*R. catamitus*	Indonesia	Sumatera	Barat	Road between Sungai Penuh and Tapan, W. of crest	-2.04294	101.31129		KY886331				
ENS 14426		*R. catamitus*	Indonesia	Sumatra	Sumatera Barat	Road between Sungai Penuh and Tapan, W. of crest	-2.04294	101.31129	SAMN05426771	KX398879				
ENS 14093		*R. catamitus*	Indonesia	Sumatra	Lampung	Hill Above Ngarip	-5.28218	104.55773		KX398875				
ENS 14091		*R. catamitus*	Indonesia	Sumatra	Lampung	Hill Above Ngarip	-5.28218	104.55773		KX398883				MF066241
ENS 14087		*R. catamitus*	Indonesia	Sumatra	Lampung	Hill Above Ngarip	-5.28218	104.55773	SAMN05426787					
ENS 14085		*R. catamitus*	Indonesia	Sumatra	Lampung	Hill Above Ngarip	-5.28218	104.55773		KX398874				
ENS 14084		*R. catamitus*	Indonesia	Sumatra	Lampung	Hill Above Ngarip	-5.28171	104.55784		KX398881				
	FMNH 232964	*R. chenfui*									GQ204529	GQ204467		
	KIZ 060821073	*R. chenfui*											EU924519	EU924547

	KIZ 060821 280	*R. chenfui*								EF564 466								
	KIZ 746	*R. chenfui*	China						KU840 563		KP9967 49							
		R. chenfui																
	KIZ 746	*R. chuyangsinensis*	Vietnam		Lam Dong				JX219 451									
ENS 20990		*R. cyanopunctatus*	Indonesia	Sumatra	Aceh	Aceh Barat	4.574 07	96.103 76	KY886 349									
ENS 20985		*R. cyanopunctatus*	Indonesia	Sumatra	Aceh	Aceh Barat Road from Medan to	4.575 39	96101 28	MF004 471									
ENS 15499		*R. cyanopunctatus*	Indonesia	Sumatra	Sumatera Utara	Berastagi	3.336 69	98.583 91	KX398 884									
	NMBE 1056480	*R. cyanopunctatus*	Malaysia	Borneo	Sarawak				KC961 084				KC9610 98	KC9611 52				
	KIZ 060821 050	*R. dennysi*										EU9246 04			EU9245 20	EU924 548		
	VNMN 4098	*R. dennysi*	Vietnam		Vinh Phuc				LC010 576									
	YPX 12448	*R. dennysi*									KP9967 57							
	ZCMV 110	*R. dennysi*												HM9989 72				
		R. dennysi								AF215 185								KM03 5412
	VNMN 4099	*R. dorsoviridis*	Vietnam		Sonla				LC010 577									
	VNMN 4103	*R. duboisi*	Vietnam		LaoCai				LC010 581									
		R. duboisi									KP9967 64							
	KIZ 060821 003	*R. dugritei*	isolate									EU9246 05				EU924 549		
	SCUM 051001 L	*R. dugritei*											GQ2857 05	GQ2857 36	GQ2857 68			
		R. dugritei	China		Sichuan				KU840 565									
		R. dugritei									KP9968 19							
	9087	*R. dulitensis*	Malaysia	Borneo	Sabah				AB847 123									
	FMNH 2357	*R. dulitensis*										GQ2045 32						

NMBE 1056482	*R. dulitensis*								KC961122	KC961158		
	R. dulitensis					AF215187						
	R. dulitensis						KR087911					
RH06085	*R. exechopygus*	Vietnam		Ubo	GQ469980							
NMBE 1056492	*R. fasciatus*								KC961104	KC961148		
NMBE 1057405	*R. fasciatus*	Malaysia	Borneo	Sarawak	KC961085					KC961147		
KIZ 060821197	*R. feae*					EF564474		EU924606			EU924522	EU924550
VNMN3462	*R. feae*	Vietnam		Kon Tum	LC010588							
	R. feae						JN700897					
NMBE 1057173	*R. gadingensis*	Malaysia	Borneo	Sarawak	KC961087				KC961102	KC961145		
FMNH 235047	*R. gauni*							GQ204531				
FMNH 273928	*R. gauni*	Malaysia	Borneo	Sarawak	JX219456							
NMBE 1056493	*R. gauni*								KC961103	KC961146		
KIZ 1039	*R. gongshanensis*	China		Yunnan	EF564568							
NMBE 1056497	*R. harrissoni*	Malaysia	Borneo	Sarawak	JN377359				KC961107	KC961149		
ABV00238	*R. helenae*	Vietnam			KX139178							
CIB 097696	*R. hongchibaensis*	China		Chongqing	JN688882							
KIZ 070521	*R. hui*							EU924607			EU924523	EU924551
KIZ 07052101	*R. hui*	China		Yunnan	EU924622							
SCUM 060425	*R. hungfuensis*				JN688879	JN688879	KP996813					EU215568
	R. hungfuensis						KP996779					
MZB 23626	*R. indonesiensis*	Indonesia	Sumatra	Jambi	AB983368							
SCUM 37941C	*R. kio*								GQ285703	GQ285734	GQ285766	

	SCUM0437979											KP996759					
	CJ VNMN4112	*R. kio*	Vietnam		CaoBang				LC010591								
		R. kio															NC032356
		R. kio															
	SDB.2010.330	*R. lateralis*	India		Bygoor				KC571277								
	M3003	*R. malabaricus*	India						GU136112								
		R. malabaricus								GU136096							
		R. malabaricus										KP996761					
		R. malabaricus											GU136142				
ENS7418		*R. margaritifer*	Indonesia	Java	Bogor	Taman Safari	-6.726	106.950833	KX398886								
ENS7417		*R. margaritifer*	Indonesia	Java	Bogor	Taman Safari Road from S. Coast of Java to	-6.726	106.950833	KX398885								
ENS16188		*R. margaritifer*	Indonesia	Java	Jawa Barat	Gunung Patuha	-7.2337	107.35601	KX398888								
ENS16162		*R. margaritifer*	Indonesia	Java	Jawa Barat	Gunung Tilu LIPI	-7.14457	107.5183	KX398889				KY886354				
ENS15850		*R. margaritifer*	Indonesia	Java		Botanical Garden	-5.55809	105.08358	KX398887								
	KIZ060821140	*R. maximus*								EF564476		EU924608			EU924524	EU924552	
	VNMN4113	*R. maximus*	Vietnam		Nghe An				LC010593								
		R. maximus									KP996734						
	KIZ060821020	*R. minimus*								EF564489		EU924609			EU924525	EU924553	
		R. minimus									KP996691						
	CIBGZ2009.05.11	*R. minimus*							LC010594								

ENS 7670		*R. modestus*	Indonesia	Sumatra	Sumatera Selatan	Dempo	-4.0505	103.14		KX398891	
ENS 19409		*R. modestus*	Indonesia	Sumatra	Aceh	High point on Meulaboh-Takengon road	4.38367	96.51633		KY886348	
ENS 19407		*R. modestus*	Indonesia	Sumatra	Aceh	High point on Meulaboh-Takengon road	4.38367	96.51633		MF004475	
ENS 18176		*R. modestus*	Indonesia	Sumatra	Bengkulu	Kaba	-3.50795	102.62886		MF004477	
ENS 17500		*R. modestus*	Indonesia	Sumatra	Selatan	Dempo	-4.04591	103.14983		KY886339	
ENS 16988		*R. modestus*	Indonesia	Sumatra	Sumatera Utara	Tujuh	-1.71076	101.36986	SAMN05426782		
ENS 16853		*R. modestus*	Indonesia	Sumatra	Sumatera Utara	Higher Elevations above Pangururan	2.61984	98.80542		KX398904	KY886356
ENS 16852		*R. modestus*	Indonesia	Sumatra	Sumatera Utara	Higher Elevations above Pangururan	2.61984	98.80542	SAMN05426799		
ENS 16851		*R. modestus*	Indonesia	Sumatra	Sumatera Utara	Higher Elevations above Pangururan	2.61984	98.80542	SAMN05426773	KX398903	
ENS 16850		*R. modestus*	Indonesia	Sumatra	Sumatera Utara	Higher Elevations above Pangururan	2.61984	98.80542		KX398902	
ENS 15993	MZB 22227	*R. modestus*	Indonesia	Sumatra	Jambi	Gunung Tujuh	-1.71675	101.3617		KX398906	

ENS 15972		*R. modestus*	Indonesia	Sumatra	Jambi	Gunung Tujuh	-1.71679	101.36174		KX398899
ENS 15970		*R. modestus*	Indonesia	Sumatra	Jambi	Gunung Tujuh upper elevations above	-1.7168	101.36175		KX398905
ENS 15540		*R. modestus*	Indonesia	Sumatra	Sumatera Utara	Pangguran upper elevations above	2.61875	98.80412		KX398900
ENS 15539		*R. modestus*	Indonesia	Sumatra	Sumatera Utara	Pangguran	2.61875	98.80412	SAMN05426791	
ENS 14776		*R. modestus*	Indonesia	Sumatra	Jambi	Gunung Kirinci above Kayu Aro	-1.7457	101.25844		KX398898
ENS 14770		*R. modestus*	Indonesia	Sumatra	Jambi	Gunung Kirinci above Kayu Aro	-1.7449	101.25845		KX398897
ENS 14766		*R. modestus*	Indonesia	Sumatra	Jambi	Gunung Kirinci above Kayu Aro	-1.7449	101.25845		KX398896
ENS 14753		*R. modestus*	Indonesia	Sumatra	Jambi	Gunung Kirinci above Kayu Aro	-1.7457	101.25844		KX398895
ENS 14747		*R. modestus*	Indonesia	Sumatra	Jambi	Gunung Kirinci above Kayu Aro	-1.7449	101.25845		KX398894
ENS 14357	MZB 22208	*R. modestus*	Indonesia	Sumatra	Jambi	Trail on Mt. Kerinci	-1.73906	101.25962		KX398893
ENS 14350		*R. modestus*	Indonesia	Sumatra	Jambi	Trail on Mt. Kerinci	-1.73906	101.25962		KX398892

ENS 7595		*R. modestus*	Indonesia	Sumatra	Bengkulu		-3.500333	102.634667	KX398890									MF066238
	A538	*R. moltrechti*	Tiawan						DQ468676									
		R. moltrechti								AF118477								
		R. moltrechti									KP996738							
	RMB 1236	*R. monticola*	Sulawesi						AY326060									
ENS 16368		*R. nigropalmatus*	Indonesia	Sumatra	Utara	Vicinity of Bandar Baru	3.30719	98.5747	KX398907									
	FMNH 230902	*R. nigropalmatus*	Malaysia	Borneo	Sabah				GQ204710			GQ204527						
	ZMH A10414	*R. nigropalmatus*											KC961094	KC961154				
		R. nigropalmatus								AF215188								
		R. nigropalmatus									KP996740							
	GZ 070658	*R. nigropunctatus*	China		Guizhou				JX219430									
	KIZ GZ070604	*R. nigropunctatus*									KP996794							
	KIZ 060821199	*R. nigropunctatus*										EU924610						
	KIZ 060821287	*R. nigropunctatus*								EF564491								
	KIZ 070610	*R. nigropunctatus*														EU924555		
	SCUM 070657L	*R. nigropunctatus*											GQ285704	GQ285735	GQ285767			
		R. norhayatii												AB728248				
		R. norhayatii	Malaysia	Malay Peninsula	Johor				AB728191									
		R. norhayatii											AB728214					
	KIZ 060821282	*R. omeimontis*										EU924612			EU924528	EU924556	EU924584	
		R. omeimontis	China		Sichuan				KU840564									

		R. omeimontis								KP996742								
	VNMN4115	*R. orlovi*	Vietnam		Nghe An				LC010600									
ENS7577		*R. pardalis*	Indonesia	Sumatra	Bengkulu		-3.684667	102.5365	KX398908									
ENS20991		*R. pardalis*	Indonesia	Sumatra	Aceh	Aceh Barat	4.57506	96.10265	MF004470									
	FMNH231366	*R. pardalis*										GQ204466						
	ZMH A10834	*R. pardalis*											KC961157					
	ZMH A13091	*R. pardalis*	Malaysia	Borneo	Sarawak				JN377370									
		R. pardalis								AF215189								
	ZMH A10168	*R. penanorum*	Malaysia	Borneo	Sarawak				JN377349			KC961100	KC961143					
	KIZ 060821289	*R. pingbianensis*									EU924613			EU924529	EU924557	EU924585		
	RaoYN080492	*R. pingbianensis*	China		Guangxi				JX219412									
ENS7647		*R. poecilonotus*	Indonesia	Sumatra	Sumatera Selatan	Dempo	-4.0505	103.14	KX398918									
ENS7606		*R. poecilonotus*	Indonesia	Sumatra	Bengkulu	Kaba	-3.5005	102.633833	KX398917									MF066240
ENS19423		*R. poecilonotus*	Indonesia	Sumatra	Aceh	Aceh Tengah	4.6654	96.80505	MF004474									
ENS17629		*R. poecilonotus*	Indonesia	Sumatra	Selatan	Dempo Gunung Patah near Desa	-4.00045	103.15836	KY886345									
ENS17576		*R. poecilonotus*	Indonesia	Sumatra	Selatan	Segamit	-4.20963	103.4285	KY886342									
ENS17501		*R. poecilonotus*	Indonesia	Sumatra	Selatan	Dempo Higher Elevations above	-4.04591	103.14983	KY886340									
ENS16854		*R. poecilonotus*	Indonesia	Sumatra	Sumatera Utara	Pangguran Higher	2.61984	98.80542	KX398919									
ENS16847		*R. poecilonotus*	Indonesia	Sumatra	Sumatera Utara	Elevations above	2.61914	98.80482	KX398901									

					Pangur uran					
ENS 16722	*R. poecilonotus*	Indonesia	Sumatra	Sumatera Utara	Batang Gadis	0.70869	99.51923	SAMN05426786		
ENS 16483	*R. poecilonotus*	Indonesia	Sumatra	Sumatera Utara	Sibayak, Foot of mountain	3.21462	98.49793	SAMN05426775		
ENS 16481	*R. poecilonotus*	Indonesia	Sumatra	Sumatera Utara	Sibayak, Foot of mountain	3.21462	98.49793		KX398914	
ENS 16480	*R. poecilonotus*	Indonesia	Sumatra	Sumatera Utara	Sibayak, Foot of mountain	3.21462	98.49793	SAMN05426789	KX398920	KY886358
ENS 15976	*R. poecilonotus*	Indonesia	Sumatra	Jambi	Gunung Tujuh	-1.71678	101.36173		KX398922	
ENS 15965	*R. poecilonotus*	Indonesia	Sumatra	Jambi	Gunung Tujuh	-1.71681	101.36176		KX398921	
ENS 15610	*R. poecilonotus*	Indonesia	Sumatra	Sumatera Utara	Vicinity of Tele	2.55397	98.59806	SAMN05426797		
ENS 15609	*R. poecilonotus*	Indonesia	Sumatra	Sumatera Utara	Vicinity of Tele	2.55397	98.59806		KX398924	
ENS 15583	*R. poecilonotus*	Indonesia	Sumatra	Sumatera Utara	Vicinity of Tele	2.56452	98.58595		KX398923	
ENS 15535	*R. poecilonotus*	Indonesia	Sumatra	Sumatera Utara	upper elevations above Pangur uran	2.61875	98.80412	SAMN05426795	KX398916	
ENS 15534	*R. poecilonotus*	Indonesia	Sumatra	Sumatera Utara	upper elevations above Pangur uran	2.61875	98.80412	SAMN05426781	KX398915	
ENS 15533	*R. poecilonotus*	Indonesia	Sumatra	Sumatera Utara	upper elevations above Pangur uran	2.61875	98.80412	SAMN05426802	KX398911	
ENS 15532	*R. poecilonotus*	Indonesia	Sumatra	Sumatera Utara	upper elevations	2.61875	98.80412	SAMN054		

						above Pangur uran upper elevations			26790			
ENS 15531		*R. poecilonotus*	Indonesia	Sumatra	Sumatera Utara	above Pangur uran	2.61875	98.80412		KX398912		
ENS 14104		*R. poecilonotus*	Indonesia	Sumatra	Lampung	Hill Above Ngarip	-5.28721	104.55401	SAMN05426784	KX398910		
ENS 14103		*R. poecilonotus*	Indonesia	Sumatra	Lampung	Hill Above Ngarip	-5.28721	104.55401	SAMN05426793	KX398909		
ENS 14102		*R. poecilonotus*	Indonesia	Sumatra	Lampung	Hill Above Ngarip	-5.28721	104.55401	SAMN05426778	KY886330		
MBH 5514		*R. promianus*	Indonesia	Sumatra	Lampung	Raja Basa	-5.808	105.627		KX398927		
ENS 16994		*R. promianus*	Indonesia	Sumatra	Lampung	Above Ngarip Town, forest above 1300 m	-5.28364	104.5567		KX398925		KY886351
ENS 14604		*R. promianus*	Indonesia	Sumatra	Lampung	Gunung Tanggamus, above Gisting	-5.42364	104.6922		KX398928		
ENS 14201		*R. promianus*	Indonesia	Sumatra	Lampung	Above Ngarip Town, forest above 1300 m	-5.28364	104.5567		KX398926		
ENS 19612		*R. prominanus*	Indonesia	Sumatra	Barat	Bukittinggi	-0.29496	100.38218		MF004473		
	SDB.2011.1010	*R. pseudomalabaricus*	India		Kerala					KC593855		
	AMNH 106539	*R. puerensis*	Vietnam		HaGiang					JN688894		
		R. puerensis									KP996728	
MBH 5320		*R. reinwardtii*	Indonesia	Java	Java Barat	Desa Su Kamahi; P.T. near	640.32	10652.78		KX398929		

						bajor; very disturbed farmland											
ENS 16447		*R. reinwardtii*	Indonesia	Java	Jawa Barat	Bandung Road from S. Coast of Java to	-7.28555	107.53059	KY886335 MF004479								
ENS 16180		*R. reinwardtii*	Indonesia	Java	Jawa Barat	Gunung Patuha Road from S. Coast of Java to	-7.24749	107.35796				KY886355					
ENS 16179		*R. reinwardtii*	Indonesia	Java	Jawa Barat	Gunung Patuha	-7.24749	107.35796	KX398930								
	KIZ 060821224	*R. reinwardtii*												EU924531	EU924559	EU924587	
	NMBE 1056517	*R. reinwardtii*											KC961155				
	ZRC1.1.5273	*R. reinwardtii*	voucher							JN700900	GQ204537						
		R. reinwardtii															
	CIB 1306010	*R. rhodopus*	China		Yunnan				LC010612								
	KIZ 060821037	*R. rhodopus*									EU924616			EU924532	EU924560	EU924588	
		R. rhodopus								JN700899							
	VNMN 3446	*R. robertingeri*	Vietnam		Kon Tum				LC010615								
	NMBE 1056519	*R. rufipes*											KC961150				
	NMBE 1057529	*R. rufipes*	Malaysia	Borneo	Sarawak				KC961086			KC961108	KC961151				
	KUHE 26251	*R. schlegelli*	Japan						AY880528						AY880658		AB202079
	KIZ 1039	*R. taronensis*									EU924617			EU924533	EU924561	EU924589	

Rao6237	*R. translineatus*	China	Xizang	JX219449								
VNMN4125	*R. vampyrus*	Vietnam	KhanhHoa	LC010616								
Rao6254	*R. verrucopus*	China	Motuo	JX219436								
CIB097685	*R. wui*	China	Hubei	JN688881								
	R. wui					KP996733						
KIZ0152	*Rana kukunoris*			KX269185	KX269185				GQ285748	GQ285780	GQ285798	GQ285816	
	Raorchestes longchuanensis				KC465839			GQ285713	GQ285744	GQ285776	GQ285813	GQ285813	KX022005
MRSN6737	*Spinomantis peraccae*			FJ559307									
SMA202	*Spinomantis peraccae*						JN133048						
	Spinomantis peraccae										AY341777		
ZRC1.1.93	*Theloderma asperum*				GQ204776		GQ204542			GQ204606	GQ204661		
KIZ0608212	*Theloderma asperum*			KT461929								EU924590	
	Theloderma asperum					KR087944							

Table A5.2: Results of demographic modeling analyses with SNP data.

R. poecilonotus

Model	log-likelihood	theta	AIC	ΔAIC	AIC Weight	nu1	nu2	nu1a	nu2a	nu1b	nu1b	m12	m21	T1	T2
Split with Secondary Contact, Asymmetrical Gene Flow, Size Change	**-71.48**	**424.18**	**158.96**	**0**	**0.46091223**	**—**	**—**	**0.274**	**0.8573**	**0.813**	**0.2955**	**0.0637**	**0.0189**	**0.346**	**0.351**
Split with Secondary Contact, Symmetrical Gene Flow, Size Change	**-72.67**	**404.49**	**159.34**	**0.38**	**0.38115557**	**—**	**—**	**0.290**	**10.27**	**0.814**	**0.2894**	**0.0539**	**—**	**0.340**	**0.445**
Split with Symmetric Migration, Size Change	-74.26	310	162.52	3.56	0.07772738	—	—	0.352	1.1927	1.079	0.5016	0.056	—	0.550	0.803
Split with No Migration, Size Change	-76.3	398.12	164.6	5.64	0.02747310	—	—	0.528	0.5803	0.890	0.2519	—	—	0.673	0.122
Split with Symmetric Migration	-78.83	87.83	165.66	6.7	0.01617080	3.088	2.119	—	—	—	—	0.0164	—	7.687	—
Split with Asymmetric Migration	-77.95	160.79	165.9	6.94	0.01434222	1.715	1.068	—	—	—	—	0.0218	0.0443	3.772	—
Split with Ancient Symmetrical Migration	-78.35	81.49	166.7	7.74	0.00961387	2	4	—	—	—	—	0.0184	—	8.298	0.178
Split, Secondary Contact, Symmetrical Gene Flow	-78.72	131.16	167.44	8.48	0.00664063	3.311	2.255	—	—	—	—	0.027	—	2	1
Split with Asymmetrical Migration, Size Change	-76.68	204.09	169.36	10.4	0.00254265	2.086	1.427	—	—	—	—	0.0332	0.0254	2.918	1.753
Split, Secondary Contact, Asymmetrical Gene Flow	-79.05	289.35	170.1	11.1	0.00175629	9	8	—	—	—	—	0.0427	0.0773	1.36	2
Split with Ancient Asymmetrical Migration, Size Change	-77.34	311.28	170.68	11.7	0.00131417	0.915	0.608	1.425	23.869	1.298	0.7287	0.0332	0.0254	1.032	1.220
Split with Ancient Asymmetrical Migration	-80.96	1345.3 6	173.92	14.9 6	0.00026007 3	0.192 8	0.113 3	—	—	—	—	0.7041	6.0715	4.791 8	0.142 4
Split with No Migration	-85.09	395.64	176.18	17.2 2	8.40123E-05	0.627	0.441 1	—	—	—	—	—	—	0.786 4	—
Split with Ancient Symmetrical Migration, Size Change	-83.58	110.23	181.16	22.2	6.9655E-06	—	—	1.217 5	4.445	2.379 9	1.5628	1.2084	—	6.214 2	2.663 3
No Divergence	-3163.64	455.47	6333.2 8	—	—	—	—	—	—	—	—	—	—	—	—

Rhacophorus sp. and *R. catamitus*

Model	log-likelihood	theta	AIC	ΔAIC	AIC Weight	nu1	nu2	nu1a	nu2a	nu1b	nu1b	m12	m21	T1	T2
Split with Secondary Contact, Symmetrical Gene Flow, Size Change	**-102.57**	**239.07**	**219.14**	**0**	**0.41904068**	**—**	**—**	**0.326**	**0.8922**	**0.376**	**3.0087**	**0.0375**	**—**	**0.881**	**0.188**
Split with Symmetric Migration	**-106.51**	**61.2**	**221.02**	**1.88**	**0.16368895**	**1.379**	**5.303**	**4**	**5**	**7**	**—**	**—**	**—**	**7.369**	**—**
Split, Secondary Contact, Symmetrical Gene Flow	-106.1	91.9	222.2	3.06	0.09073725	0.912	3.553	—	—	—	—	0.0171	—	3.101	—
Split with No Migration, Size Change	-105.24	302.27	222.48	3.34	0.07888317	1	5	0.206	—	0.279	—	—	—	2	1.393
Split with Secondary Contact, Asymmetrical Gene Flow, Size Change	-103.35	246.41	222.7	3.56	0.07066624	—	—	3	0.3591	6	1.6534	—	—	0.269 7	0.344
Split with Symmetric Migration, Size Change	-104.47	154.68	222.94	3.8	0.06267533	—	—	0.102 4	—	0.342 5	—	—	—	0.158 7	0.880 7
Split, Secondary Contact, Asymmetrical Gene Flow	-106.12	79.79	224.24	5.1	0.03271939	1.041	4.078	1.342 8	1.0769	0.524 3	2.3292	0.0217	—	0.979	1.316 3
Split with Asymmetric Migration	-107.23	155.56	224.46	5.32	0.02931115	0.528 1	2.049 1	—	—	—	—	0.0173 0.0345	0.0105 0.0143	0.679 9 2.275 6	4.732 3 —

Split with Asymmetrical Migration, Size Change	-104.44	229.43	224.88	5.74	0.02375915 7	—	—	6.981 1	0.2439	0.352 3	1.5655	0.0344	0.0263	0.173 5	1.051 5
Split with Ancient Symmetrical Migration, Size Change	-105.49	201.54	224.98	5.84	0.02260040 9	—	—	0.428 6	0.545	0.409 6	2.1719	0.4849	—	2.155 6	0.669
Split with Ancient Asymmetrical Migration, Size Change	-106.12	824.48	228.24	9.1	0.00442808 9	—	—	0.270 4	0.0387	0.093 5	0.5074	0.0439	12.554 9	0.865 1	0.201 6
Split with Ancient Asymmetrical Migration	0	84.17	230.42	11.2 8	0.00148879 6	0.970 8	3.823 9	—	—	—	—	0.0441	0.0108	4.659 7	0.531 8
Split with Ancient Symmetrical Migration	-117.21	241.27	244.42	25.2 8	1.35761E-06	0.363 9	1.260 4	—	—	—	—	0.2866		0.006 8	1.003 8
Split with No Migration	-126.05	314.37	258.1	38.9 6	1.4528E-09	0.263 4	0.790 7	—	—	—	—	—	—	0.653 3	—
No Divergence	-4818.1	336.07	9642.2	—	—	—	—	—	—	—	—	—	—	—	—

R. catamitus

Model	log-likelihood	theta	AIC	ΔAIC	AIC Weight	nu1	nu2	nu1a	nu2a	nu1b	nu1b	m12	m21	T1	T2
Split with No Migration, Size Change	**-28.98**	**709.86**	**69.96**	**0**	**0.47130188 3**	—	—	**0.184 8**	**1.4868**	**1.000 5**	**0.3618**	—	—	**0.271 3**	—
Split with Ancient Symmetrical Migration, Size Change	**-29.26**	**576.47**	**72.52**	**2.56**	**0.13103950 3**	—	—	**0.293 9**	**1.84**	**3.451 6**	**0.1215**	**0.9385**	—	**0.200 5**	—
Split with Symmetric Migration, Size Change	**-29.46**	**409.69**	**72.92**	**2.96**	**0.10728607 1**	—	—	**0.431 5**	**1.1945**	**1.729**	**0.0522**	**1.6678**	—	**0.345 7**	—
Split with Secondary Contact, Symmetrical Gene Flow, Size Change	-29.56	604.61	73.12	3.16	0.09707645 2	—	—	0.054 8	1.5334	0.741 4	0.0762	0.3677	—	0.634 5	—
Split with Asymmetrical Migration, Size Change	-28.95	344.61	73.9	3.94	0.06572627 9	—	—	0.123 8	4.6593	1.489 6	0.114	0.0252	1.9644	1.135	—
Split with Ancient Asymmetrical Migration, Size Change	-29.06	551.76	74.12	4.16	0.05887984 4	—	—	0.254 8	8.1346	1.049 9	0.702	0.1482	6.5257	0.441 6	—
Split with Secondary Contact, Asymmetrical Gene Flow, Size Change	-29.07	273.62	74.14	4.18	0.05829398	—	—	0.829 8	3.8916	3.881 8	0.2186	0.0264	3.1236	0.399 2	—
Split with Asymmetric Migration	-34.52	304.22	79.04	9.08	0.00503039 7	1.045 4	0.085 9	—	—	—	—	0.01	3.2749	—	—
Split, Secondary Contact, Asymmetrical Gene Flow	-33.66	273.63	79.32	9.36	0.00437321 7	1.213 7	0.733 3	—	—	—	—	0.0103	3.4224	0.113 7	—
Split with No Migration	-38.66	599.87	83.32	13.3 6	0.00059185 1	0.479 9	0.953 4	—	—	—	—	—	—	—	—
Split with Symmetric Migration	-38.72	581.39	85.44	15.4 8	0.00020505	0.495 9	0.011	—	—	—	—	1.0279	—	—	—
Split with Ancient Symmetrical Migration	-37.97	371.56	85.94	15.9 8	0.00015969 3	0.766 7	0.390 6	—	—	—	—	1.6876	—	0.969 4	—
Split with Ancient Asymmetrical Migration	-38.72	611.32	89.44	19.4 8	2.77505E-05	0.466 9	0.242	—	—	—	—	2.1913	0.0234	0.895 5	—
Split, Secondary Contact, Symmetrical Gene Flow	-40.96	430.67	91.92	21.9 6	8.0306E-06	0.697 7	0.169 2	—	—	—	1.6567	—	—	0.370 1	—
No Divergence	-1563.03	925.54	3132.06	—	—	—	—	—	—	—	—	—	—	—	—

R. modestus

Model	log-likelihood	theta	AIC	ΔAIC	AIC Weight	nu1	nu2	nu1a	nu2a	nu1b	nu1b	m12	m21	T1	T2
Split with No Migration	**-28.14**	**1118.3**	**62.28**	**0**	**0.43189744**	**0.081 3**	**0.305 6**	—	—	—	—	—	—	**0.370 1**	—
Split with Symmetric Migration	**-27.81**	**859.47**	**63.62**	**1.34**	**0.22100562**	**0.110 2**	**0.55**	—	—	—	—	**0.0223**	—	**0.814 5**	—
Split with Asymmetric Migration	-27.72	826.15	65.44	3.16	0.08896012	0.118 8	0.527 9	—	—	—	—	0.0135	0.1289	1.021 3	—

Split, Secondary Contact, Symmetrical Gene Flow	-27.82	694.53	65.64	3.36	0.08049444	0.136 1	0.764 4	—	—	—	—	0.0206	—	0.929 4	0.296 7
Split with Ancient Symmetrical Migration	-28.03	968.58	66.06	3.78	0.06524753	0.096 1	0.441 4	—	—	—	—	0.0465	—	0.578	0.043 6
Split, Secondary Contact, Asymmetrical Gene Flow	-27.53	531.2	67.06	4.78	0.03957463	0.184 9	1.039 6	—	—	—	—	0.0333	0.4228	1.888	0.031 5
Split with Ancient Asymmetrical Migration	-28.01	1394.1 1	68.02	5.74	0.02448812	0.067 9	0.254 2	—	—	—	—	0.5243	1.2979	0.494 8	0.165 7
Split with No Migration, Size Change	-28.6	1117.5 2	69.2	6.92	0.01357443	—	—	0.571 9	0.0475	0.073 7	21.646 5	—	—	0.079	0.292 7
Split with Symmetric Migration, Size Change	-27.83	689.87	69.66	7.38	0.01078534	—	—	2.135 4	0.687	0.135 8	0.7703	0.0186	—	0.043	1.227 4
Split with Secondary Contact, Symmetrical Gene Flow, Size Change	-27.94	451.87	69.88	7.6	0.00966188	—	—	0.448 6	1.1292	0.205 5	1.3766	0.0125	—	1.498 2	0.949 2
Split with Ancient Symmetrical Migration, Size Change	-28.16	1140.1 4	70.32	8.04	0.00775384	—	—	0.259 1	0.9437	0.078 9	0.288	14.262 2	—	0.149 8	0.346 2
Split with Asymmetrical Migration, Size Change	-28.11	330.63	72.22	9.94	0.00299873	—	—	1.396 4	1.2271	0.289 2	1.5678	0.0106	0.0354	1.198 9	2.828 8
Split with Ancient Asymmetrical Migration, Size Change	-28.2	1076.9 5	72.4	10.1 2	0.00274063	—	—	1.914 2	0.7481	0.083 9	0.3244	0.5081	0.3151	0.045 2	0.380 4
Split with Secondary Contact, Asymmetrical Gene Flow, Size Change	-29.41	2110.3 7	74.82	12.5 4	0.00081725	—	—	1.239 1	2.3126	0.047 1	0.1174	0.0101	1.4782	0.424 8	3.628 3
No Divergence	-5039.1	757.23	10084. 2												

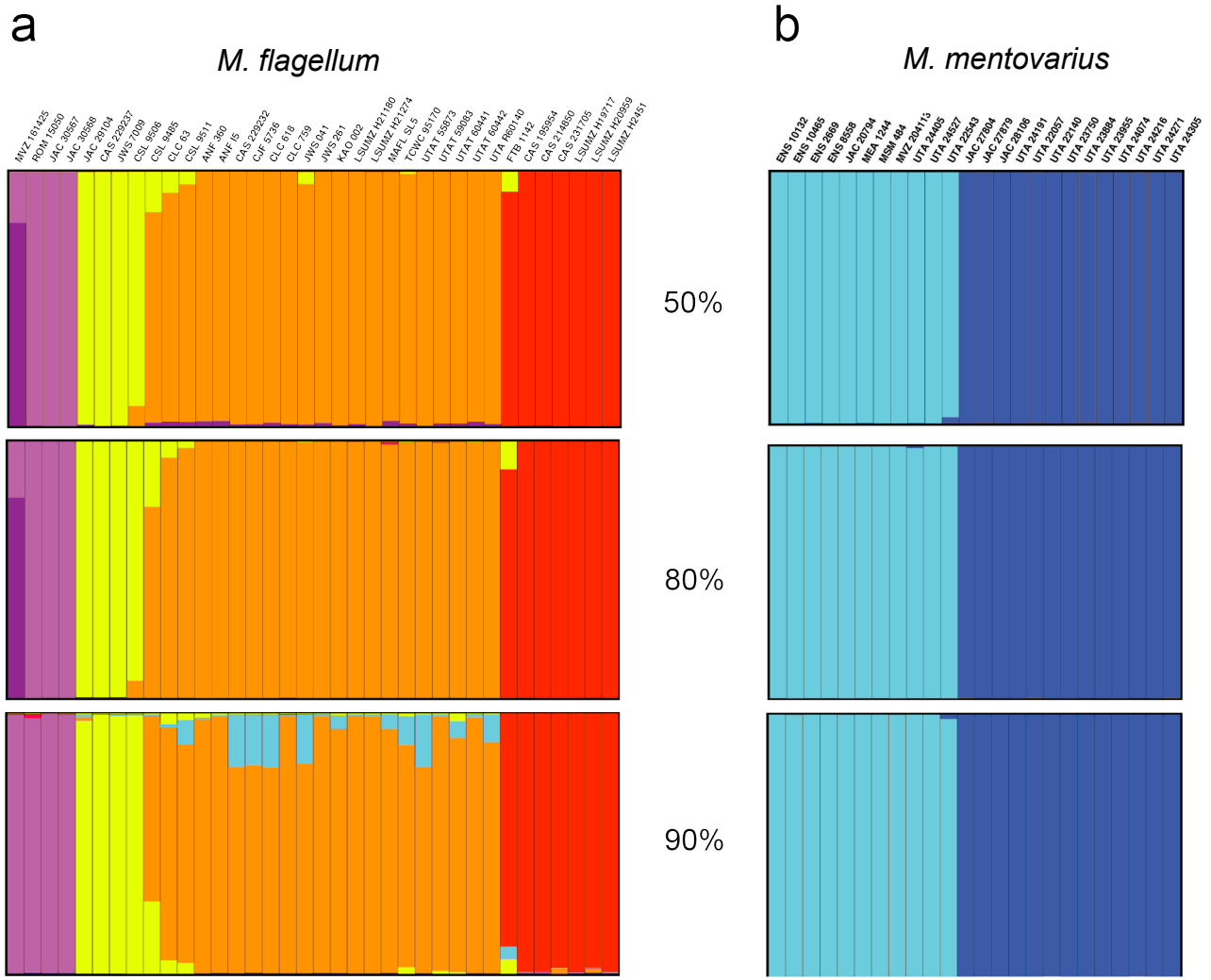

Fig. A2.1: Structure plots at three missing data thresholds for *M. flagellum* and *M. mentovarius*

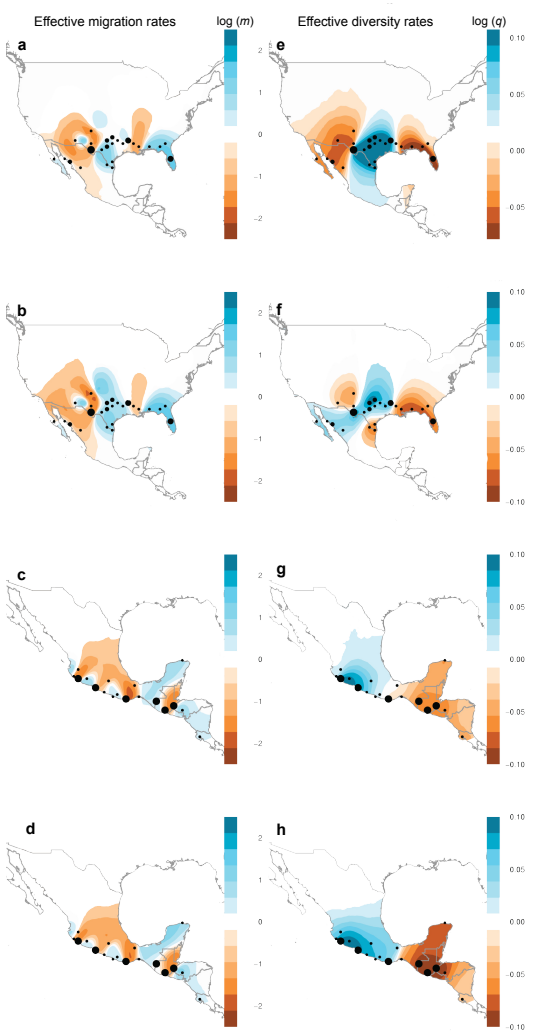

Fig. A2.2: EEMS Plots at various missing data thresholds for *M. flagellum* and *M. mentovarius*

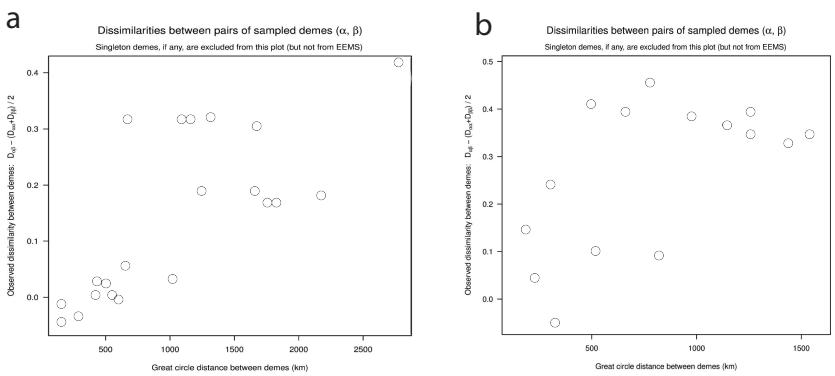

Fig. A2.3 EEMS outputs for A) *M. flagellum,* and B) *M. mentovarius*

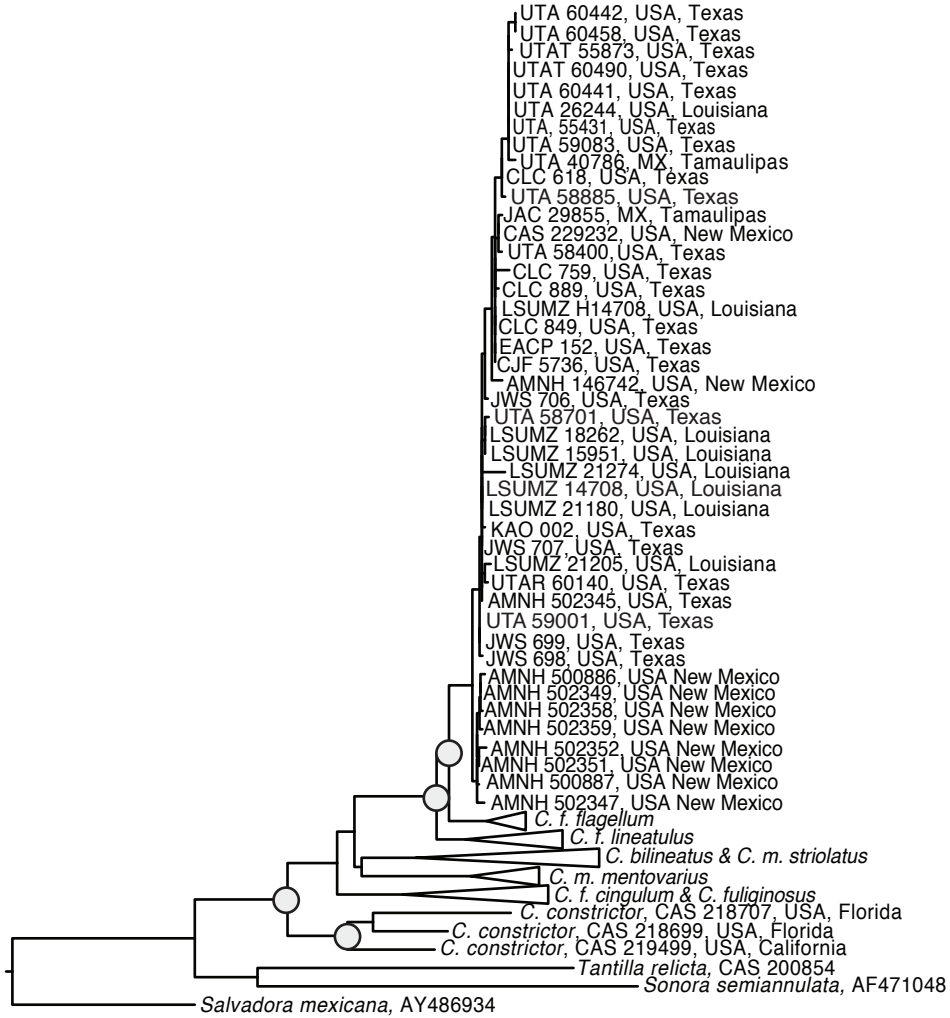

Fig. A3.1: Maximum likelihood phylogeny generated from mtDNA. The full clade representing *Coluber flagellum testaceus* is shown. All other clades are collapsed. Nodes with at least 70% bootstrap support are shown with grey circles.

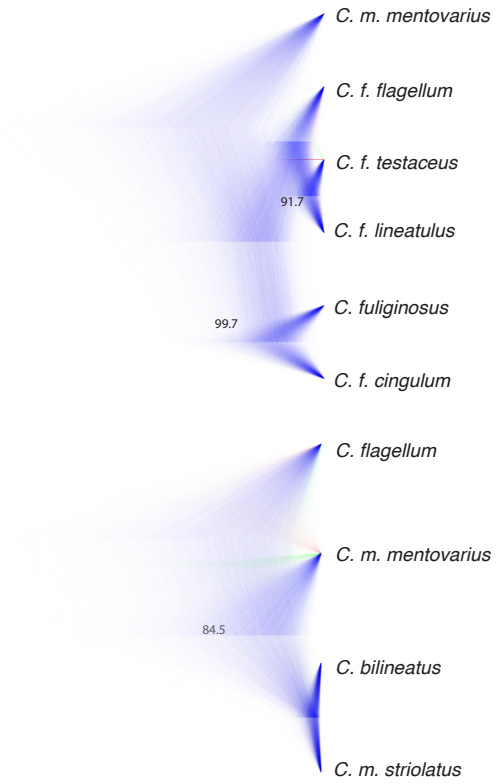

Fig. A3.2: Species trees generated using SNAPP based on the best-supported models from our Bayes Factor delimitation analysis shown in Fig. 3 for datasets C and E ($\leq$ 20% missing loci). Support values are labeled for each node that is not fully supported.

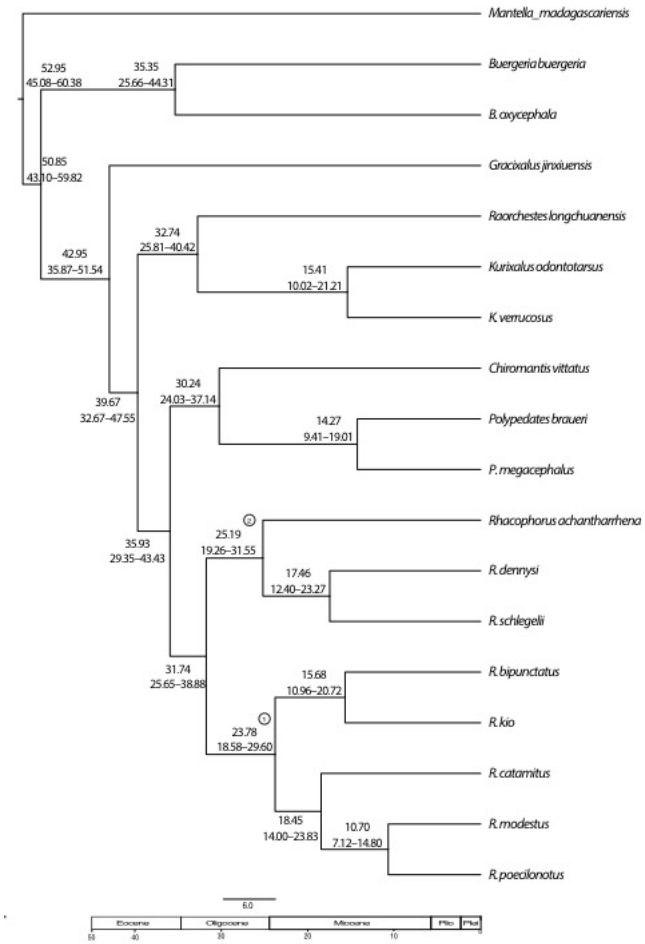

Fig. A5.1. Dated Bayesian phylogeny estimated using partial mitochondrial genomes for 18 species. Clades 1 and 2 are labeled within *Rhacophorus*. All nodes were fully resolved. Mean clade age and the 95% HDP are shown at each node.

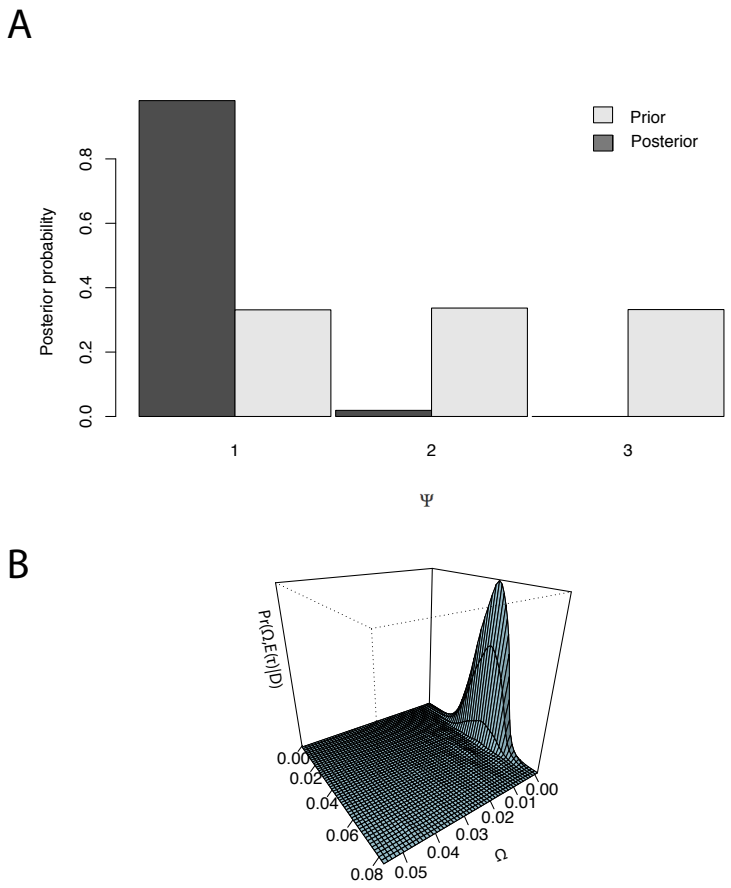

Fig. A5.2. Results of the msBayes sister-species pair analysis with *R. achantharrhena* and *R. prominanus* removed. A) Histogram of posterior probability of one synchronous divergence event. B) Joint posterior probability of the average divergence time and the variance in divergence times/average divergence time for the three sister-species pairs showing the highest probability of one divergence event.