Title

Spatiotemporal environmental variation and its effect on diversity and diversification at macro- and micro-evolutionary scales

Author

Wogan, Guinevere October

Publication Date

2011

Peer reviewed|Thesis/dissertation

Spatiotemporal environmental variation and its effect on diversity and diversification at macro- and micro-evolutionary scales

by

Guinevere October Uhler Wogan

A dissertation submitted in partial satisfaction of the

requirements of the for the degree of

Doctor of Philosophy

in

Integrative Biology

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Jimmy A. McGuire, Chair
Professor Craig C. Moritz
Professor Montgomery Slatkin
Professor Rosemary G. Gillespie

Fall 2011

Spatiotemporal environmental variation and its effect on diversity and diversification at macro- and micro-evolutionary scales

by Guinevere October Uhler Wogan

ABSTRACT

Spatiotemporal environmental variation and its effect on diversity and diversification at macro- and micro-evolutionary scales

by

Guinevere October Uhler Wogan

Doctor of Philosophy in Integrative Biology

University of California, Berkeley

Professor Jimmy A. McGuire, Chair

Species evolve in landscapes and environments that change through time. This spatial backdrop has profound effects on the diversification and merging of lineages. Populations become isolated by climate change, moving continents, and rising oceans, all of these factors can impact the evolutionary dynamics of lineages. While the spatial dimension of evolution has been a subject of interest since the inception of biogeography, the approaches to addressing how spatial heterogeneities affect lineages have advanced considerably in recent decades as genetic tools, computational tools, and new methodologies have allowed for explicit spatial hypotheses to be generated using GIS methods that can then be tested with genetic data.

The focus of this dissertation examines amphibian evolution at deep temporal scales in relation to a changing global landscape, and sets the stage for examining the spatial dimension across the Asian landscape at shallow temporal scales. Chapter one investigates amphibian evolution and diversification within an explicit spatiotemporal framework in order to understand how spatial variation drives evolutionary patterns of amphibians at global scales. At macroevolutionary scales, global amphibian diversity is strongly correlated with the area of forest rather than the longevity of forests, and signatures of the extensive forest area of the Eocene underlies the dramatic amphibian latitudinal diversity gradient.

Chapter two models the changing Asian landscape throughout the Pleistocene in order to understand how glacial and interglacial cycles impact the distribution of habitat types throughout the region. Satellite data on current habitats form the basis for this work. Palynological data are used to validate the models, providing a measure of confidence that the palaeo-predictive models are doing a good job in hind-casting habitats through changing climatic conditions. Asia differs from many other systems in not having been covered in extensive ice sheets during glacial periods, and provides a very different set of habitat dynamics, setting a unique stage for evolutionary dynamics. Habitats across Asia responded in dramatic fashion to changing palaeo-climates, with some habitats

undergoing massive expansions and others contracting. The past 140 thousand years have witnessed a highly dynamic landscape.

Chapter three delves deeper into palaeo-ecological modeling, and spatially locates stable habitat refugia across Asia for multiple tropical forest, temperate forest, and non-forested habitat types. These refugia are examined with respect to climatic stability and latitude to evaluate if stable climates give rise to stable habitats. Refugia are further examined with respect to terrestrial vertebrate species richness to understand if stable refugia have helped structure contemporary diversity patterns. Strong relationships between habitat stability and species richness were found, indicating that the spatial backdrop has played a pivotal role in contemporary diversity patterns. This indicates that these refugia may be important buffers against climate change and are probably important for the conservation of diversity in Asia.

Chapter four focuses in on glacial and interglacial habitat dynamics on the Sundashelf, where dramatic changes in sea-level affect the connectivity of landmasses providing potential migration corridors between mainland SE Asia and insular Asia. I examine how these changing sea levels impact the distribution of habitats across the region and how in turn these habitat changes structure genetic data across the area. I found that high levels of genetic diversity correspond with stable habitats.

Chapter five turns its attention to phylogenetics of the family Dicroglossidae, a widespread Old World frog family with high diversity across Asia. This chapter examines the generic assignments and familial monophyly and evaluates how nucleotide substitution models and partitioning affect the phylogenetic informativeness of markers and how this impacts phylogenetic inference. Both partitioning scheme and model selection were found to have profound effects on the phylogenetic informativenss of the genetic data. It was found that there are systemic biases among model selection criteria with the BIC commonly selecting highly parameterized models. It was also found that models of nucleotide substitution are often inadequate in capturing the complexity of empirical data. The monophyly of the Dicroglossidae was supported in these analyses. Many of the genera within the family were found to be monophyletic with the exceptions of the frogs of the Painii group and the *Fejervarya*.

Together, these chapters provide the spatial foundation which sets the stage for research evaluating diversification processes in Dicroglossid frogs across the Old World and within Asia at biogeographic, phylogeographic and population genetics time scales. These analyses provide the first deeper time model-based evaluation of landscape processes for Asia.

Dedicated to Lewis

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

When I started graduate school at UC Berkeley, I was sure that the central core of my research would revolve around species delimitation of Asian frogs. While this is still a component of my research (although not presented in this dissertation), my time here has broadened my interests significantly and has given me the tools to address a whole suite of other types of questions. It is the community at UC Berkeley and the Department of Integrative Biology, and in the Museum of Vertebrate Zoology that makes this possible. The exchange of ideas in seminars, classes, and over coffee make this one of the most vibrant scientific communities that I have had the good fortune to experience. It is the people that make it so.

Jim McGuire gave me breathing room to explore and engage in research independent of his own. That freedom has been invaluable in my development as a researcher. Despite the pressures that come with a career at a top-notch research university, Jim always made time to discuss ideas, and is a strong advocate for his students and their research. His enthusiasm and passion for exploration, herpetology, and phylogenetics is a continual source of inspiration. Thank you for giving me the opportunity to be part of the lab, it's been great!

Craig Moritz inspired much of the work I have undertaken in this dissertation. The approach he has implemented in his own research program has left an indelible mark on my perception of how truly integrative science is done. I am lucky to have had a chance to learn from him during his time at Berkeley. His role as both dissertation committee member and qualifying exam committee member both have left deep impressions on how I think about science.

Monty Slatkin has been a tremendous force in preparing me as an applied population geneticist. The very thoughtful approach he holds towards teaching very difficult topics sets a wonderful example of how to make even the most difficult topics approachable. My favorite question from Monty on my qualifying exams at the very end of a grueling few hours was "Why don't things disperse from islands very often?" Answer they are surrounded by water.

Rosemary Gillespie always has insightful comments and a smile. She has been a great role model and inspiration. Rosie manages not just excellent academic productivity, but also manages to engage in academic service and outreach. The GK12 program, which is funded through Rosie's efforts, was one of the most rewarding teaching experiences in which I had the opportunity to participate.

Although not part of my dissertation committee, Jeff Boore was the genomics master of my qualifying exams and did a great job guiding me from thinking about traditional sequencing approaches to thinking about genomics and their application to evolutionary questions. Other IB professors that have been particularly influential on my intellectual development through seminars, classrooms, co-authoring, or teaching are Tony Barnosky, Charles Marshall, David Ackerly, John Huelsenbeck, Kevin Padian, Marvalee Wake and David Wake. Tony Barnosky had the amazing idea to try to publish the ideas stemming from an extinction seminar -that resulted in our Nature paper. My interactions with both Tony and Charles have sparked a deep interest in integrating a palaeontological perspective into my work, and both have greatly influenced my work-

chapter one was in fact derived from a project I undertook as part of a seminar with Charles. David Ackerly's seminars on climate change and range changes and R analyses have continually introduced new approaches and views into my thinking. John Huelsenbeck and his 10-sided die taught me to think like a programmer. Kevin Padian and his in-depth knowledge of Darwin led me through two readings of Origin of Species, and an appreciation for the historical context of the ideas that form the basis for all modern evolutionary biology. David and Marvalee Wake both have a very thoughtful and in-depth knowledge of amphibians that has been inspiring. From ESPM I have enjoyed the many seminars I have taken with John Harte and Steve Beissinger.

The graduate student and postdoctoral community at UC Berkeley, and in particular in the MVZ is filled with incredible researchers and friends. I want to thank those that I was fortunate to overlap with on a daily basis for years, and whom I hope to have continued interactions for years to come. From the McGuire lab Adam Leache, Matt Fujita, Jon Fong, Matt Brandley, Tom Devitt, Shobi Lawalata-Dobbs, Sean Reilly, Dan Portik and most recently Sarah Hykin and Philip Skipworth provided a sounding board for ideas and great discussions. Postdocs Bryan Stuart, Chris Witt, Cori Richards, and Alison Davis-Rabosky all brought unique ideas to the table and greatly enriched our lab group. Sean Rovito, Juan Parra, Jay McEntee, Roberta Damasceno, Angela Ribiero, Ricardo Periera. Julie Woodruff, Sonal Singhal, Wly de Santos, have been great company, office mates, sounding boards, and inspirations. Thanks also to the MVZ staff Carol Spencer, Michelle Koo, Anna Ippolito, Lydia Smith, and Chris Conroy for making it easy to accomplish the work.

My work on Asian amphibians would not be possible without the folks from the herpetology department at California Academy of Sciences, Alan Leviton and Jens Vindum in particular. Jens fished my cv out of the trash and gave me a job in the herpetology department when I arrived in CA, he has been a major supporter from early on, and taught me how to be a good field biologist, and make beautiful specimens with careful data. Al took me in as a masters student and didn't complain when I moved on to my doctoral work before finishing. He has always made time for me, and has always advised me on how to incorporate a geological context into my research. With Al leading the way we have co-authored a book and several papers, I look forward to continuing our work on Asian snakes. George R. Zug and Jeffery Wilkinson have both been great colleagues along the way.

The MVZ and Integrative Biology funded much of my graduate research. Although most of my genetics work is not presented here, it would not have been possible without these pots of gold. The Kristina Louie fund of CAS also supported museum work and laboratory work. Fieldwork in Asia was supported by the California Academy of Sciences through grants to Alan Leviton, Jens Vindum and George Zug.

Lastly, none of this would have been possible without the support of Lewis, who was always encouraging and a voice of reason. Thanks for everything! Big time!

CHAPTER ONE

Examining the Causes of the Latitudinal Diversity Gradient in Amphibians

INTRODUCTION

Why are there so many species in the tropics as opposed to the temperate and boreal regions? The latitudinal diversity gradient is a very striking spatial pattern and has existed for the past 325 million years, becoming more pronounced throughout the past 60 million years (Mittelbach et al., 2007). In both terrestrial and marine systems, most taxonomic groups are more speciose at low latitudes (Hillebrand, 2004; Rosenzweig, 1995). Understanding the factors that promote and maintain diversity is a central line of investigation in explaining the latitudinal diversity gradient (henceforth LDG), and well over one-hundred different hypotheses have been proposed to explain the underlying processes (Mittelbach et al., 2007). However, the complex interaction of current environments, differing life-histories, and evolutionary dynamics, makes teasing apart the dynamics of the LDG difficult, especially when direct fossil inference is not possible.

Amphibians are an ancient species rich group that has been heavily impacted by recent extinction (Barnosky et al., 2011; Stuart et al., 2008; Stuart et al., 2004; Wake and Vredenburg, 2008), making them among the most imperiled of the vertebrate groups (Barnosky et al., 2011; Hoffmann et al., 2010; Stuart et al., 2008; Stuart et al., 2004; Wake and Vredenburg, 2008). The majority of extant amphibian species are found in forest habitats, with 82% percent of all known amphibian species inhabiting tropical forests (Stuart et al., 2004); this has translated into a particularly pronounced latitudinal diversity gradient with 91% of species in the tropical biomes, 8% in the temperate biomes, and less than 1% of species found in boreal biomes.

While a link between current amphibian diversity and forest habitats is apparent based on the above associations of tropical forests and high amphibian diversity, the nature of the relationship between the two is less understood. One possibility for example is that current environmental factors such as precipitation and temperature or available energy drive concordant patterns of amphibian and tropical plant diversity (Currie et al., 2004). Separate analyses of contemporary amphibian diversity and contemporary plant diversity suggest that all three of these climatic factors are important indicators structuring global diversity patterns (Buckley and Jetz, 2007; Kreft and Jetz, 2007). However, Qian and Ricklefs (2008) demonstrated a positive association between vascular plant richness and amphibian species richness even after controlling for contemporary environmental parameters such as climate, topography, sample area, and geographic region, suggesting that the concordance is more than coincidence. They further suggest that concordance arises due to large-scale ecological and evolutionary processes acting similarly on multiple images (Qian and Ricklefs, 2008).

Global analyses of amphibians suggest that historical factors play a dramatic role in shaping contemporary diversity patterns (Buckley and Jetz, 2007; Qian and Ricklefs, 2008), and the fossil record has revealed a link between the evolutionary dynamics of amphibians and rainforests. Sahney et al. (2010) found that global aridification led to the fragmentation and subsequent collapse of tropical rainforest systems during the

Carboniferous (305 mya) which in turn led to the extinction of a disproportionate number of early amphibian lineages.

Given the contemporary and deep time association between amphibians and forests, I tested how forest occurrence has structured modern forest-dwelling amphibian diversity by analyzing the relationship between contemporary amphibian diversity and forest extent over the past 60 million years. I tested two competing hypotheses relating forest to amphibian diversity: (i) forest area alone correlates with diversity (akin to a species-area relationship explanation), (ii) forest age and area correlates with diversity (suggesting both the longevity and area of the forest are important) (Fine and Ree, 2006).

Both of these hypotheses can be lumped under "time and area" category of latitudinal gradient hypotheses (Mittelbach et al., 2007), but are further dissected here to distinguish between forest area and forest area combined with forest age as primary factors underlying the LDG in amphibians. The general relationship between habitat area diversity has been well enumerated since MacArthur and Wilson's seminal work on island biogeography (MacArthur and Wilson, 1967), but area hypotheses as general explanations of the LDG date back to Darwin (1859) and Wallace (1878). Area hypotheses posit that the large area of the tropics has enabled increased speciation or decreased extinction (reviewed in Mittelbach et al. 2007). The longevity/stability of tropical systems has been hypothesized to be an important contributing factor to the LDG since longevity allows for greater species accumulation (Stebbins, 1974). High latitude regions have undergone much more perturbation through time due to the formation of continental ice sheets and orbital influence such as Milankovich cycles (Jansson and Dynesius, 2002), which has meant that extinction-colonization dynamics likely dominate (Marshall, 2007). The relative roles of abiotic and biotic interactions in structuring diversity and diversification has been a topic of recent interest, and is an area of exciting investigation on the LDG (Jablonski, 2008; Marshall, 2007; Rabosky, 2009; Schemske et al., 2009).

A third hypothesis, dubbed the "time to speciation" hypothesis posits that time, specifically time since colonization, drives diversity patterns (Stephens and Wiens, 2003). This idea has already been extensively tested in amphibians (Hua and Wiens, 2010; Kozack and Wiens 2006; Kozak and Wiens 2007; Smith et al., 2006; Wiens, 2007; Wiens et al., 2006), and invalidated as an explanation for global amphibian diversity patterns (Wiens, 2007). However, for some individual lineages "time to speciation" effects do explain diversity patterns (Wiens et al., 2006).

METHODS

Estimates of forest area through time were obtained from Fine and Ree (2006) which were compiled from (Beerling and Woodward, 2001; Scotese, 2003; Willis and McElwain, 2002). Six different time slices were assessed: Eocene, Oligocene, Miocene, Last Glacial Maximum, Mid-Holocene, and the present. Forests were grouped into three major biomes boreal; (encompassing boreal and sub-arctic forest), temperate (encompassing temperate and sub-antarctic forest), and tropical (encompassing moist and

dry lowland and montane tropical and subtropical forest). Forest areas were estimated for the Afrotropics, Neotropics, Asian tropics, Eastern and Western North American temperate, South American temperate, Asian temperate, European temperate, Australian temperate, North American boreal, and Eurasian boreal forests.

Three different area calculations were utilized for statistical tests. First, the log of forest area was used. The second and third approach were inspired by or directly follow methods in Fine and Ree (2006). Both a discretized area parameter, and an area-time composite parameter were derived for each of the above forests using Image J software following methods outlined in Fine and Ree (2006). Bilinear interpolation was utilized to estimate missing forest area values for particular time points, thereby maintaining the integrity of each forest model. The discrete area parameter differs from the area-time composite parameter in that it treats the area at each time period as a separate value, while the area-time parameter adds between times thus allowing the combined effect of time and area to be evaluated (Fine and Ree, 2006).

Amphibian diversity counts were obtained from the Global Amphibian Assessment (Stuart et al., 2004) now integrated into the IUCN Red List (IUCN, 2011). All forest dwelling amphibians were extracted and then subdivided into the appropriate continent and forest type to accord with the eleven forest areas.

The log of area, log of discretized area, or log of area-time composite parameter and the log of amphibian diversity were used for statistical analyses. Analyses were carried out for all amphibians combined, and for each order separately. Pearson correlations and linear regressions consisted of the forest area estimated for each of the eleven forest types at a single time and the contemporary amphibian species diversity in each of those eleven forest types. Regressions treated forest area as the independent factor and amphibian diversity as the dependent factor. AIC scores were calculated to compare regression models across area and area time models. 216 models were evaluated for each of four separate analyses: all amphibians, anurans, salamanders, and caecilians.

RESULTS

Contemporary amphibian species diversity is largely concentrated in the tropics, with extra-tropical centers of diversity in southeast Africa, North America, and eastern Australia (Figure 1). Biome areas have changed dramatically since the Eocene. Tropical and temperate forests have been dramatically reduced in area and their locations have shifted. The emergence of the boreal forest biome during the Miocene, and its expansion since then, marks the transition to a very different global forest structure (Figure 2). Although individual biome reconstructions differ in their estimates of forest extent, they are in general agreement.

Pearson correlations did not find any significant relationship between contemporary forest area and amphibian diversity, this is also true for the mid-Holocene time period (Table 1). Pearson correlations did however reveal a tight link between global amphibian diversity and forest area, which was particularly strong during the Eocene and Oligocene. Several of the models also support a strong relationship for the Miocene. A

single statistical model finds evidence for a relationship between these factors during the LGM (Table 1). Regression analyses reveal that contemporary amphibian diversity is not explained by contemporary or recent forest occurrence, however amphibian diversity is explained by forest occurrence during the Eocene, Oligocene, and Miocene in most reconstructions. AIC model comparisons overwhelmingly favored forest area rather than time-integrated forest area, thus global amphibian diversity is best explained by forest area rather than time-integrated area.

Evaluation of correlation patterns at the Ordinal level reveals that anurans are largely driving these findings (Table 2). Anuran diversity and forest occurrence are strongly correlated during the Eocene and Oligocene in all palaeo-reconstructions in all models. Support for a Miocene correlation also exists in some of the models. Salamander diversity is correlated with forest occurrence in only one of the palaeo-environmental reconstructions and only during the Eocene and Oligocene. A correlation between caecilian diversity and forest occurrence is recovered in one palaeo-environmental reconstruction during the Miocene and LGM time periods (Table 2). There are no correlations among any of the amphibian orders and mid-Holocene or contemporary forest occurrences.

DISCUSSION

The early Tertiary was warmer and wetter than today, and both temperate and tropical forests of that time are thought to have been greater in extent than now (Figure 2) (Beerling and Woodward, 2001; Fine and Ree, 2006; Morley, 2000; Scotese, 2003; Willis and McElwain, 2002). Rainforests underwent diversification and expansion during the greenhouse conditions of the Cretaceous to the mid-Eocene reaching their maximal extent in the Eocene, then receding when icehouse conditions set in (Morley, 2000). This suggests that rainforest expansion promoted diversification of amphibians, and supports the hypothesis that modern amphibians co-diversified with the angiosperms (Roelants et al., 2007). During this time several major amphibian lineages diversified, notably hyloid and ranoid frogs (Roelants et al., 2007), and Plethodontid salamanders (Wiens et al., 2007), driving elevated diversification estimates among amphibians (Roelants et al., 2007), the signatures of which are still evident in modern diversity patterns.

Past rather than present forest environments better explain amphibian species richness patterns, which could imply the presence of a deep-time extinction debt (Kuussaari et al., 2009; Lister and Stuart, 2008; Looy et al., 2001). Extinction debt can be thought of as a non-equilibrium state following a habitat perturbation, in this case the global restructuring of forests. However, this time lag lies on the order of tens of millions of years. Can extinction debts persist so long? Looy et al. (2001) found evidence for end Permian extinction debts persisting 25 my after ecosystem collapse. However, Jablonski (2002) posits that extinction debt should persist on shorter time scales than observed in the amphibian case.

Alternatively, this finding could also be explained by ecological interactions such as highly efficient species packing. Species packing could arise due to functional

equivalency (Hubbell, 2001) or spatial competition (Tilman, 1994). Sorting among these alternative explanations will require estimates of long-term clade dynamics. While palaeo-diversity estimates from fossils would be ideal, amphibians have a fragmentary fossil record, especially those from tropical wet forests, so elucidating their diversification is particularly challenging. To better understand the mechanisms underlying the LDG in amphibians, analyses evaluating the relative roles of speciation, extinction, and dispersal within and between biomes, respectively, would help illuminate the findings presented here. So too would analyses to distinguish between ecological interactions and non-equilibrium dynamics.

From a conservation standpoint, the association of modern amphibian diversity with Tertiary tropical and temperate forest maxima speaks to the need to reduce destruction of these habitats as a primary step to conserving amphibians. Current global forest cover is estimated at 32,688,000 km^2 (Hansen et al., 2010), while most reconstructions of Tertiary forest cover suggest two to three fold greater coverage at that time. Contemporary forest biomes are roughly 1/3 each boreal, temperate, and tropical (Olson et al., 2001), while during the Tertiary, the boreal forest biome did not exist, and tropical forests comprised 56- 71% of the forest biome (Beerling and Woodward, 2001; Scotese, 2003; Willis and McElwain, 2002) (Figure 2). This drastic difference in the global structure of forest biomes, e.g. the expansion of the amphibian depauperate boreal biome and the shrinking of the amphibian rich tropical biome, and the overall global reduction of forest cover, means first that amphibians have already gone through a major perturbation, and second that further deforestation could hasten extinction. Habitat loss is the biggest threat to amphibian diversity (Stuart et al., 2004), and deforestation continues to be a major conservation issue across the planet with 1,011,000 km^2 of forest lost between 2000-05 (Hansen et al., 2010). Synergies among deforestation, climate change, disease, and introduced species have set up a system of feedbacks that are particularly destructive to amphibian diversity.

In contemporary faunas, habitat stability in recent geological time is an important predictor of diversity, particularly of low dispersal organisms such as amphibians (Carnaval et al., 2008; Carnaval and Moritz, 2008; Graham et al., 2006; Sandel et al., 2011; Wogan, chapter 3). The analyses here reveal that forest area is more important than forest area plus longevity in structuring amphibian diversity patterns over deep temporal scales, perhaps demonstrating that the relative importance of different environmental factors varies over differing temporal scales. However, forest longevity and forest stability, though related are not identical factors. It would be of interest to evaluate diversity within forest patches that are predicted to have remained relatively stable through time at these deeper time scales to distinguish between the longevity of a broadly defined forest region (e.g. Asian tropics) and the spatio-temporal stability of particular patches within those broadly defined classifications. For example, while the Asian tropics are hyper-diverse and have a continual existence throughout the temporal scale evaluated here, only a few locations (e.g. Borneo, Peninsular Malaysia) are predicted to have remained stable. Identification of "deep-time" forest refugia may add another important facet to our understanding of global amphibian diversity patterns.

LITERATURE CITED

Barnosky, A. D., N. Matzke, S. Tomiya, G. O. U. Wogan, B. Swartz, T. Quental, C. R. Marshall, J. L. McGuire, B. Mersey, K. C. Maguire, and E. A. Ferrer. 2011. Has the Earth's sixth mass extinction already arrived? Nature 471:51-57.

Beerling, D. J., and F. I. Woodward. 2001. Vegetation and the Terrestrial Carbon Cycle: Modelling the first 400 million years. Cambridge University Press, Cambridge.

Buckley, L. B., and W. Jetz. 2007. Environmental and historical constraints on global patterns of amphibian richness. Proc. Biol. Sci. 274:1167-73.

Carnaval, A. C., M. Hickerson, C. Haddad, M. T. Rodrigues, and C. Moritz. 2008. Stability predicts genetic diversity in the Brazialian Atlantic Forest Hotspot. Science 323:785-789.

Carnaval, A. C., and C. Moritz. 2008. Historical climate modelling predicts patterns of current biodiversity in the Brazilian Atlantic forest. J Biogeography 35:1187-1201.

Currie, D. J., G. G. Mittelbach, H. V. Cornell, R. Field, J.-F. Guegan, B. A. Hawkins, D. M. Kaufman, J. T. Kerr, T. Oberdorff, E. O'brien, and J. R. G. Turner. 2004. Predictions and tests of climate-based hypotheses of broad-scale variation in taxonomic richness. Ecol Letters 7:1121-1134.

Darwin, C. 1859. On the origin of species by means of natural selection, or the preservation of favoured races in the struggle for life, first edition. John Murray, London.

Fine, P., and R. Ree. 2006. Evidence for a time-integrated species-area effect on the latitudinal gradient in tree diversity. Am. Nat. 168:796-804.

Graham, C., C. Moritz, and S. Williams. 2006. Habitat history improves prediction of biodiversity in rainforest fauna. Proc. Natl. Acad. Sci. U.S.A. 103:632-6.

Hansen, M. C., S. V. Stehman, and P. V. Potapov. 2010. Quantification of global gross forest cover loss. Proceedings of the National Academy of Sciences 107:8650-8655.

Hillebrand, H. 2004. On the generality of the latitudinal diversity gradient. American Naturalist 163:192-211.

Hoffmann, M., C. Hilton-Taylor, A. Angulo, M. Bohm, T. M. Brooks, S. H. M. Butchart, K. E. Carpenter, J. Chanson, B. Collen, N. A. Cox, W. R. T. Darwall, N. K. Dulvy, L. R. Harrison, V. Katariya, C. M. Pollock, S. Quader, N. I. Richman, A. S. L. Rodrigues, M. F. Tognelli, J.-C. Vie, J. M. Aguiar, D. J. Allen, G. R. Allen, G. Amori, N. B. Ananjeva, F. Andreone, P. Andrew, A. L. A. Ortiz, J. E. M. Baillie, R. Baldi, B. D. Bell, S. D. Biju, J. P. Bird, P. Black-Decima, J. J. Blanc,

F. Bolanos, W. Bolivar-G, I. J. Burfield, J. A. Burton, D. R. Capper, F. Castro, G. Catullo, R. D. Cavanagh, A. Channing, N. L. Chao, A. M. Chenery, F. Chiozza, V. Clausnitzer, N. J. Collar, L. C. Collett, B. B. Collette, C. F. C. Fernandez, M. T. Craig, M. J. Crosby, N. Cumberlidge, A. Cuttelod, A. E. Derocher, A. C. Diesmos, J. S. Donaldson, J. W. Duckworth, G. Dutson, S. K. Dutta, R. H. Emslie, A. Farjon, S. Fowler, J. Freyhof, D. L. Garshelis, J. Gerlach, D. J. Gower, T. D. Grant, G. A. Hammerson, R. B. Harris, L. R. Heaney, S. B. Hedges, J.-M. Hero, B. Hughes, S. A. Hussain, J. Icochea M, R. F. Inger, N. Ishii, D. T. Iskandar, R. K. B. Jenkins, Y. Kaneko, M. Kottelat, K. M. Kovacs, S. L. Kuzmin, E. La Marca, J. F. Lamoreux, M. W. N. Lau, E. O. Lavilla, K. Leus, R. L. Lewison, G. Lichtenstein, S. R. Livingstone, V. Lukoschek, D. P. Mallon, P. J. K. Mcgowan, A. Mcivor, P. D. Moehlman, S. Molur, et al. 2010. The Impact of Conservation on the Status of the World's Vertebrates. Science 330:1503-1509.

Hua, X., and J. J. Wiens. 2010. Latitudinal variation in speciation mechanisms in frogs. Evolution 64:429-443.

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

IUCN. 2011. International Union for Conservation of Nature Red List IUCN.

Jablonski, D. 2002. Survival without recovery after mass extinction. Proceedings of the National Academy of Science 99:8139-8144.

Jablonski, D. 2008. Biotic interactions and macroevolution: extensions and mismatches across scales and levels. Evolution 62:715-739.

Jansson, R., and M. Dynesius. 2002. The fate of clades in a world of recurrent climatic change: Milankovitch oscillations and evolution. Annual Reviews in Ecology and Systematics 33:741-777.

Kozack, K., and J. J. Wiens 2006. Does niche conservatism promote speciation? A case study in North American salamanders. Evolution 60:2604-2621.

Kozak, K. H., and J. J. Wiens 2007. Climatic zonation drives latitudinal variation in speciation mechanisms. Proc. Biol. Sci. 274:2995-3003.

Kreft, H., and W. Jetz. 2007. Global patterns and determinants of vascular plant diversity. P Natl Acad Sci Usa 104:5925-5930.

Kuussaari, M., R. Bommarco, R. K. Heikkinen, A. Helm, J. Krauss, R. Lindborg, E. Ockinger, M. Pärtel, J. Pino, F. Rodà, C. Stefanescu, T. Teder, M. Zobel, and I. Steffan-Dewenter. 2009. Extinction debt: a challenge for biodiversity conservation. Trends in Ecology & Evolution 24:564-71.

Lister, A. M., and A. J. Stuart. 2008. The impact of climate change on large mammal distribution and extinction: evidence from the last glacial/interglacial transition. Compte Rendue Geoscience 340:615-620.

Looy, C. V., R. J. Twitchett, D. L. Dilcher, J. H. A. Van Konijnenburg-Van Cittert, and H. Visscher. 2001. Life in the end-Permian dead zone. Proceedings of the National Academy of Sciences 98:7879-7883.

MacArthur, R. H., and E. O. Wilson. 1967. The theory of Island Biogeography, 2001 edition. Princeton University Press, Princeton.

Marshall, C. R. 2007. Explaining latitudinal diversity gradients. Science 317:451-3; author reply 451-3.

Mittelbach, G., D. Schemske, H. Cornell, A. Allen, J. Brown, M. Bush, S. Harrison, A. Hurlbert, N. Knowlton, H. Lessios, C. McCain, A. McCune, L. McDade, M. McPeek, T. Near, T. Price, R. Ricklefs, K. Roy, D. Sax, D. Schluter, J. Sobel, and M. Turelli. 2007. Evolution and the latitudinal diversity gradient: speciation, extinction and biogeography. Ecol. Lett. 10:315-31.

Morley, R. J. 2000. Origin and evolution of tropical rain forests. John Wiley & Sons, LTD, Chichester.

Olson, D., E. Dinerstein, E. Wikramanayake, N. Burgess, G. Powell, E. Underwood, J. D'amico, I. Itoua, H. Strand, and J. Morrison. 2001. Terrestrial ecoregions of the world: a new map of life on earth. BioScience 51:933-938.

Qian, H., and R. E. Ricklefs. 2008. Global concordance in diversity patterns of vascular plants and terrestrial vertebrates. Ecol Letters 11:547-553.

Rabosky, D. L. 2009. Ecological limits and diversification rate: alternative paradigms to explain the variation in species richness among clades and regions. Ecol. Lett. 12:735-743.

Roelants, K., D. Gower, M. Wilkinson, S. Loader, S. Biju, K. Guillaume, L. Moriau, and F. Bossuyt. 2007. Global patterns of diversification in the history of modern amphibians. Proc. Natl. Acad. Sci. U.S.A. 104:887-92.

Rosenzweig, M. L. 1995. Species diversity in space and time. Cambridge University Press, Cambridge.

Sahney, S., M. J. Benton, and H. J. Falcon-Lang. 2010. Rainforest collapse triggered Carboniferous tetrapod diversification in Euramerica. Geology 38:1079-1082.

Sandel, B., L. Arge, B. Dalsgaard, R. G. Davies, K. J. Gaston, W. J. Sutherland, and J.-C. Svenning. 2011. The influence of Late Quaternary climate-change velocity on species endemism. Science early online October 6, 2011:6.

Schemske, D. W., G. G. Mittelbach, H. V. Cornell, J. M. Sobel, and K. Roy. 2009. Is there a Latitudinal Gradient in the Importance of Biotic Interactions? Annual Review Ecology Evolution Systematics 40:245-269.

Scotese, C. R. 2003. PALEOMAP project http://www.scotese.com.

Smith, S., P. Stephens, and J. Wiens. 2006. Replicate patterns of species richness, historical biogeography, and phylogeny in Holarctic treefrogs. Evolution 59:2433-50.

Stebbins, G. L. 1974. Flowering plants: evolution above the species level. The Belknap Press of Harvard University Press, Cambridge.

Stephens, P., and J. Wiens. 2003. Explaining species richness from continents to communities: the time-for-speciation effect in emydid turtles. Am. Nat. 161:112-28.

Stuart, S., M. Hoffman, J. Chanson, N. Cox, R. Berridge, P. Ramani, and B. Young. 2008. Threatened amphibians of the world. Lynx Edicions, Barcelona.

Stuart, S. N., J. S. Chanson, N. A. Cox, B. E. Young, A. S. L. Rodrigues, D. L. Fischman, and R. W. Waller. 2004. Status and trends of amphibian declines and extinctions worldwide. Science 306:1783-6.

Tilman, D. 1994. Competition and Biodiversity in Spatially Structured Habitats. Ecology 75:2-16.

Wake, D. B., and V. Vredenburg. 2008. Are we in the midst of the sixth mass extinction? A view from the world of amphibians. Proceedings of the National Academy of Sciences 105:11466.

Wallace, A. R. 1878. Tropical nature, and other essays. R. Clay, Sons, and Taylor, London.

Wiens, J. 2007. Global patterns of diversification and species richness in amphibians. Am. Nat. 170 Suppl 2:S86-106.

Wiens, J., C. Graham, D. Moen, S. Smith, and T. Reeder. 2006. Evolutionary and ecological causes of the latitudinal diversity gradient in hylid frogs: treefrog trees unearth the roots of high tropical diversity. The American naturalist 168:579-96 L3 - papers://FD10C007-F381-43DE-81EE-3D9B213DD40F/Paper/p1630.

Wiens, J., G. Parra-Olea, M. García-París, and D. Wake. 2007. Phylogenetic history underlies elevational biodiversity patterns in tropical salamanders. Proc. Biol. Sci. 274:919-28.

Wiens , J. J., C. H. Graham, D. S. Moen, S. A. Smith, and T. W. Reeder. 2006. Evolutionary and ecological causes of the latitudinal diversity gradient in hylid

frogs: treefrog trees unearth the roots of high tropical diversity. Am. Nat. 168:579-96.

Willis, K. J., and J. C. McElwain. 2002. The evolution of plants. Oxford University Press, Oxford.

Wogan, G. 2011. Chapter 3. Palaeo-vegetation reconstruction and habitat stability analyses pinpoint potential Pleistocene refugia in Asia. Pages 41 *in* Integrative Biology University of California, Berkeley, Berkeley.

Figure 1. Global contemporary amphibian diversity. Species richness was calculated utilizing IUCN Red List spatial distribution data (2010). Species richness is scaled from high diversity (red) to low diversity (blue).

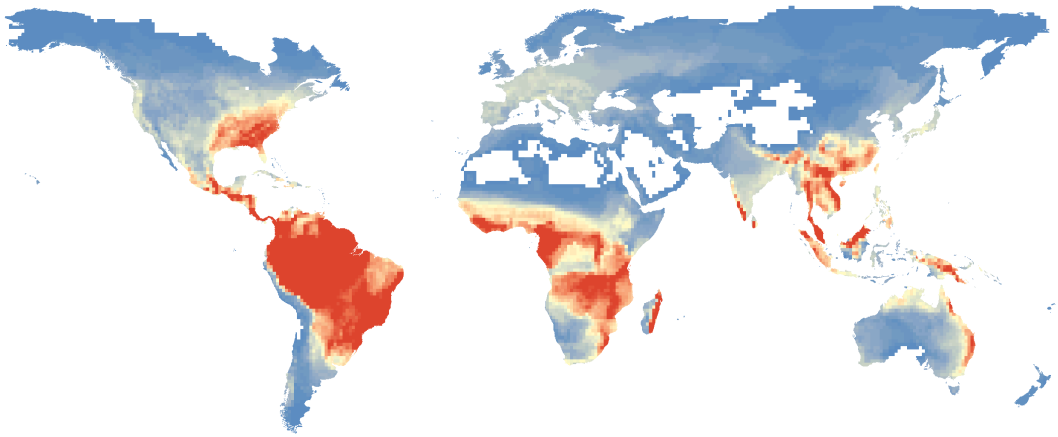


Figure 2. The global distribution of three major forest biomes throughout the Cenozoic to now. Estimates of biome area and location come from palaeontological reconstructions utilizing differing proxies. The pie charts below each globe depict the relative proportion of each of the three biomes at each time slice as reconstructed from differing sources. The last panel again depicts contemporary amphibian species diversity and the pie chart depicts the relative proportion of diversity within each of the three forest biomes. Figure adapted from Fine and Ree 2007.

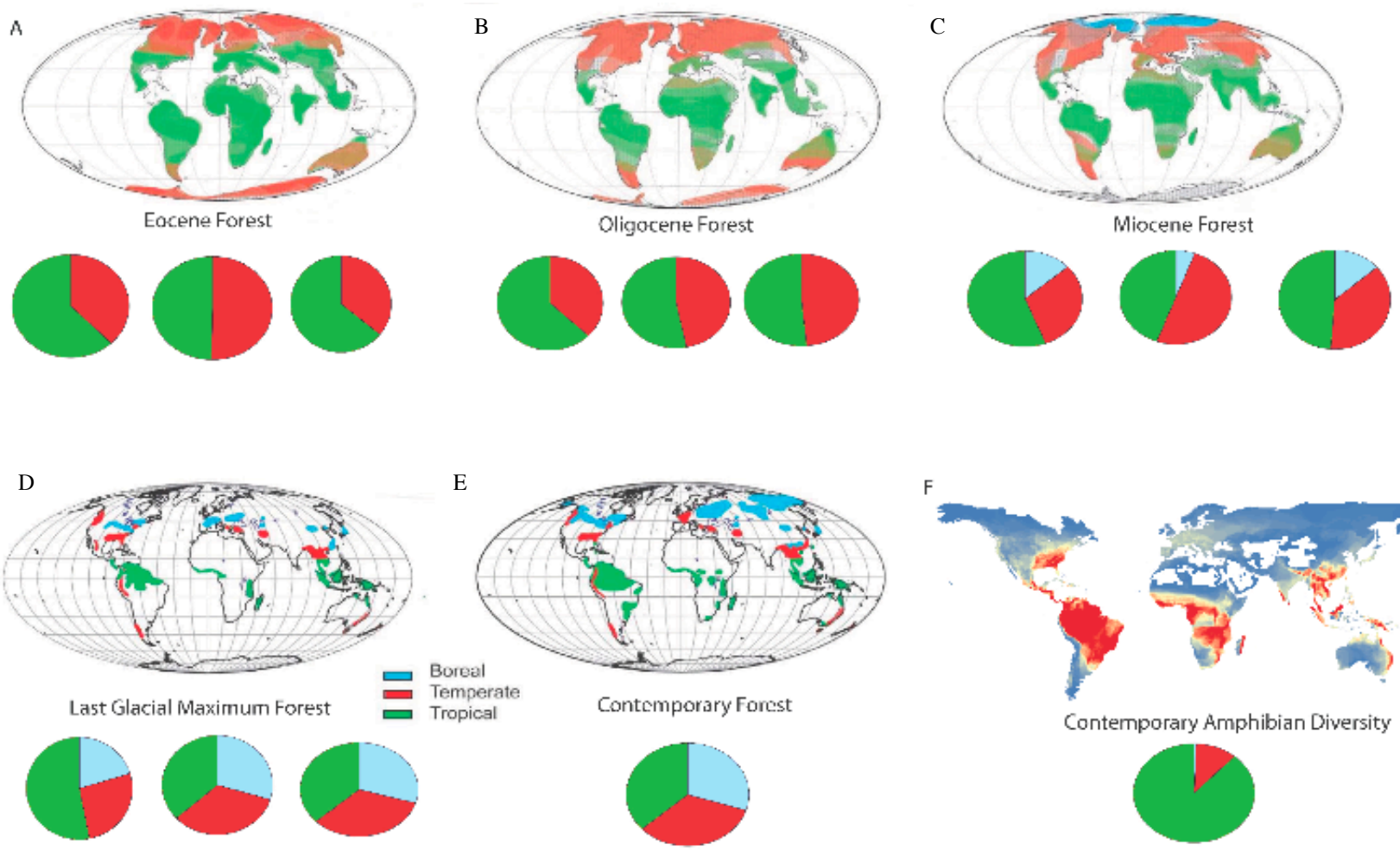

A
Eocene Forest
B
Oligocene Forest
C
Miocene Forest
D
Last Glacial Maximum Forest
Boreal
Temperate
Tropical
E
Contemporary Forest
F
Contemporary Amphibian Diversity

Table 1. Statistical analyses examining the relationship between amphibian diversity and forest area, discretized forest area, and forest area plus longevity.

Table 1.

		Log Area				Discrete Area Log				Cumulative Time Area Log			
		All Amphibians			AIC	All Amphibians			AIC	All Amphibians			AIC
Model Source	Epoch	p-value	R2	adjusted R2	model score	p-value	R2	adjusted R2	model score	p-value	R2	adjusted R2	model score
Beerling and Woodward	Eocene	**0.0009***	0.7241	0.6934	15.7567	**0.0015***	0.6889	0.6574	16.9774	**0.0011***	0.7056	0.6234	**12.3746**
(2001)	Oligocene	**0.0010***	0.7169	0.6854	16.0397	**0.0013***	0.7056	0.6666	16.6788	**0.0054***	0.6929	0.6490	**11.7393**
	Miocene	**0.0223**	0.4575	0.3972	23.1942	0.0579	0.3481	0.2710	25.2846	0.2447	0.1444	0.0521	28.1736
	LGM	0.0801	0.3016	0.2240	25.9720	0.1828	0.1849	0.0978	27.6303	0.3469	0.0961	-0.0015	28.7783
	Mid-Holocene	0.4523	0.0642	0.0640	28.0347	0.2267	0.1600	0.0640	28.0347	0.4726	0.0576	-0.0458	29.2547
	Now	0.4263	0.0716	-0.0316	29.1038	0.3368	0.1024	0.0029	28.7298	0.3368	0.1024	0.0029	28.7298
Scotese (2003)	Eocene	**0.0074***	0.5680	0.5199	**20.6896**	**0.0079***	0.5625	0.5129	20.8508	**0.0102**	0.5329	0.4865	21.4304
	Oligocene	**0.0112**	0.5292	0.4769	21.6345	**0.0097***	0.5476	0.4921	**21.3102**	**0.0101**	0.5329	0.4874	21.4113
	Miocene	**0.0027***	0.6497	0.6108	18.3817	**0.0016***	0.6889	0.6507	**17.1928**	**0.0241**	0.4489	0.3874	23.3712
	LGM	0.4253	0.0719	-0.0312	29.1001	**0.0144**	0.5041	0.4490	**22.2049**	0.1049	0.2704	0.1837	26.5296
	Mid-Holocene	0.4261	0.0717	-0.0315	29.1028	0.2026	0.1764	0.0816	27.8254	0.2028	0.1764	0.0815	27.8274
	Now	0.4263	0.0716	-0.0316	29.1038	0.2030	0.1764	0.0813	27.8293	0.2030	0.1764	0.0813	27.8293
Willis and McElwain	Eocene	**0.0018***	0.6786	0.6429	**17.4339**	**0.0022***	0.6724	0.6273	17.9058	**0.0064***	0.5776	0.5344	20.3536
(2002)	Oligocene	**0.0094***	0.5460	0.4956	21.2336	**0.0049***	0.6084	0.5587	**19.7626**	**0.0078***	0.5625	0.5149	20.8046
	Miocene	**0.0473**	0.3693	0.2992	24.8508	**0.0005***	0.7569	0.7251	**14.5597**	**0.0369**	0.3969	0.3329	24.3084
	LGM	0.4202	0.0734	-0.0295	29.0820	0.0521	0.3600	0.2861	25.0556	0.1419	0.2209	0.1373	27.1369
	Mid-Holocene	0.4252	0.0719	-0.0312	29.0997	0.2601	0.1369	0.0425	28.2840	0.2652	0.1369	0.0395	28.3182
	Now	0.4263	0.0716	-0.0316	29.1038	0.2710	0.1296	0.0362	28.3569	0.2710	0.1296	0.0362	28.3569

Table 2. Statistical results examining the relationship between the three major amphibian groups and forest occurance.

Table 2.

		Log Area								Discrete Time/Area Log								Cumulative Time/Area Log							
		All Amphibians		Anurans		Salamanders		Caecilians		All Amphibians		Anurans		Salamanders		Caecilians		All Amphibians		Anurans		Salamanders		Caecilians	
Model Source	Epoch	p-value	R2	p-value	R2	p-value	R2	p-value	R2	p-value	R2	p-value	R2	p-value	R2	p-value	R2	p-value	R2	p-value	R2	p-value	R2	p-value	R2
Beerling and Woodward	Eocene	**0.0009***	0.82	**0.0134**	0.71	**0.0148**	0.71	0.1426	0.47	**0.0015***	0.83	**0.0107**	0.73	**0.0196**	0.69	0.1223	0.49	**0.0011***	0.84	**0.0044***	0.78	0.1103	0.51	0.1287	0.49
(2001)	Oligocene	**0.0010***	0.85	**0.0040***	0.79	0.1133	0.50	0.1200	0.50	**0.0013***	0.84	**0.0099***	0.74	**0.0177**	0.69	0.1164	0.5	**0.0011***	0.84	**0.0044***	0.78	0.1131	0.5	0.1276	0.49
	Miocene	**0.0223**	0.68	**0.0400**	0.62	0.0922	0.53	0.0552	0.59	0.0579	0.59	0.1093	0.51	0.0552	0.59	0.2507	0.38	0.2447	0.38	0.2876	0.35	0.1791	0.44	0.204	0.42
	LGM	0.0801	0.55	0.1066	0.51	0.2633	0.37	0.0686	0.57	0.1828	0.43	0.2585	0.37	0.1045	0.52	0.3555	0.31	0.3469	0.31	0.3808	0.29	0.2337	0.39	0.2375	0.39
	Mid-Holocene	0.4523	0.25	0.4197	0.27	0.4996	0.23	0.1998	0.42	0.2267	0.4	0.2157	0.41	0.343	0.32	0.0976	0.52	0.4726	0.24	0.4947	0.23	0.3072	0.34	0.2828	0.36
	Now	0.4263	0.27	0.5492	0.20	0.1973	0.42	0.4686	0.24	0.3368	0.32	0.3819	0.29	0.1966	0.42	0.2823	0.36	0.3368	0.32	0.3819	0.29	0.1966	0.42	0.2823	0.36
Scotese (2003)	Eocene	**0.0074***	0.75	**0.0147**	0.71	0.2867	0.35	0.2056	0.41	**0.0079***	0.75	**0.015**	0.71	0.2664	0.37	0.2136	0.41	**0.0102**	0.73	**0.0186**	0.69	0.2686	0.37	0.2362	0.39
	Oligocene	**0.0112**	0.73	**0.0209**	0.68	0.3339	0.32	0.2409	0.39	**0.0097***	0.74	**0.0174**	0.7	0.2831	0.36	0.2267	0.4	**0.0101**	0.73	**0.0186**	0.69	0.2627	0.37	0.2367	0.39
	Miocene	**0.0027***	0.81	**0.0065***	0.76	0.1350	0.48	0.1129	0.51	**0.0016***	0.83	**0.0038***	0.79	0.2187	0.4	0.1065	0.51	**0.0241**	0.67	**0.0376**	0.63	0.0948	0.56	0.1043	0.52
	LGM	0.4253	0.27	0.5478	0.20	0.1974	0.42	0.4677	0.25	**0.0144**	0.71	**0.0229**	0.67	0.1168	0.5	0.1007	0.52	0.1049	0.52	0.1439	0.47	0.0877	0.54	0.1983	0.42
	Mid-Holocene	0.4261	0.27	0.5488	0.20	0.1974	0.42	0.4683	0.24	0.2026	0.42	0.2742	0.36	0.071	0.56	0.2864	0.35	0.2028	0.42	0.2745	0.36	0.071	0.56	0.287	0.35
	Now	0.4263	0.27	0.5492	0.20	0.1973	0.42	0.4686	0.24	0.203	0.42	0.2749	0.36	0.071	0.56	0.2875	0.35	0.203	0.42	0.2749	0.36	0.071	0.56	0.2875	0.35
Willis and McElwain	Eocene	**0.0018***	0.82	**0.0094***	0.74	0.0667	0.57	0.0911	0.53	**0.0022***	0.82	**0.0106**	0.73	0.0663	0.57	0.0984	0.52	**0.0064***	0.76	**0.0131**	0.72	0.2485	0.38	0.1893	0.43
(2002)	Oligocene	**0.0094***	0.74	**0.0171**	0.70	0.2523	0.38	0.2229	0.40	**0.0049***	0.78	**0.0113**	0.73	0.222	0.4	0.1811	0.44	**0.0078***	0.75	**0.015**	0.71	0.2639	0.37	0.2033	0.42
	Miocene	**0.0473**	0.61	**0.0469**	0.61	0.4354	0.26	**0.0281**	0.66	**0.0005***	0.87	**0.0014***	0.83	0.2393	0.39	**0.0294**	0.65	**0.0369**	0.63	**0.046**	0.61	0.2165	0.41	**0.0496**	0.6
	LGM	0.4202	0.27	0.5399	0.21	0.1998	0.42	0.4549	0.25	0.0521	0.6	**0.0482**	0.61	0.4448	0.26	**0.0284**	0.66	0.1419	0.47	0.1641	0.45	0.2041	0.42	0.1213	0.5
	Mid-Holocene	0.4252	0.27	0.5471	0.20	0.1983	0.42	0.4649	0.25	0.2601	0.37	0.3295	0.32	0.1017	0.52	0.3065	0.34	0.2652	0.37	0.3381	0.32	0.0983	0.52	0.3215	0.33
	Now	0.4263	0.27	0.5492	0.20	0.1973	0.42	0.4686	0.24	0.271	0.36	0.3479	0.31	0.0952	0.53	0.3385	0.32	0.271	0.36	0.3479	0.31	0.0952	0.53	0.3385	0.32

CHAPTER 2

SHIFTING HABITATS ACROSS GLACIAL AND INTERGLACIAL ASIA

INTRODUCTION

The past 2 million years of Earth's history have been characterized by glacial-interglacial climatic cycles. As the climate has cooled and warmed, glacial advances and retreats and corresponding rising and falling of sea levels caused dramatic changes in the connectivity, spatial configuration of habitats, and available area of terrestrial landscapes. The cyclical nature of the glacial-interglacial climatic episodes has been well characterized throughout the Quaternary (the past 1.8 million years encompassing the Pleistocene and Holocene epochs), with early Quaternary glaciation cycles occurring at 41,000 year intervals and then shifting to 100,000 year cycles about 800,000 years ago (Jansson and Dynesius, 2002; Raymo et al., 1998). The impact of glacial and interglacial cycles on faunal and floral distributions is particularly well documented in North America (Shafer et al., 2010), and Europe (Hewitt, 2000; Hewitt, 2004; Petit et al., 2008), however, much less is known about those responses in Asia. Few studies have explicitly evaluated how Asian diversity distributions changed with glacial and interglacial periods, and no syntheses have been compiled across taxonomic groups.

Glacial advance in northern Asia was much less dramatic than that observed in North America and Europe, where respectively the Laurentide and Weichselian Ice sheets advanced far to the south. In Asia glaciations were primarily restricted to high elevation mountains (Ehlers and Gibbard, 2004), although in northern Asia large glaciated areas existed in the Altai and Khangai Mountain ranges (Lehmkuhl et al., 2004), the Siberian mountains (Sheinkman, 2004), and may have been 900 times more extensive during the Pleistocene than now (Zamaruyev, 2004) . Glacial regions have been identified on the Tibetan Plateau (Qinghai-Xizang Plateau) (Klinge and Lehmkuhl, 2004; Kuhle, 2004; Zhou et al., 2004), Nepal (Fort, 2004; Konig, 2004), Japan (Sawagaki et al., 2004), Taiwan (Bose, 2004), northern India (Owen, 2004), and Pakistan (Kamp and Haserodt, 2004). Even some of the high elevation mountains in tropical Asia were glaciated during the Pleistocene glaciations. Many of the central mountains on the isle of New Guinea supported glaciers (Peterson et al., 2004), and three mountains in Sumatra (Mts. Kemiri, Leuser, Bandahara) as well as Mt Kinabalu (Borneo) are all thought to have supported icecaps, which suggests that additional high elevation mountains in the region may have as well (Hope, 2004). (Figure 1)

While large ice sheets may not have formed in Asia, glacial-interglacial climatic oscillations had substantial effects on regional and global climates. The uplift of the Tibetan Plateau in combination with the Northern Hemisphere glaciations caused dramatic changes in the East Asian and Indian monsoonal patterns (Zhisheng et al., 2001), which in turn is thought to have led to the development of the Thar Desert and intensification of desertification in the Sahara (Kuhle, 2004). The ongoing aridification of Northern Asia has also been linked to glaciation cycles (Guo et al., 2002; Wu et al., 2011). At a more localized level, the glaciations resulted in shifting tree-lines and dramatic oscillations in vegetation types (Kramer et al., 2009).

Here I utilized a GIS based modeling approach to predict how tropical and temperate Asian habitats responded to glacial-interglacial conditions during the past 120 thousand years, and validate the models using palynological data. Models developed for this time-frame can be extrapolated as indicative of glaciation cycles throughout the Quaternary as the magnitude and amplitude of glaciations was relatively stable (Gibbard and Cohen, 2008). The use of GIS predictive models to predict the distribution of habitat types has been informative in understanding how habitats in other parts of the world have responded to climate in the past (Carnaval et al., 2009; Graham et al., 2006), and for predicting how they might respond to future climate changes (Loarie et al., 2008).

METHODS

Recently developed land cover classifications in Asia and SE Asia using SPOT-VEGETATION data (Bingfang et al., 2000; Bingfang et al., 2004; Stibig et al., 2007) provide the basis for this project. The 26 land classifications for insular and mainland SE Asia and South Asia (Stibig et al., 2007) were combined with 24 classifications developed for China (Bingfang et al., 2000; Bingfang et al., 2004). Using these classifications and current climatic conditions, I generated predictive models for natural land cover classification categories. A few natural habitat types could not be modeled. Those include habitats such as freshwater swamps and estuaries whose locations are determined by complex physiographic processes, and for the boreal forests (see Allen et al. (2010) for a recent analyses of climate change on the distribution of this biome). I converted each land category to point files that were then reduced so that only one unique point was retained per grid cell. Since habitats were not always contiguous, and varied greatly in area, rather than standardizing a set number of points within each category, I scaled presence to absence points in a 1:5 ratio. Random background points were generated from outside the presence boundaries. This was repeated for each category independently. The climatic parameters at these points were extracted to build environmental envelopes for each classification category. Maximum entropy was utilized to generate the predictive models (Phillips et al., 2006; Phillips et al., 2004). The maximum entropy modeling approach has been found to perform closest to modeling efforts that explicitly incorporate physiological constraints when modeling plant distributions (Hijmans and Graham, 2006). All runs were set with the following parameters: 25% of points were reserved for testing the model, duplicates were removed, random number seeds were used to initiate each run. The following statistical analyses were performed: receiver operating characteristic (ROC) curves were used to test the ability of the model to distinguish among cover classification categories under current climate conditions, area under the curves (AUC) were used to evaluate the models ability to classify occurrences from nonoccurrence, and jackknife sampling was used to test the contribution of each climatic variable to the predictive ability of the model.

The WorldClim v. 1.4 current climate model was utilized to predict present distributions (Hijmans et al., 2005). The Echam 3 global palaeoclimatic model (Klimarechenzentrum(DKRZ)Modellbetreuungsgruppe, 1992) was utilized, to model the climate at the Last Glacial Maximum (henceforth LGM) at 21ky. A second global palaeoclimatic model was used to model climatic conditions during the Last Interglacial

Period (henceforth LIG) from between 120-140ky bp (Otto-Bliesner et al., 2006). Climate data were downscaled to 30 second arc grid resolution. Because climatic seasonality is critical in structuring forest communities, 19 BioClim parameters were utilized for the palaeoclimatic models (Table 1).

In order to calculate the area of each habitat at each time slice, a threshold was required to determine presence or absence. Threshold levels varied among habitat models due to differences in the strength of predictions. Thresholds vary due to variations among models and characteristics reflecting the ecology of the modeling subject (Liu et al., 2005). For each habitat, presence/absence was determined by utilizing the threshold balancing sensitivity and specificity. The number of pixels with predicted occurrences was then multiplied by the area of each pixel (roughly 1 km^2) to obtain the predicted area of each habitat type at each time slice.

RESULTS

Seventeen habitat types were modeled for each of three time slices (present, LGM, LIG) (Figures 1-17). The ROC model scores ranged from 0.849 -0.999 indicating that the predictive models were able to accurately distinguish among vegetation types (Table 2). Palynological data from 51 LGM localities were used to validate the LGM GIS models. Of the 51 LGM pollen cores, seven were for habitat types either not modeled or not successfully modeled here (e.g. desert, freshwater swamp), and one from a site that is now submerged. Of the remaining 44 sites, 31 corroborate the predictive models (Table 3). Forest models generally were in better accord with pollen data than non-forests habitats, with only two non-corroborating pollen sites, thereby indicating that forest models are potentially more reliable than models for non-forest habitats. In particular, the grassland model and grassland pollen data showed the greatest number of disparities, with twelve mis-matches.

Tropical Forest Habitats (Figure 2)

Tropical lowland evergreen broadleaf forests (rainforests) are megadiverse assemblages distributed broadly but patchily across Asia. The data display a clear trend of increasing climatic suitability from the LIG to now, however the actual forest extent is much reduced. Predictions for the current extent of this forest type are 3.2 times greater than the actual extent. During the LIG, the predicted distribution of forest was most extensive in insular Asia. Lowland rainforest coverage is predicted on Sumatra, the Mentawai Islands, Borneo, New Guinea, the Philippines, Sulawesi and the Molucca Islands. In SE Asia, rainforest is predicted in Vietnam-Lao, the Mekong Delta region, the Isthmus of Kra, and Perak Malaysia. At the LGM, there was a predicted dramatic northward expansion of forest. Sri Lanka is predicted to have been covered by rainforest. In India, forest is predicted to have expanded northwards in the west with isolated patches of rainforest appearing in Andhra Pradesh, Chattisgarh, and Orissa. Rainforest is also predicted in NE India, and in several large tracts in Myanmar, Laos, Vietnam, southern Cambodia, and the Laos-Thailand-Cambodia border. The forest on the Malay Peninsula is predicted to have expanded northwards in peninsular Myanmar, and all along the eastern coast, connecting formerly isolated patches. Isolated rainforest also is

predicted to have appeared appeared in western Malaysia. In insular Asia forest expansion is predicted for Sumatra which connected some of the forest patches predicted from the LIG, and forest expansion is predicted in western Java and the Philippines. Marginal forest contraction is indicated in Borneo and New Guinea, with severe contraction (e.g. disappearance) predicted for Sulawesi, the Lesser Sundas, and the Molucca Islands indicating much drier conditions in this region during the LGM.

Tropical and subtropical montane evergreen broadleaf forest is widely distributed throughout the region, currently covering 435,034 km^2. Like the lowland rainforests, montane rainforests are predicted to be increase in area from the LIG to the present, but again, much greater current area is predicted than actually exists. During the LIG montane evergreen forests are predicted to have covered much of Sulawesi, the Molucca Islands, and interior Borneo, but was largely absent from northern SE Asia, although a few isolated patches were scattered across the mountainous regions of southern India, Sri Lanka, eastern Myanmar, Yunnan, China, northern Laos, the Cardamom Mountains of Cambodia, central and southern Vietnam, peninsular Thailand, Myanmar, Malaysia, Java, and western Sumatra. By the LGM however, montane forests are predicted to have expanded in mainland Asia with forests extending throughout the Chin and Rakhine mountains in western Myanmar, and throughout Yunnan and southern Sichuan China, across northern Laos, and the southern edge of the Himalayas in NE India and Nepal. They are also predicted to have expanded marginally in southern India and Sri Lanka. Forest expansion across the mountains and into lower elevation regions of Sumatra and Java is predicted, but conversely, contraction is predicted for Borneo. Severe forest contractions are predicted in Sulawesi, and the Lesser Sundas and the Moluccan Islands.

Tropical broadleaf deciduous forests have high stability in mainland Asia. This habitat type is predicted to cover an area seven times greater than its current extent of 750,497 km^2. During the LIG forests are predicted across much of mainland SE Asia (except for the Malay Peninsula), as well as across most of India, however, a large break is predicted between these forests across Bangladesh and the Gangetic plains of northern India. Much of the forest is predicted to have been contiguous in India, however a large isolated patch of forest is predicted to have existed in Madhya Pradesh. Sri Lanka, Hainan, and some of the Philippine Islands also appear to have had extensive forest tracts at this time. The LGM saw dramatic changes in the predicted distribution of this forest type characterized by distributional shifts, massive expansions, and severe contractions. In most of India this forest is predicted to have expanded and shifted. The isolated patch in Madhya Pradesh became connected, and new isolated patches appeared in Rajasthan. An arm of forest skirting the southern edge of the Himalayas is predicted to have stretched into Himachal Pradesh, with isolated patches of forest appearing in Pakistan. In mainland SE Asia, much of the forest is predicted to have contracted leaving isolated patches in northeast India, Myanmar, southern China and northern Vietnam. Larger patches of forest may have remained throughout Vietnam, Cambodia, and eastern Thailand. Forest contraction is predicted in Hainan but expansion is predicted for the Philippines. Most dramatically, this forest type is predicted to have occurred in insular Asia, in particular Borneo, Sulawesi, Java, the northern tip of Sumatra and the southern tip of Malaysia, indicating drier conditions in this region at the LGM.

At present swamp forest is found in Borneo, Sumatra, Sulawesi, Bangka Island, and New Guinea, and in the Tonle Sap region of Cambodia. At the LIG the predicted range of this forest type was greater, with swamp forests predicted on the western coast of the Malay Peninsula, Sumatra, Bankga, Java, Bali, Borneo, Sulawesi, the Moluccas, New Guinea, and Palawan Island (Philippines). At the LGM, swamp forest expansion is predicted in western Borneo but with contraction in the east. They are predicted to have disappeared from Sulawesi, Java, Bali, most of the Moluccas, and Palawan. In New Guinea, ranges shifted and western forests expanded. Much of the swamp forest of Sumatra is predicted to have contracted leaving two small isolated patches, but forest on Bangka expanded. One of the Mentawai islands is predicted to have had swamp forest during the LGM.

Mangrove Forests are one of the more imperiled habitats in Asia. At present, they cover only 55,348 km^2 across the region. The current extent is predicted to be 1,396,061km^2, with mangrove habitat stretching almost continuously from western India to insular Asia. In insular Asia, mangrove habitat is predicted to occur along most of the coasts of most of the islands (excepting a few of the Lesser Sundas). At the LIG mangroves were predicted primarily along the coasts of the Sundaic landmasses, with extensive mangroves along the Mekong Delta, peninsular Malaysia, Sumatra, and Borneo. Mangrove forests are also predicted on Sulawesi, particularly along the northern coasts, in a single patch on Java, and on Timor, the Moluccan Islands, New Guinea, and throughout the Philippines. Mangroves were also predicted along the Western coast of India, Sri Lanka, and along the southern coast of Myanmar. At the LGM mangrove forests are predicted to have expanded significantly in mainland Asia, with an almost continuous mangrove corridor predicted from western India to the Malay Peninsula. A dramatic break in southeast India would have separated the eastern and western Indian mangrove forests. The Gangetic plains and Bangladesh are predicted to have been covered in mangrove forest, and connected with mangroves stretching along the coast of Myanmar. On the Malay Peninsula, extensive mangrove habitat is predicted on the northwestern coast, but they are not predicted for Peninsular Malaysia. In insular Asia mangrove is predicted on the northern coast and eastern coasts of Sumatra, Bangka, Borneo, some of the Philippine Islands (most extensively on Luzon), the eastern coasts of Sulawesi, the Moluccas, and New Guinea.

Temperate Forests (Figure 3)

Temperate deciduous broadleaf forest is presently found in eastern China and covers 611,993km^2. The area predicted for this habitat is considerably greater at 3,647,689 km^2. Using the strict threshold approach, this habitat was not predicted for either the LGM or LIG time slices.

At present, lowland temperate evergreen broadleaf forest is found in southeast China, Hainan and Taiwan Islands, and NE India. It is predicted to occur much more widely than this, with significant predicted occurrences in north to central Vietnam, northern Myanmar, and southern Japan. At the LIG, this forest type was much less extensive than now, with isolated habitat patches predicted in northern Vietnam, eastern China, Taiwan, Hainan, and southern Japan. At the LGM this habitat is predicted to have

undergone massive expansion in eastern China, South Korea, Japan, northern Myanmar and northern India. In China, the forest is predicted to have extended as far west as Sichuan. Conversely, forests are predicted to have contracted on Hainan and Taiwan.

Temperate mountain broadleaf evergreen forest is presently distributed throughout eastern China, and the mountainous regions of Myanmar, Bhutan, Nepal and Northern India as well as Taiwan and Hainan. Its distribution is predicted to be much greater in eastern China, and the mountainous regions of Vietnam, and Myanmar. It is also predicted in southern Japan and South Korea. At the LIG this forest distribution was predicted to be farther south, with southern extensions along the mountainous regions of Myanmar and Vietnam as well as patches in eastern China, Taiwan, Hainan and southern Japan. At the LGM, the predicted distribution expanded northward in China and Japan, appearing also on the Korean peninsula. It also expanded west along the southern edge of the Himalayas through Bhutan, and Nepal into Himachal Pradesh, India. Two new isolated patches are predicted in continental India in Maharashtra and at the border of Orissa and Andra Pradesh States.

Temperate evergreen coniferous forest is distributed in the mountains of eastern China stretching from the far southern border to the northern state of Nei Mongol Zizhiqu and patchily through the mountains to the western edge of the Taklamankan Desert. The forest follows the southern edge of the Himalayas west into the mountains of Pakistan. The predicted occurrence includes South Korea, Japan, northern Vietnam, Laos, Myanmar, and the high peaks in Orissa, Vanachal, and Chhattisgarh States of eastern India. During the LIG the predicted general distribution of this forest was shifted farther south relative to the current distribution. While some forest was present in southern China, much less of this region is predicted to have been suitable. Instead, Vietnam, Lao, Thailand, and Myanmar form a large forest patch connected to the southern China forest across the Vietnam-Yunnan border. A patchy distribution of forest is predicted along the Himalayas stretching west to Pakistan. A large forest patch is predicted in Bangladesh, Varanachal and Bihar States, India, a second isolated patch is predicted in the mountains of Orissa, Chhattisgharh, and Andra Pradesh, and a third large forest patch is predicted across the Pakistan, Punjab, Haryana, and Greater Rajasthan region. In the east, forest is predicted in Taiwan, Hainan, Japan, and South Korea. With the onset of the LGM, dramatic shifts were again observed with predicted forest shifting north across its entire range. SE Asian forest largely disappeared leaving small fragments in northern Vietnam, Myanmar and NE India. In India, forest is predicted to have retracted to the southern edge of the Himalayas. In eastern coastal China, a massive range shift is predicted with forest taking hold around Hebei, Lianing, Shandong, Tianjin, and Jiangsu. The Fujian forest is predicted to have expanded north into Zheijing. In interior eastern China, a large forest expansion from the Vietnam border north to Shaanxi and west to Xizang Zhzhqu predicted. In Japan, forest is predicted to have expanded northward dramatically.

Non-forest habitats (Figures 4, 5)

Alpine grasslands currently span over 2,000,000 km^2, although their predicted current extent is twice what we observe. Alpine grasslands are presently primarily restricted to the mountains of southeastern China and the Tibetan Plateau. At the LIG,

this habitat was reduced in reference to its current distribution. Most of the predicted alpine habitat was restricted to the mountains of Sichuan and Xizang Zizhiqu, but with southern extensions into Yunnan, northern Myanmar and NE India. Alpine habitats may have reached as far west as Nepal. At the LGM the habitat is predicted to have contracted along much of its range, with a large alpine area emergent in the Hubei region.

Grasslands are broadly distributed in temperate regions of Asia, currently covering 866411 km^2. The predicted distribution of this habitat is vast covering the majority of China, parts of SE Asia, and the far western edge of India and Pakistan. The predicted area of this habitat is substantially greater than the actual extent. During the LIG, this habitat was also predicted to have been extensive throughout temperate Asia, although the predicted distribution is shifted south relative to much of current distribution. Large grassland areas are predicted in Myanmar, Laos, Vietnam, NE India, and Bangladesh as well as the southern Chinese border. Some smaller grassland habitats are predicted across northern China, and subcontinental India. At the onset of the LGM the models indicate a severe contraction, with grassland habitats covering only 217828 km^2. Grassland habitats were predicted primarily in eastern China close to the coast, Hainan, northern Vietnam, and as an isolated patch in southern Pakistan.

Temperate Meadow habitats are currently found primarily in NE China. The predicted extent of this habitat is greater with extensive meadow habitat predicted in Mongolia, southern Russia, and in a belt crossing north of the Tibetan Plateau reaching as far west as Kyrgyzstan. This habitat was predicted with suitability scores below the thresholds for both the LIG and LGM and as such its paleodistribution can not be estimated with confidence.

Sparse woods are primarily found in eastern China, but range as far west as Yunnan Province. During the LIG the distribution of this habitat is predicted to have been shifted substantially, with habitat predicted in central Vietnam, Hainan, Taiwan, Yunnan, Guangdong, Fujian, and Guangxi. At the LGM, the habitat was predicted to have been confined to northern Vietnam, Guangxi, Guangdong, and Hainan.

Deciduous thorny scrub is currently found throughout most of subcontinental India. At the LIG, a swath of thorny scrub is predicted from western India to southeast India, isolating the western coast of India. A single patch of this habitat is predicted in the Central Dry Zone of Myanmar. The predicted and actual contemporary distributions of this habitat are similar. At the LGM, the predicted distribution shifted north centering in Pakistan and western India (Greater Rajasthan, Madhya Pradesh and Gujarat). The Myanmar habitat is predicted to be absent in this interval, but a new patch is predicted at the Thai-Laos-Cambodia border.

Desert and Sparse Grassland is widely distributed across much of temperate Asia. The expected occurrence of this habitat is greater than the observed extent. Two isolated habitat patches are predicted, one encompassing western India and Pakistan, the second across the much of central China. During the LIG and LGM this grass/desert habitat is predicted to have been restricted to lowlands in western India and Pakistan, with a

contraction occurring during the LGM. From the LGM to present a marked expansion is predicted across temperate Asia, which is consistent with continued continental aridification.

Deserts are widespread in temperate Asia, covering 894,042 km^2. Desert habitats were not predictable in the past, with suitability indices below threshold values for both the LGM and LIG time frames. Desert and sparse grassland habitats however were predicted at both the LGM and LIG. At the LGM and LIG, this habitat was largely predicted to have been restricted to Pakistan, which marks a drastic range reduction as compared to now, although the habitat is predicted to be more extensive during the LIG than the LGM. An interesting finding is that based on contemporary climatic conditions, the extent of this habitat type is predicted to be much more extensive than observed, in particular much more of western China is predicted to be covered in this habitat type.

Deciduous shrublands are found throughout mainland Asia, and are likely to represent bamboo groves (Stibig et al., 2007). Presently much of this habitat is centered in eastern China and SE Asia. The contemporary distribution of this habitat is predicted to be much greater across eastern China and throughout mainland SE Asia. During the LIG, this habitat is predicted to have shifted south and west, and is predicted across much of mainland SE Asia as well as eastern China, Korea, Japan, and southern India, Sri Lanka, and the Philippines. The LGM distribution of this habitat moved farther into the insular areas of Asia, with predicted distribution in Borneo, Java, some of the smaller Indonesian Islands, New Guinea, and the Philippines. In mainland Asia, much of the eastern coast of China, and the Vietnam coast is predicted to have consisted of this habitat. Farther inland a large patch is predicted in Thailand, the Ayeyarwaddy Delta region of Myanmar, the southern edge of the Himalayas and southern India and Sri Lanka.

Sparse shrubland is currently found in Java, Borneo, the Lesser Sundas, New Guinea, and the Philippines. It is predicted to have a much wider current day distribution throughout the Sundaic landmasses (Borneo, Java, Sumatra, peninsular Malaysia) as well as Sulawesi, the Lesser Sundas, New Guinea and some of the Philippine Islands. During the LGM, this habitat is predicted to have been very extensive across Sumatra, Peninsular Malaysia, Northern Borneo, and New Guinea, with smaller habitat patches in western Java, and on some of the smaller islands, as well as a few predicted patches in southern Cambodia and the Myanmar/Bangladesh border. At the LIG the distribution was predicted to have been shifted south with far less of this habitat in New Guinea and Peninsular Malaysia, but with extensive patches predicted in Sumatra, southern Java, southern Borneo, Sulawesi, and the Philippines.

DISCUSSION

Predicted areas through time show some clear patterns (Figure 6). While several of the habitats are stable in area (although not necessarily location) through time, a few habitats show dynamic patterns, in particular, several of the non-forest habitats are predicted to have severely contracted during the LGM and to have expanded rapidly

between the LGM and present. Natural non-forest habitats comprise a great majority of the contemporary landscape in temperate Asia and mainland SE Asia (less so in Insular Asia). Non-forest habitats consist of grasslands, alpine habitats, deciduous thorny scrubland, desert and sparse desert grasslands, deciduous shrublands, sparse shrubland, and sparse woods. They are however predicted to have a smaller extent than they actually have. This likely reflects anthropogenic activities since both grasslands and shrublands are often early successional communities after habitat destruction.

All but one habitat (deciduous thorny scrubland) are predicted to have a greater area than the actual present day habitat. This reflects at least in part the wide-scale conversion of land from natural habitat to agricultural, urban, and otherwise anthropogenically modified habitats. Stibig (2007) estimated that cropland, plantations and urban centers accounted for 4,282,000 km^2, this number increases to 5,268,700 km^2 if fragmented forests, regrowth cover, abandoned shifting cultivation, and cropland shrub mosaic land cover is included.

Overall patterns emerge for the grouped habitat categories. At the LIG, predicted non-forest habitats were greater in extent than all forest types combined, however, at the LGM there was a contraction of these non-forest habitats and an expansion of both temperate and tropical forests (Figure 7). From the LGM until the present, the relative proportions of the three habitat categories imply an expected increase of non-forest habitats, and reduction of both tropical and temperate forests (Figure 7); however this pattern is partially driven by the over prediction of grassland habitat. Currently, the actual extent of non-forest natural habitats exceeds the combined extent of both temperate and tropical forests.

The distribution of vegetation types at the Last Glacial Maximum has been of interest to researchers trying to reconstruct temperature histories, monsoonal patterns, as well as phylogeographic histories from around the globe. Global reconstructions of vegetation types have been based on many different sources of data and compiled from many different types of analyses. For Asia, few paleo-vegetation reconstructions exist, but we can compare them against the GIS based reconstructions presented here. Three reconstructions come from an extensive palynological database compiled by the Biome Group (Prentice et al., 2000). These analyses implement a “biomitization” approach in which pollen abundance data are utilized to determine biome type (Prentice et al., 2000). The first of these analyses spans across Asia and Australia and provides reconstructions for the modern, 6ky and 18ky time periods, but includes only a handful of pollen sites for SE Asia (Pickett et al., 2004). The reconstructions at the LGM for the sites within the geographic realm of this study include sites in Thailand, Java and Sulawesi, all of which are reconstructed as having been covered in tropical deciduous broadleaf forests and woodlands. A single site on Sumatra was recovered as having been xerophytic scrubland, and several sites on New Guinea were recovered as having cool-temperate rainforest. All of our models accord with these analyses. The second study focused on China, and added 37 palynological records for the region at the LGM (Yu et al., 2000). The LGM reconstructions suggest accord with our models for all biomes except grassland. Yu et al. (2000) suggest that temperate deciduous forests which currently dominate the eastern

coast of China may have been farther North than present. However pollen data were not available to confirm or refute their hypothesis. The models generated here suggest that those forests were shifted north as they contracted at the LGM. The third study focused on east Asia, particularly China (Harrison et al., 2001). Again, there is strong concordance among the Biome models and the GIS models, however the GIS based models recovered more extensive temperate deciduous and coniferous forests. We reconstructed sparse woods and deciduous shrublands in the region, which loosely correspond in location and extent with their non-forest and savannah-woodlands biomes. Ray and Adams (2001) generated a global LGM reconstruction approach which incorporated multiple paleoclimatic proxies. The distribution of vegetation differs markedly from the models generated here, which may result form the difference in scale (global versus regional) as well as the difference in modeling approach. Some significant differences include their inferred restriction of tropical forest to Borneo and the eastern Sunda shelf, and their finding of extensive grassland habitat predicted across India.

Two limitations of the models generated here are the consistent over-prediction of habitats, and the changing coastlines through out this time. Over-prediction among habitat types is most problematic when thinking about the broad habitat classes; tropical forest, temperate forest, and non-forest. Recent analyses have shown that tropical forest and savanna (non-forest) habitats may simply be alternative stable states controlled by environmental factors such as precipitation (Hirota, 2011; Staver et al., 2011). Quantification of overlap among the predictive models at both habitat and broad habitat classes will help better understand the nature of the observed over-prediction. The changing coastlines associated with dramatic sea-level changes are another confounding factor in precise palaeo-reconstruction. In particular, reconstruction of some habitats such as mangrove habitats will be greatly affected by changing coastlines, and this is not accounted for here. Chapter four however focuses on the effects of changing sea-levels across the Sundaland region (Wogan, 2011).

The paleo-vegetation reconstructions provided here are the first to model multiple habitats across Asia within a predictive modeling framework. The inclusion of both glacial and interglacial reconstructions provide insight into the dynamic architecture of the Asian landscape under differing climatic regimes. These reconstructions can provide biogeographers and phylogeographers with a means of generating *a priori* hypotheses that can be tested utilizing genetic and other biodiversity metrics. Lastly, paleo-vegetative models can aid in the identification of critical areas for conservation by laying the basis for identification of refugia, centers of neo-endemism, paleo-endemism, and evolutionary hotspots.

LITERATURE CITED

Allen, J. R. M., T. Hickler, J. S. Singarayer, M. T. Sykes, P. J. Valdes, and B. Huntley. 2010. Last glacial vegetation of northern Eurasia. Quaternary Science Reviews 29:2604-2618.

Ansari, M. H., and A. Vink. 2007. Vegetation history and palaeoclimate of the past 30 kyr in Paistan as inferred from the palynology of continental margin sediments off the Indus Delta. Review of Palaeobotany and Palynology 145 201–216.

Anshari, G., A. P. Kershaw, and S. v. d. Kaars. 2001. A late Pleistocene and Holocene pollen and charcoal record from peat swamp forest, Lake Sentarum Wildlife Reserve, West Kalimantan, Indonesia. Palaeogeography, Palaeoclimatology, Palaeoecology 171:213-228.

Bingfang, W., X. Wenting, H. Huiping, and Y. Changzhen. 2000. The Land Cover Map for China in the year 2000European Commision Joint Research Center, GLC Database.

Bingfang, W., X. Wenting, H. Huiping, Y. Changzhen, and X. Wenbo. Year. Combining Spot4-Vegetation and Meterological Data Derived Land Cover Map in China i*n* Geoscience and Remote Sensing Symposium, IEEE International IGARSS '04. Proceedings. 2004 2004:4.

Bose, M. 2004. Traces of glaciation in the high mountains of Taiwan. Pages 347-352 i*n* Quaternary Glaciations-Extent and Chronology (J. Ehlers, and P. L. Gibbard, eds.). Elsevier, Amsterdam.

Carnaval, A. C., M. J. Hickerson, C. F. B. Haddad, M. T. Rodrigues, and C. Moritz. 2009. Stability Predicts Genetic Diversity in the Brazilian Atlantic Forest Hotspot. Science 323:785-789.

Ehlers, J., and P. L. Gibbard (eds) 2004. Quaternary glaciations- extent and chronology. Elsevier, Amsterdam.

Fort, M. 2004. Quaternary glaciation in the Nepal Himalaya. Pages 261-278 i*n* Quaternary Glaciations-Extent and Chronology (J. Ehlers, and P. L. Gibbard, eds.). Elsevier, Amsterdam.

Gibbard, P., and K. M. Cohen. 2008. Global chronostratigraphical corelation table for the last 2.7 million years. Episodes 31:243-247.

Graham, C. H., C. Moritz, and S. E. Williams. 2006. Habitat history improves prediction of biodiversity in rainforest fauna. Proc. Natl. Acad. Sci. U.S.A. 103:632-6.

Guo, Z., W. Ruddiman, Q. Hao, H. Wu, and Y. Qiao. 2002. Onset of Asian desertification by 22 Myr ago inferred from loess deposits in China. Nature 416:159-163.

Harrison, S., G. Yu, H. Takahara, and I. Prentice. 2001. Palaeovegetation - Diversity of temperate plants in east Asia. Nature 413:129-130.

Herzschuh, U., A. Kramer, S. Mischke, and C. Zhang. 2009. Quantitative climate and vegetation trends since the late glacial on the northeastern Tibetan Plateau deduced from Koucha Lake pollen spectra. Quaternary Research 71:162-171.

Hewitt, G. 2000. The genetic legacy of the Quaternary ice ages. Nature 405:907-913.

Hewitt, G. 2004. Genetic consequences of climatic oscillations in the Quaternary. Philos T Roy Soc B 359:183-195.

Hijmans, R., S. E. Cameron, J. L. Parra, P. G. Jones, and A. Jarvis. 2005. Very high resolution interpolated climate surfaces for global land areas. International Journal of Climatology 25:1965-1978.

Hijmans, R. J., and C. H. Graham. 2006. The ability of climate envelope models to predict the effect of climate change on species distributions. Global Change Biol 12:2272-2281.

Hirota. 2011. Global resiliance of tropical forest and savanna to critical transitions. Science 334:232-.

Hope, G., and J. Tulip. 1994. A long vegetation history from lowland Irian Jaya, Indonesia. Palaeogeography, Palaeoclimatology, Palaeoecology 109:385-398.

Hope, G. S. 1976. The vegetational history of Mt. Wilhelm, Papua New Guinea. Journal of Ecology 64:627-663.

Hope, G. S. 2004. Glaciation of Malaysia and Indonesia, excluding New Guinea. Pages 211-214 i*n* Quaternary Glaciations-Extent and Chronology (J. Ehlers, and P. L. Gibbard, eds.). Elsevier, Amsterdam.

Jansson, R., and M. Dynesius. 2002. The fate of clades in a world of recurrent climatic change: Milankovitch oscillations and evolution. Annual Reviews in Ecology and Systematics 33:741-777.

Kamp, U., and K. Haserodt. 2004. Quaternary glaciations in the high mountains of northern Pakistan. Pages 293-311 i*n* Quaternary Glaciations-Extent and Chronology (J. Ehlers, and P. L. Gibbard, eds.). Elsevier, Amsterdam.

Klimarechenzentrum(DKRZ)Modellbetreuungsgruppe, D. 1992. The ECHAM3 atmospheric general circulation model. Pages 184 i*n* DKRZ Tech. Report Deutsches Klimarechenzentrum, Hamburg, Germany.

Klinge, M., and F. Lehmkuhl. 2004. Pleistocene glaciations in southern and eastern Tibet. Pages 361-369 i*n* Quaternary Glaciations-Extent and Chronology (J. Ehlers, and P. L. Gibbard, eds.). Elsevier, Amsterdam.

Konig, O. 2004. The glaciation of the Rolwaling Himal and the Kangchenjunga Himal during the Last Glacial Maximum (Nepal, E-Himalaya). Pages 279-284 i*n* Quaternary Glaciations-Extent and Chronology (J. Ehlers, and P. L. Gibbard, eds.). Elsevier, Amsterdam.

Kramer, A., U. Herzschuh, S. Mischke, and C. Zhang. 2009. Late glacial vegetation and climate oscillations on the southeastern Tibetan Plateau inferred from the Lake Naleng pollen profile. Quaternary Research:1-12.

Kuhle, M. 2004. The high glacial (Last ice age and LGM) ice cover in High and Central Asia. Pages 175-199 i*n* Quaternary Glaciations - Extent and Chronology (J. Ehlers, and P. L. Gibbard, eds.). Elsevier, Amsterdam.

Lehmkuhl, F., M. Klinge, and G. Stauch. 2004. The extent of Late Pleistocene glaciation in the Altai and Khangai Mountains. Pages 243-254 i*n* Quaternary Galciations- Extent and Chronology (J. Ehlers, and P. L. Gibbard, eds.). Elsevier, Amsterdam.

Liew, P. M., C. M. Kuo, S. Y. Haung, and M. H. Tseng. 1998. Vegetation change and terrestrial carbon storage in eastern Asia during the Last Glacial Maximum as indicated by a new pollen record form central Taiwan. Global and Planetary Change 16-17:85-94.

Liu, C., P. M. Berry, T. P. Dawson, and R. G. Pearson. 2005. Selecting thresholds of occurrence in the prediction of species distributions. Ecography 28:385-393.

Loarie, S. R., B. E. Carter, K. Hayhoe, S. Mcmahon, R. Moe, C. A. Knight, and D. D. Ackerly. 2008. Climate change and the future of California's endemic flora. PLoS ONE 3:e2502.

Morley, R. J., and J. R. Flenley. 1987. Late Cainozoic vegetational and environmental changes in the Malay archipelago. Pages 50-59 i*n* Biogeographical evolution of the Malay archipelago (T. C. Whitemore, ed.) Clarendon Press, Oxford.

Newsome, J., and J. Flenley. 1988. Late Quaternary Vegetational History of the Central Highlands of Sumatra. II. Palaeopalynology and Vegetational History. J Biogeography 15:555-578.

Otto-Bliesner, B. L., S. J. Marshall, J. T. Overpeck, G. H. Miller, and A. Hu. 2006. Simulating Arctic climate warmth and icefield retreat in the last interglaciation. Science 311:1751-3.

Owen, L. A. 2004. Late Quaternary glaciation of northern India. Pages 201-209 i*n* Quaternary Glaciations-Extent and Chronology (J. Ehlers, and P. L. Gibbard, eds.). Elsevier, Amsterdam.

Penny, D. 2001. A 40,000 year palynological record from north-east Thailand; implications for biogeography and palaeo-environmental reconstruction. Palaeogeogr Palaeocl 171:97-128.

Peterson, J. A., S. Chandra, and C. Lundberg. 2004. Landforms from the Quaternary glaciation of Papua New Guinea: an overview of ice extent during the Last Glacial Maximum. Pages 313-319 i*n* Quaternary Glaciations-Extent and Chronology (J. Ehlers, and P. L. Gibbard, eds.). Elsevier, Amsterdam.

Petit, R. J., F. S. Hu, and C. W. Dick. 2008. Forests of the past: A window to future changes. Science 320:1450-1452.

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

Phillips, S. J., M. Dudık, and R. E. Schapire. 2004. A Maximum Entropy Approach to Species Distribution Modeling Proceedings of the 21st International Conference on Machine Learning:8.

Pickett, E. J., S. P. Harrison, G. Hope, K. Harle, J. R. Dodson, A. P. Kershaw, I. C. Prentice, J. Backhouse, E. A. Colhoun, D. D'Costa, J. Flenley, J. Grindrod, S. Haberle, C. Hassell, C. Kenyon, M. Macphail, H. Martin, A. H. Martin, M. McKenzie, J. C. Newsome, D. Penny, J. Powell, J. I. Raine, W. Southern, J. Stevenson, J.-P. Sutra, I. Thomas, S. c. d. Kaars, and J. Ward. 2004. Pollen-based reconstructions of biome distributions for Australia, Southeast Asia and the Pacific (SEAPAC region) at 0, 6000 and 18,000 14C yr BP. Journal of Biogeography 31:1381-1444.

Prentice, I. C., D. Jolly, and Biome_6000_Participants. 2000. Mid-Holocene and Glacial-Maximum Vegetation Geography of the Northern Continents and Africa. Journal of Biogeography 27:507-519.

Qin, J., G. Wu, H. Zheng, and Q. Zhou. 2008. The palynology of the first hard clay layer (Late Pleistocene) from the Yangtze delta, China. Review of Palaeobotany and Palynology 149:63-72.

Ray, N., and J. M. Adams. 2001. A GIS-based vegetation map of the world at the last glacial maximu, (25,000-15,000bp). Internet Archaeology 11:1-43.

Raymo, M., K. Ganley, S. Carter, D. Oppo, and J. McManus. 1998. Millennial-scale climate instability during the early Pleistocene epoch. Nature 392:699-702.

Sawagaki, T., T. Aoki, H. Hirohiko, S. Iwasaki, S. Iwata, and K. Hirakawa. 2004. Late Quaternary glaciations in Japan. Pages 217-225 i*n* Quaternary Glaciations- Extent and Chronology (J. Ehlers, and P. L. Gibbard, eds.). Elsevier, Amsterdam.

Shafer, A. B. A., C. I. Cullingham, S. D. Côté, and D. W. Coltman. 2010. Of glaciers and refugia: a decade of study sheds new light on the phylogeography of northwestern North America. Molecular Ecology 19:4589-4621.

Sheinkman, V. 2004. Quaternary glaciation in the high mountains of central and north-east Asia. Pages 325-335 i*n* Quaternary Glaciations-Extent and Chronology (J. Ehlers, and P. L. Gibbard, eds.). Elsevier, Amsterdam.

Staver, A. C., S. Archibald, and S. A. Levin. 2011. The global extent and determinants of savanna and forest as alternative biome states. Science 334:230-232.

Stibig, H.-J., A. S. Belward, P. S. Roy, U. Rosalina-Wasrin, S. Agrawal, P. K. Joshi, R. Beuchle, S. Fritz, S. Mubareka, and C. Giri. 2007. A land-cover map for South and Southeast Asia derived from SPOT-VEGETATION data. J Biogeography 34:625-637.

Stujits, I., J. C. Newsome, and J. R. Flenley. 1988. Evidence for late quaternary vegetational change in the Sumatran and Javan highlands. Review of Palaeobotany and Palynology 55:207-216.

Van Der Kaars, S., and M. A. C. Dam. 1995. A 135,000-year record of vegetational and climatic change from the Bandung area, West Java, Indonesia. Palaeogeography, Palaeoclimatology, Palaeoecology 117:55-72.

van der Kaars, W. A. 1991. Palynology of eastern Indonesian marine piston-cores: A Late Quaternary vegetational and climatic record for Australasia. Palaeogeography, Palaeoclmatology, Palaeoecology 85:239-302.

White, J. C., D. Penny, L. Kealhofer, and B. Maloney. 2004. Vegetation changes from the late Pleistocene through the Holocene from three areas of archaeological significance in Thailand. Quaternary International 113:111–132.

Wogan, G. O. U. 2011. Chapter 4. Reconstructing ancient pathways: The Habitats of Sundaland during Glacials and Interglacialsi*n* Integrative Biology University of California, Berkeley, Berkeley.

Wu, F., X. Fang, M. Herrmann, V. Mosburgger, and Y. Miao. 2011. Extended drought in the interior of Central Asia since the Pliocene reconstructed from sporopollen records. Global and Planetary Change 76:16-21.

Yu, G., X. Chen, J. Ni, R. Cheddadi, J. Guiot, H. Han, S. P. Harrison, C. Huang, M. Ke, Z. Kong, S. Li, W. Li, P. Liew, G. Liu, J. Liu, Q. Liu, K.-B. Liu, I. C. Prentice, W. Qui, G. Ren, C. Song, S. Sugita, X. Sun, L. Tang, E. V. Campo, Y. Xia, Q. Xu, S. Yan, X. Yang, J. Zhao, and Z. Zheng. 2000. Palaeovegetation of China: a pollen data-based synthesis for the mid-Holocene and last glacial maximum. Journal of Biogeography 27:635-664.

Zamaruyev, V. 2004. Quaternary glaciation of north-eastern Asia. Pages 321-323 i*n* Quaternary Glaciations-Extent and Chronology (J. Ehlers, and P. L. Gibbard, eds.). Elsevier, Amsterdam.

Zhisheng, A., J. E. Kutzbach, W. L. Prell, and S. C. Porter. 2001. Evolution of Asian monsoons and phased uplift of the Himalaya-Tibetan plateau since Late Miocene times. Nature 411:62-6.

Zhou, S. Z., L. Jijun, S. Q. Zhang, J. D. Zhao, and J. X. Cui. 2004. Quaternary glaciations in China. Pages 105-113 i*n* Quaternary Glaciations- Extent and chronology Part III:

South America, Asia, Africa, Australia, Antarctica (J. Ehlers, and P. L. Gibbard, eds.). Elsevier, Amsterdam.

Figure 1. Quaternary glaciation patterns in Asia at the max (A) and at the last glacial maximum (B), bottom panel (C) shows the tropical mountain top glaciers across the Sundaic region. Glacial data obtained from digital data accompanying Ehlers and Gibbard (2004) and compiled by Andrew Reagan and Michelle Koo.

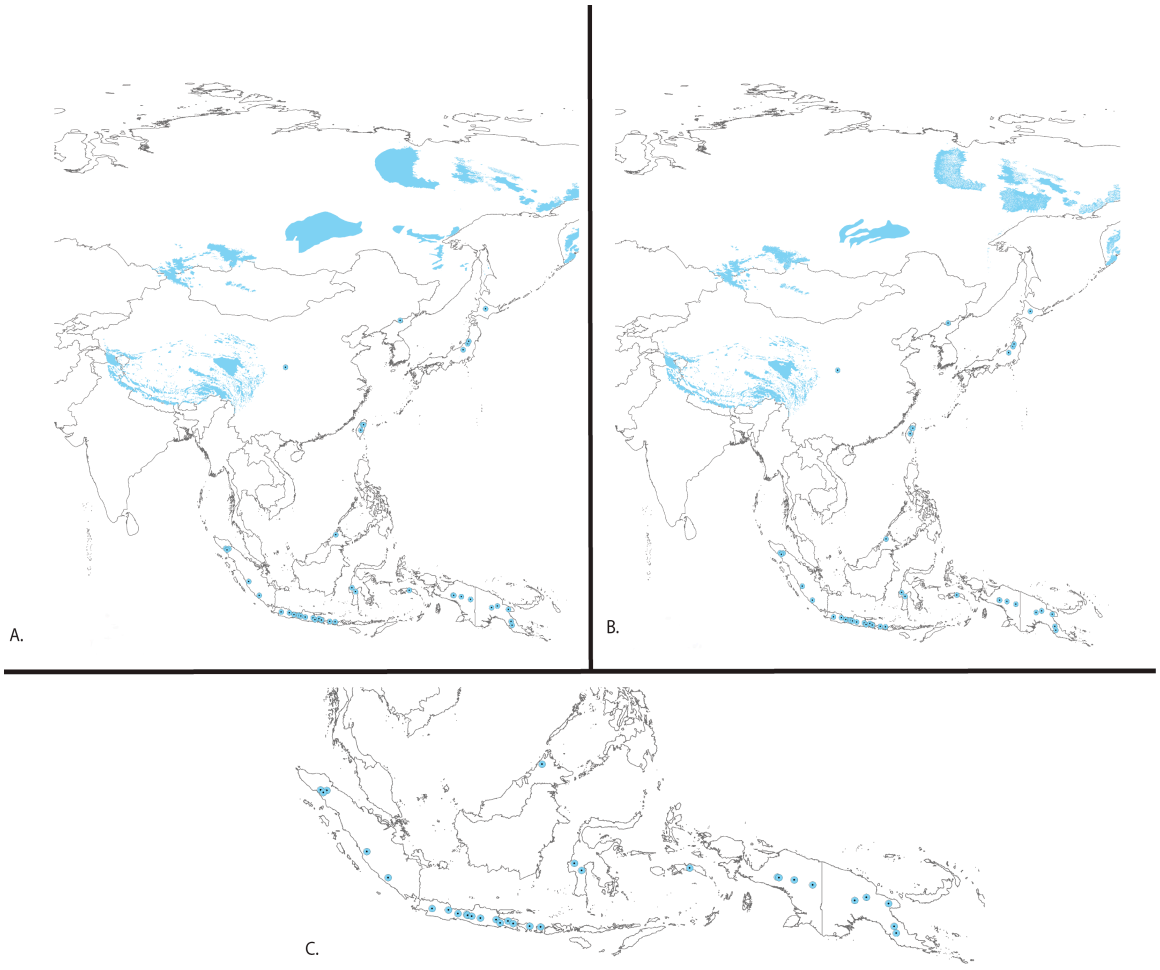

A.
B.
C.

Figure 2. Tropical Forest dynamics. Rows depict the actual and predicted distributions of each of five different tropical forest types through time. The first column is the present actual distribution, the second column is the current predicted distribution, column three is the predicted distribution during the Last Glacial Maximum (21k), and column four is the predicted distribution during the Last Interglacial (120k).

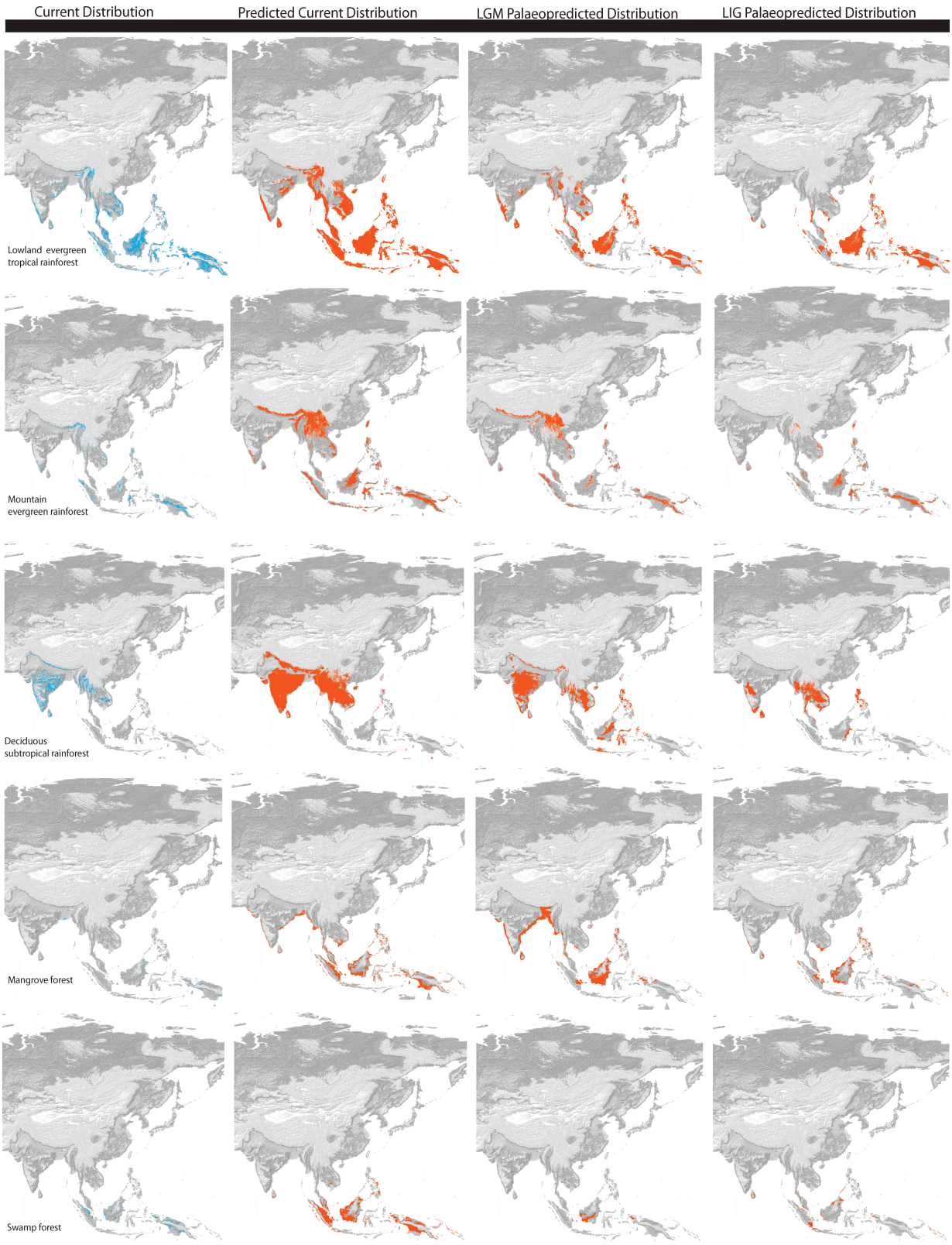

Current Distribution
Predicted Current Distribution
LGM Palaeopredicted Distribution
LIG Palaeopredicted Distribution
Lowland evergreen tropical rainforest
Mountain evergreen rainforest
Deciduous subtropical rainforest
Mangrove forest
Swamp forest

Figure 3. Temperate Forest dynamics. Rows depict the actual and predicted distributions of each of three different temperate forest types through time. The first column is the present actual distribution, the second column is the current predicted distribution, column three is the predicted distribution during the Last Glacial Maximum (21k), and column four is the predicted distribution during the Last Interglacial (120k).

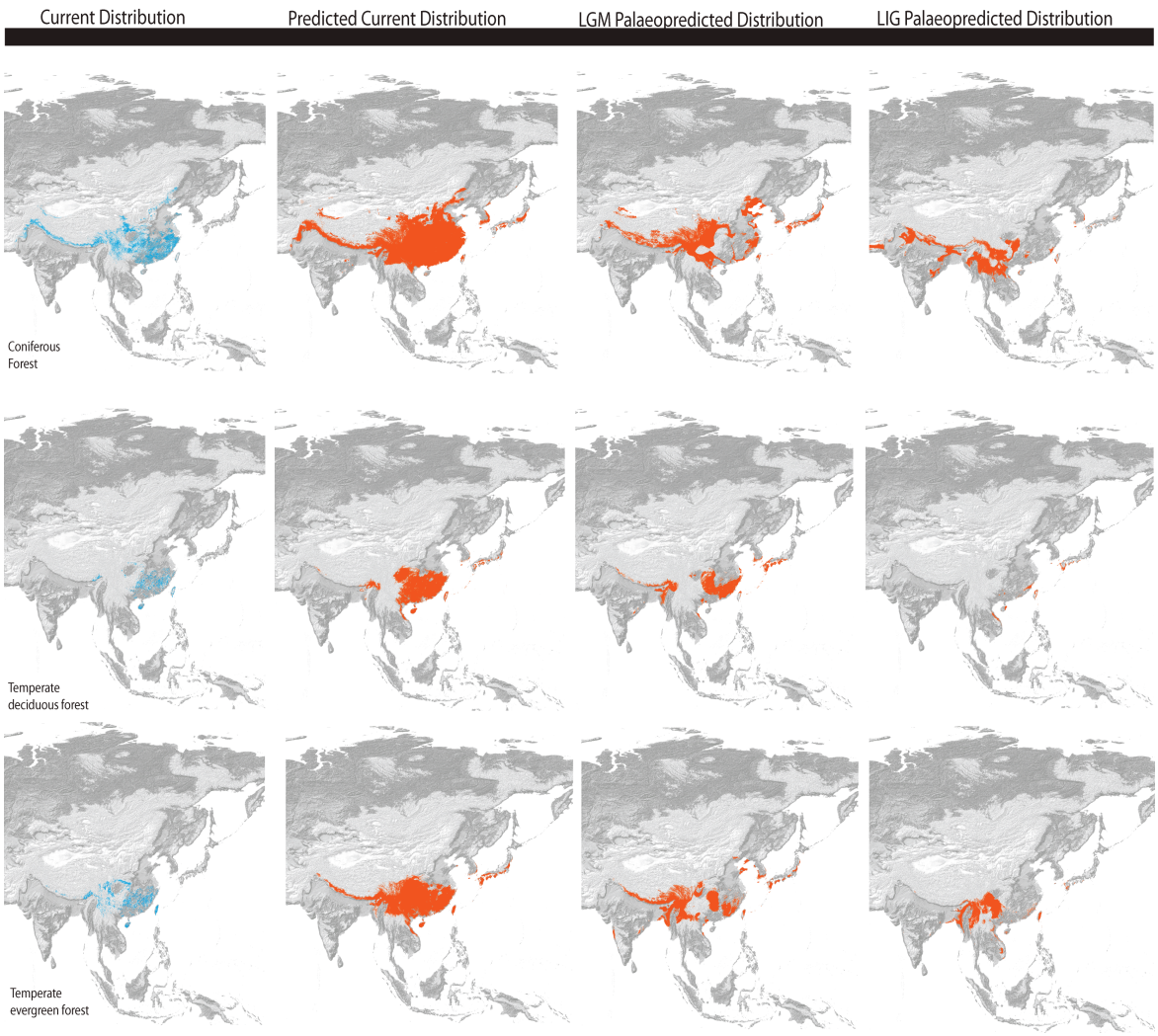

Current Distribution
Predicted Current Distribution
LGM Palaeopredicted Distribution
LIG Palaeopredicted Distribution
Coniferous Forest
Temperate deciduous forest
Temperate evergreen forest

Figure 4. Non-forest habitats A. Rows depict the actual and predicted distributions of each of four non-forest habitats through time. The first column is the present actual distribution, the second column is the current predicted distribution, column three is the predicted distribution during the Last Glacial Maximum (21k), and column four is the predicted distribution during the Last Interglacial (120k).

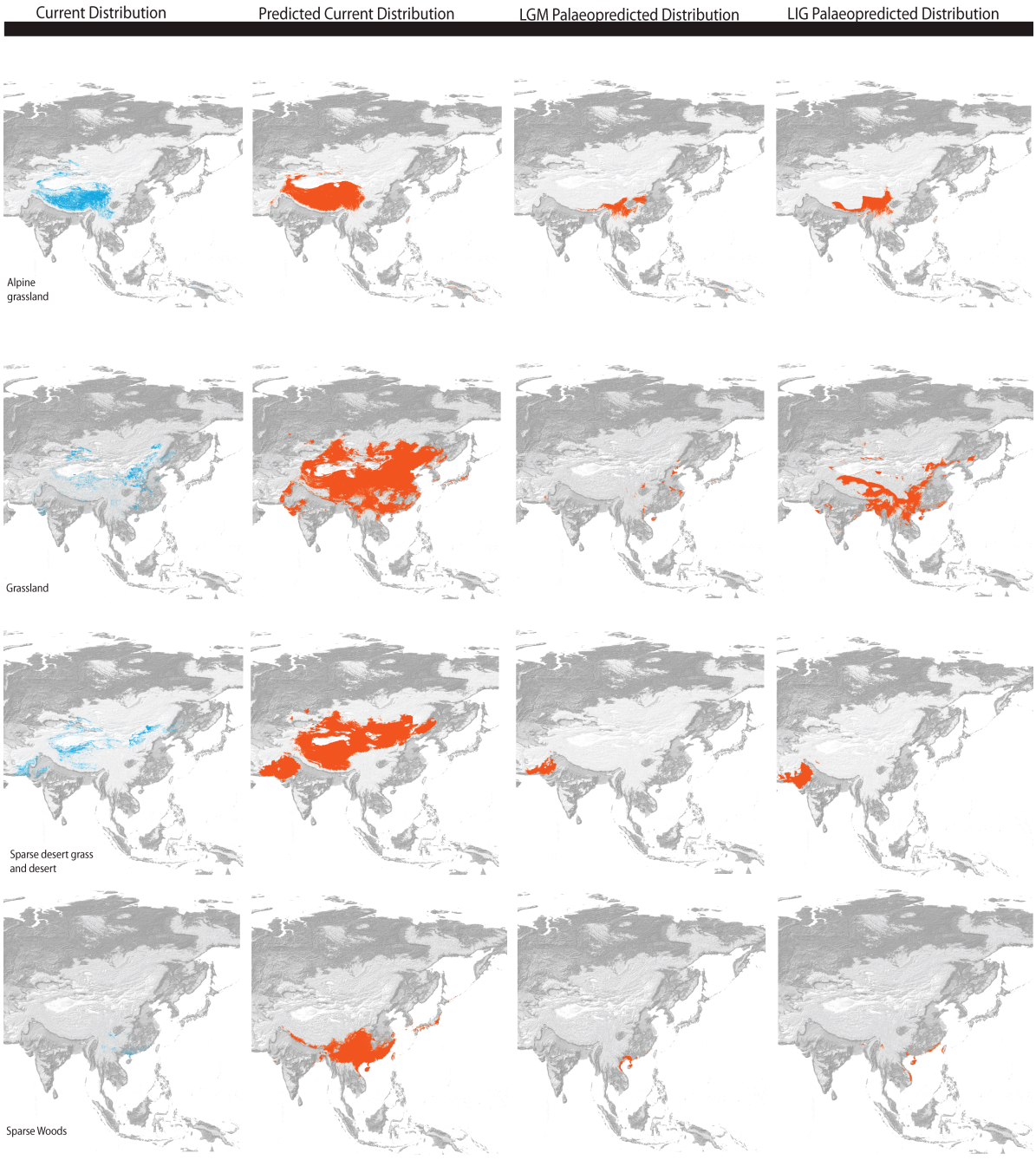

Current Distribution
Predicted Current Distribution
LGM Palaeopredicted Distribution
LIG Palaeopredicted Distribution
Alpine grassland
Grassland
Sparse desert grass and desert
Sparse Woods

Figure 5. Non-forest habitats 2. Rows depict the actual and predicted distributions of each of three non-forest habitats through time. The first column is the present actual distribution, the second column is the current predicted distribution, column three is the predicted distribution during the Last Glacial Maximum (21k), and column four is the predicted distribution during the Last Interglacial (120k).

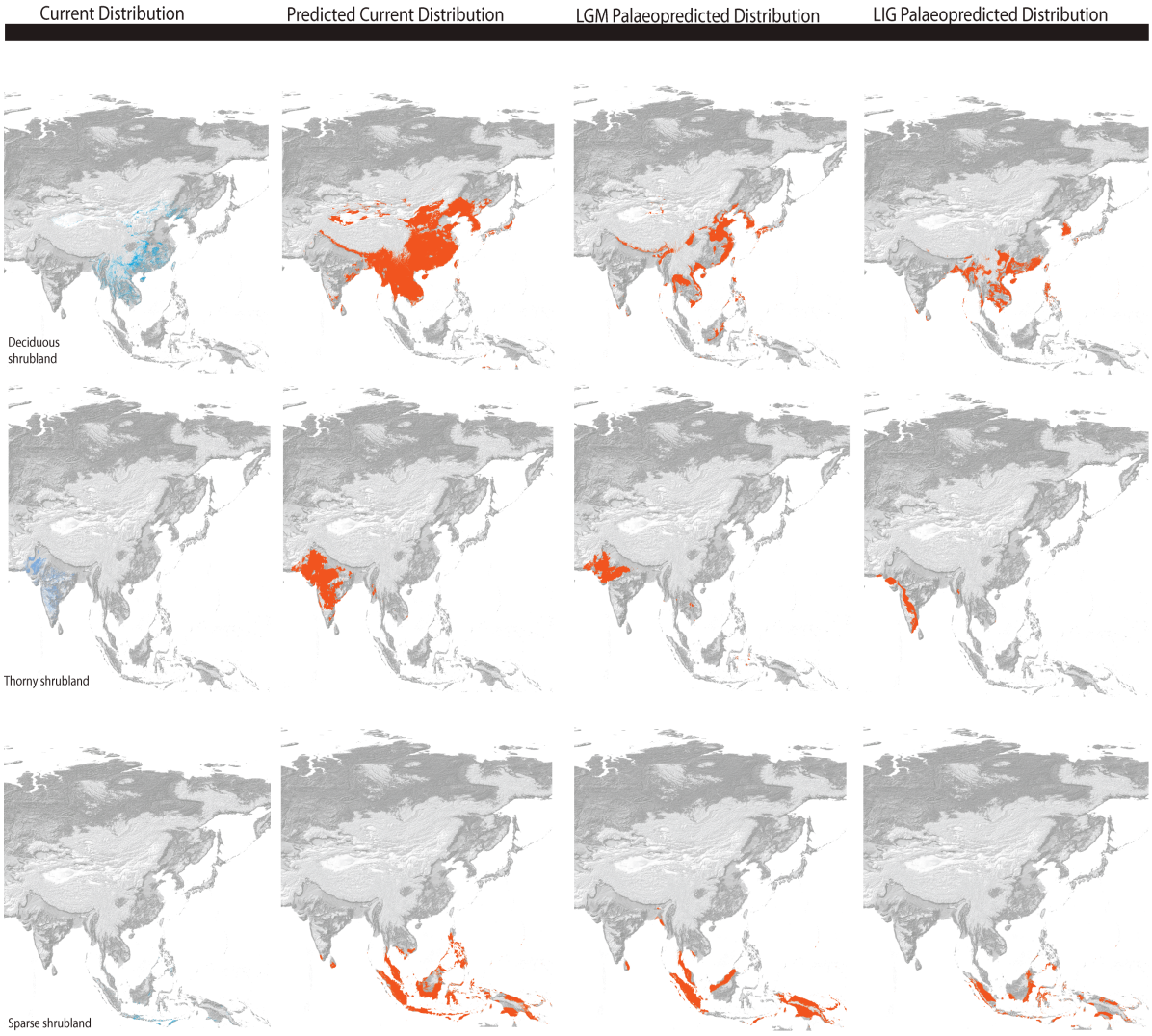
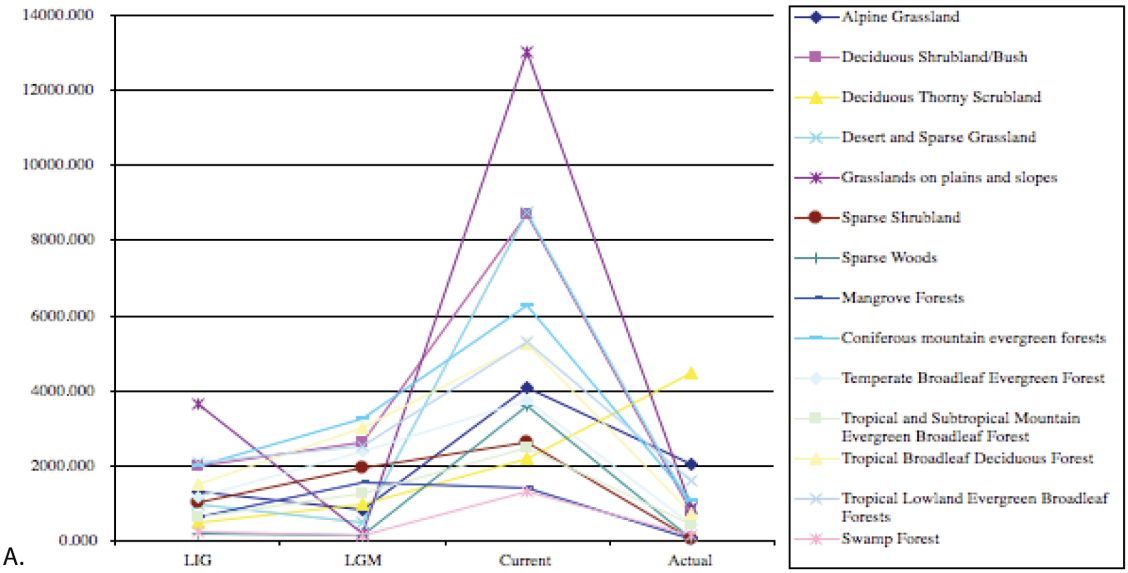

Current Distribution
Predicted Current Distribution
LGM Palaeopredicted Distribution
LIG Palaeopredicted Distribution
Deciduous shrubland
Thorny shrubland
Sparse shrubland

Figure 6. Habitat dynamics for 14 natural habitat types throughout Asia from the last glacial-interglacial cycle. LIG references the Last Interglacial, LGM references the Last Glacial Maximum, current references the predicted distribution of habitats which removes anthropogenic affects from estimates, and actual is the real distribution of each habitat estimated from satellite imaging. Panel A depicts changes in predicted area over time for each habitat, note that deciduous thorny shrubland is the only habitat with a higher actual distribution than current predicted distribution. Panel B depicts the relative extent of each habitat through time.

A.


14000.000
12000.000
10000.000
8000.000
6000.000
4000.000
2000.000
0.000
LIG
LGM
Current
Actual
Alpine Grassland
Deciduous Shrubland/Bush
Deciduous Thorny Scrubland
Desert and Sparse Grassland
Grasslands on plains and slopes
Sparse Shrubland
Sparse Woods
Mangrove Forests
Coniferous mountain evergreen forests
Temperate Broadleaf Evergreen Forest
Tropical and Subtropical Mountain Evergreen Broadleaf Forest
Tropical Broadleaf Deciduous Forest
Tropical Lowland Evergreen Broadleaf Forests
Swamp Forest

B.

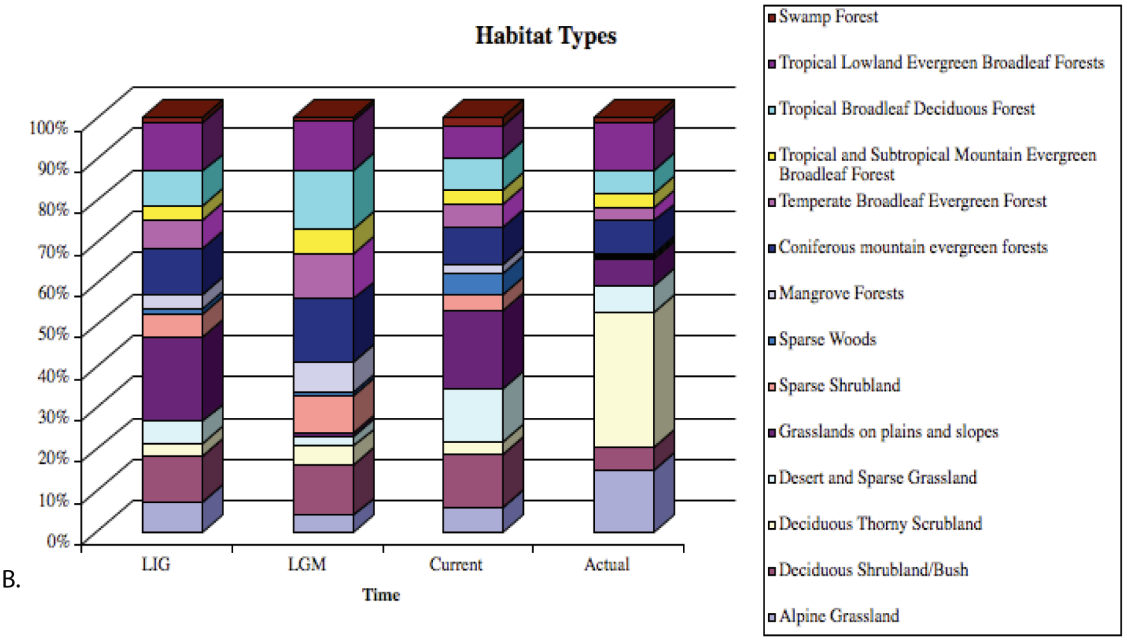
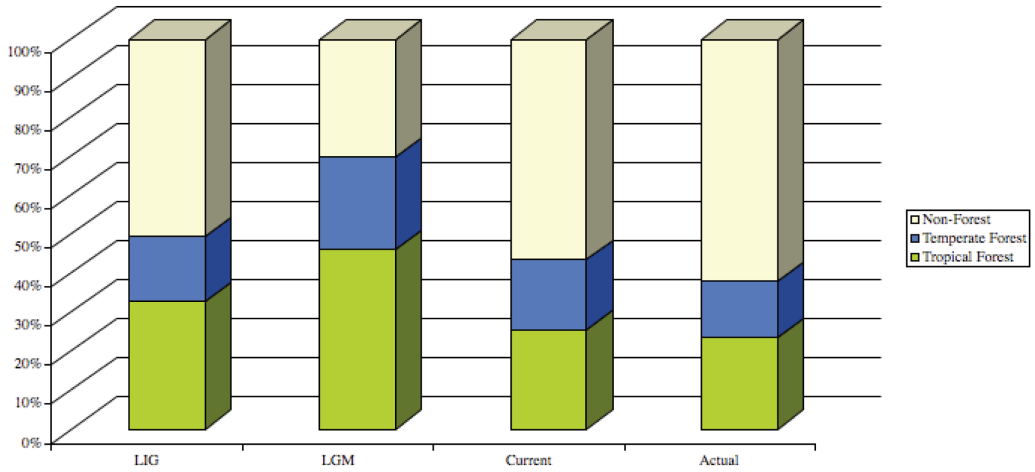

Habitat Types
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
LIG
LGM
Current
Actual
Time
Swamp Forest
Tropical Lowland Evergreen Broadleaf Forests
Tropical Broadleaf Deciduous Forest
Tropical and Subtropical Mountain Evergreen Broadleaf Forest
Temperate Broadleaf Evergreen Forest
Coniferous mountain evergreen forests
Mangrove Forests
Sparse Woods
Sparse Shrubland
Grasslands on plains and slopes
Desert and Sparse Grassland
Deciduous Thorny Scrubland
Deciduous Shrubland/Bush
Alpine Grassland

Figure 7. Relative proportion of general habitat categories through time. Panel A. shows the proportion of tropical forest, temperate forest, and non-forest habitats. Note that since these are natural habitats only, the actual distribution would be a fraction of that from earlier time-periods since anthropogenically modified habitat categories are not depicted. Panel B depicts the dynamics of tropical versus temperate habitats through time.

A.


100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
LIG
LGM
Current
Actual
Non-Forest
Temperate Forest
Tropical Forest

B.

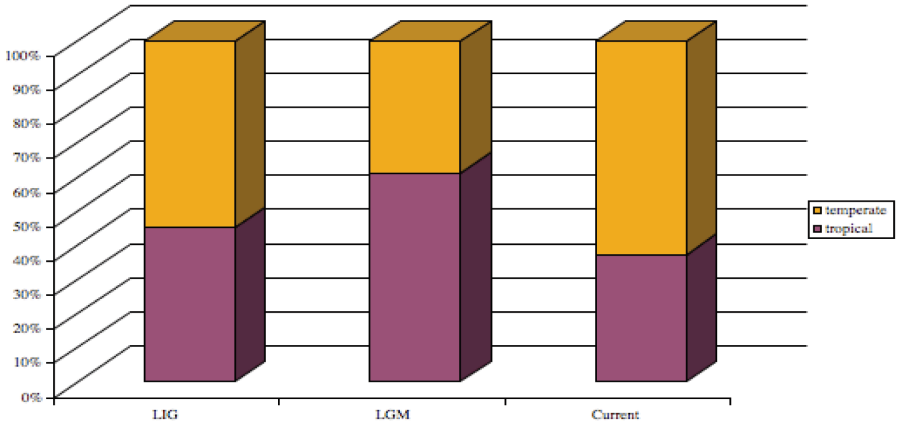

100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
LIG
LGM
Current
temperate
tropical

Table 1. Bioclimatic Variables used in these analyses.

Bioclim 1	Annual mean temperature
Bioclim 2	Mean diurnal range
Bioclim 3	Isothermality
Bioclim 4	Temperature seasonality
Bioclim 5	Max temperature of warmest month
Bioclim 6	Min temperature of coldest month
Bioclim 7	Temperature annual range
Bioclim 8	Mean temperature of wettest quarter
Bioclim 9	Mean temperature of driest quarter
Bioclim 10	Mean temperature of warmest quarter
Bioclim 11	Mean temperature of coldest quarter
Bioclim 12	Annual precipitation
Bioclim 13	Precipitation of wettest month
Bioclim 14	Precipitation of driest month
Bioclim 15	Precipitation seasonality
Bioclim 16	Precipitation of wettest quarter
Bioclim 17	Precipitation of driest quarter
Bioclim 18	Precipitation of warmest quarter
Bioclim 19	Precipitation of coldest quarter

Table 2. The natural habitats evaluated in these analyses, and the AUC score from predictive models.

Habitat Name	Habitat Class	AUC score
Alpine grassland	Non-forest	0.996
Coniferous Mountain Evergreen Forest	Temperate Forest	0.991
Deciduous Shrubland/Bush	Non-forest	0.988
Deciduous Thorny Scrubland	Non-forest	0.854
Sparse Desert Grassland	Non-forest	0.989
Grasslands on plains and slopes	Non-forest	0.982
Mangrove Forests	Tropical Forest	0.999
Sparse Shrubland	Non-forest	0.998
Sparse Woods	Non-forest	0.999
Temperate Broadleaf Deciduous Forest	Temperate Forest	0.996
Temperate Broadleaf Evergreen Mountain Forest	Temperate Forest	0.996
Temperate Meadow	Non-forest	0.991
Tropical-Subtropical Broadleaf Mountain Evergreen Forest	Tropical Forest	0.998
Tropical Broadleaf Deciduous Forest	Tropical Forest	0.993
Tropical Broadleaf Evergreen Lowland Forests	Tropical Forest	0.991
Swamp forest	Tropical Forest	0.943

Table 3. Palynological Data from across Asia.

Original Source	Location	Habitat at the LGM 21KY	Agreement with Predictive Model?
Ansari and Vink 2007	Pakistan, Indus Delta	open scrubland and savannah	Yes
Anshari et al 2001	Indonesia, West Kalimantan Lake Pemerak (Borneo)	lowland evergreen rainforest	Yes
Herzschuh et al 2009	China, Tibet, Koucha Lake	desert grassland/alpine grassland	N/A
Hope 1976	New Guinea, Mt Wilhelm	subalpine grassland	Yes
Hope and Tulip 1994	Indonesia, Irian Jaya (New Guinea Isl), Cyclops Mountains	lowland evergreen rainforest	Yes
Liew et al 1998	Taiwan, Toushe Peat Bog	semi-open forest and lowland grasslands	Yes
Morley & Flenley 1987	Malaysia, Subung-Kuala Lumpur vic.	open	Yes
Nasu et al 1995	Japan, Lake Nojiri	coniferous forest	no
Newsome and Flenley 1988	Indonesia, Sumatra, Danau di Atas	herbaceous swamp / mountain evergreen forest	Yes
Penny 2001	Thailand, Khorat Plateau, Nong Pa Kho	pine-oak forest	Yes
Qin et al 2008	China, Yangtze Delta, Yayoa (HMD core)	freshwater lake or swamp	N/A
Stujits et al. 1988	Indonesia, Sumatra, Danau di Atas	mountain evergreen forest	Yes

Stujits et al. 1988	Indonesia, Java, Situ Bayongbon	mountain evergreen forest	Yes
van der Kaars 1991	Indonesia, Halmahera Island (Molucca Sea core G4-K12p1)	lowland evergreen rainforest	no
van der Kaars 1991	Western new Guinea Ceram Trench cores G5-2053P and 56P)	deciduous forest/ grassland	submerged
van der Kaars and Dam 1995	Indonesia, West Java, Bandung Plain	mountain evergreen forest	Yes
White et al 2004	Thailand, Nong Thale Song Hong	open forest	Yes
White et al 2004	Thailand, Kwan Phayao	pine and oak woodland	Yes
Yu et al. 2000	China, Daluoba	desert	N/A
Yu et al. 2000	China, Kekexili	desert	N/A
Yu et al. 2000	China, Lop Nur K1	desert	N/A
Yu et al. 2000	China, Suzhou	desert	N/A
Yu et al. 2000	China, Tianshuihai	desert	N/A
Yu et al. 2000	China, Hanjiang-CH2	broadleaved evergreen forest	no
Yu et al. 2000	China, Beizhungcun2	grassland	no
Yu et al. 2000	China, Fuxian	grassland	no

Yu et al. 2000	China, Hahai	grassland	no
Yu et al. 2000	China, Nanshan	grassland	no
Yu et al. 2000	China, Nuoergai	grassland	no
Yu et al. 2000	China, Puzhen	grassland	no
Yu et al. 2000	China, Tianshuigoa	grassland	no
Yu et al. 2000	China, Wajianggou	grassland	no
Yu et al. 2000	China, Weinan	grassland	no
Yu et al. 2000	China, Zhabuye	grassland	no
Yu et al. 2000	China, Wasong	broadleaved evergreen forest	no
Yu et al. 2000	China, Erhai(Z27)	broadleaved evergreen forest	Yes
Yu et al. 2000	China, Leizhou Core TY1	broadleaved evergreen forest	Yes
Yu et al. 2000	China, Manxi Core M	broadleaved evergreen forest	Yes
Yu et al. 2000	China, Toushe Lake	broadleaved evergreen forest	Yes
Yu et al. 2000	China, Zhujiang delta PK16	broadleaved evergreen forest	Yes

Yu et al. 2000	China, Zhujiang delta PK19	broadleaved evergreen forest	Yes
Yu et al. 2000	China, Beijing	grassland	Yes
Yu et al. 2000	China, Mengun	grassland	Yes
Yu et al. 2000	China, Yangerzhuang	grassland	Yes
Yu et al. 2000	China, Hulun Lake	taiga	Yes
Yu et al. 2000	China, Jiuzhoutai	taiga	Yes
Yu et al. 2000	China, Jianghua Plain	temperate deciduous forest	Yes
Yu et al. 2000	China, Bailiangdong2	temperate deciduous forest	Yes
Yu et al. 2000	China, Chuangye	temperate deciduous forest	Yes
Yu et al. 2000	China, Erhair (Z18)	temperate deciduous forest	Yes

CHAPTER 3

PALAEO-VEGETATION RECONSTRUCTION AND HABITAT STABILITY ANALYSES PINPOINT POTENTIAL PLEISTOCENE REFUGIA IN ASIA

INTRODUCTION

Climatic fluctuations throughout the Quaternary Period have left an indelible mark on the distribution patterns of many species, with species distributions responding in accord with changing glacial and interglacial climatic conditions (Darwin, 1859; Hewitt, 2000; Hewitt, 2004). When climates change, habitats and biomes respond by contraction and expansion. Under unfavorable conditions, habitats may become highly restricted into refugia, and the organisms that comprise those habitats and the organisms reliant upon those habitats are expected to respond in similar fashion. Pinpointing the location of these refugia, and elucidating the role they play in generating and maintaining biodiversity has been the subject of much research over the past half century, and is especially relevant as we forge into a future of rapid climate change. Refugial areas are predicted as areas with high levels of habitat stability (Graham et al., 2006; Hugall et al., 2002), and these regions, in turn, often harbor high levels of genetic diversity and species endemism (Carnaval et al., 2008; Graham et al., 2006), making them potentially valuable regions for conservation. Here, refugia are defined simply as constricted areas possessing an appropriate habitat type that have been relatively stable through time, and specifically reference macro-refugia rather than microrefugia (refer to Ashcroft 2010 for discussion).

Asia is presently undergoing an extremely high rate of habitat loss and deforestation (Achard and Estreguil, 1995; Achard et al., 2002; Mackinnon, 2005; Sodhi et al., 2004). Although Asia encompasses ten biodiversity hotspots (Mittermeier et al., 2004; Mittermeier et al., 1999; Myers et al., 2000) and is comprised of a wide variety of natural habitat types, the biota of the region is poorly understood. Large-scale multi-country conservation and protection measures are required to help protect critical habitats and accompanying diversity, but targeting regions to protect is difficult given the current state of knowledge (Mackinnon, 2005). Although it has been widely demonstrated that species respond individualistically to climate change, the lack of basic knowledge on species distributions throughout Asia, the large geographic area, and the rapid loss of habitat throughout the regions, leaves large scale habitat modeling as an appealing alternative to individual modeling of species distributions under climate change. This approach assumes that habitats may have had different component species, but that the overall characteristics of the assemblage are consistent through time (Mayle et al., 2004). It should be noted that many of the early definitions of refugia were based on assemblages rather than on individualistic distributions of taxa (Bennett and Provan, 2009; Haffer, 1969).

Satellite imagery has provided insight into the current distribution of habitat types from throughout mainland and insular Asia (Achard and Estreguil, 1995; Bingfang et al., 2004; Stibig et al., 2007), providing a unique opportunity to evaluate how these habitat distributions have changed through time using environmental niche modeling (Wogan ch. 2). Here I identify areas of habitat stability from throughout the past 140 ky years that may have served as refugia, and validate the potential refugia with palynological data.

A similar approach has been previously implemented to evaluate the stability of two forest systems; the Australian Wet Tropics Region (Graham et al., 2006; Hugall et al., 2002; VanDerWal et al., 2009), and the Brazilian Atlantic Forest (Carnaval et al., 2008). Although these systems focused on forest habitats and forest endemics, any region exhibiting habitat stability for a particular natural habitat type may have potentially served as a refuge for species comprising or dependent upon that particular habitat type. In this study, refugial dynamics are examined for both forest and non-forest habitat types throughout mainland and insular Asia. While refugial dynamics are better understood for tropical forest systems in Australia (Graham et al., 2006; Moritz et al., 2009; Phillips et al., 2004a; VanDerWal et al., 2009; Williams and Pearson, 1997) and South America (Carnaval and Moritz, 2008; Colinvaux et al., 2000; Haffer, 1969; Hoorn et al., 2010; Lessa et al., 2003), and for temperate forests in Europe (Hewitt, 2000; Hewitt, 2004; Stewart et al., 2010) and North America (Shafer et al., 2010), much less is known about refugial dynamics in Asia.

Asian refugia have primarily been delimited based on phylogeographic analyses of individual taxa or species groups. As such, the area of extent of any particular study group is generally limited, and in most cases is focused on a single habitat type. This approach to delineating refugia has identified multiple tropical rainforest refugia in the Sundaland region particularly in Borneo and Peninsular Malaysia (Cannon and Manos, 2003; Gathorne-Hardy et al., 2002; Malohlava and Bocak, 2010; Quek et al., 2007). For mainland SE Asia rainforest refugia are poorly known, although recent effort has delineated more broadly defined forest refugia using phylogeographic approaches (Morgan et al., 2011), and parsimony analyses of endemicity (Reddy, 2005). Both non-forest habitat refugia and temperate forest refugia remain poorly characterized.

I evaluate the correlation between regions of habitat stability and regions of climatic stability through time, and between regions of habitat stability and latitude. It might be expected that refugia will primarily be found in places where climate has remained stable through time. A series of palaeoecological analyses by Williams (2007; 2001) have neatly demonstrated the spatial correspondence between non-analog climates and non-analog plant communities, and areas with the greatest turnover in climatic conditions are likely to be regions where diversity is most impacted (Ackerly et al., 2010; Jackson and Overpeck, 2000). Loarie et al. (2009) examined the global velocity of future predicted climate change, demonstrating that high latitude flat regions had the highest global velocities, and Sandel et al. (2011) linked regions of low palaeo-climatic velocity with high rates of vertebrate endemism. Weir and Schluter (2007) provide evidence that increased turnover caused by greater environmental perturbation at high latitudes underlies the latitudinal diversity gradient (but see Marshall (2007) for an alternative interpretation), and Jansson and Dynesius (2002) argue that orbitally driven climate changes drive global diversity patterns. These studies suggest that stable climates are important in structuring global diversity patterns.

Early concepts of refugia rest on the assumption that stable tropical forest regions will have higher species richness (Haffer, 1969; Mayr and O'Hara, 1986) and can thus explain distributional patterns of contemporary biotas. This assumption has only

recently been validated as new tools have become available for spatial modeling and genetic analyses. It has now been demonstrated that tropical forest refugia do in fact have higher diversity than non-stable forest regions in the Australian Wet Tropics, and in Brazil's Atlantic Wet Tropics (Carnaval and Moritz, 2008; Graham et al., 2006; Williams and Pearson, 1997), and the expectation is that this will also be true for other refugia. Since little is known about the role of refugia in regards to Asian biodiversity patterns, I examine patterns of vertebrate species richness as they relate to habitat stability. Lastly, I evaluate the impact of stability on vertebrate community structure.

METHODS

Climate Stability: Since climate and climate stability are thought to be critical components in determining the distribution of stable habitats I examined climate change from the LIG to present. Climate stability surfaces were calculated across Asia using the Standard Euclidean Distances (SED) implemented as in Williams et al (2007). This measure helps pinpoint spatially where climates are similar and different, and provides some insight into the stability of climate space through time. However, it should be noted that since this measure is calculated using two time points, climate variation during the interval is not explicitly taken into account. Here, SEDs were calculated between the LIG and present, the LIG and the LGM, and the LGM and present using four bioclimatic variables: annual mean temperature, annual precipitation, temperature seasonality, and precipitation seasonality. Next, novel and disappearing climates were calculated from the LIG to present, LIG to the LGM, and the LGM to present using the two-dimensional histogram method developed in Ackerly et al. (2010), but with some modifications. Since the area under investigation here is very large, climate space was sub-sampled using the same 5000 random points above. Since only a single climate reconstruction is available for the LIG, and thus it is not possible to calculate the standard deviation across scenarios as in Ackerly et al. (2010), the number of bins was set to 15 for temperature and 10 for precipitation.

Palaeo-predictive models: Land cover classifications derived from remote sensing satellite data form the foundation of our understanding of contemporary habitat distribution in Asia (Bingfang et al., 2000; Bingfang et al., 2004; Stibig et al., 2007). Species distribution models for three time periods (current, last glacial maximum, and the last interglacial) were generated using Maximum Entropy models (Phillips et al., 2006; Phillips et al., 2004b). A threshold approach was used to determine presence and absence for each habitat at each of the predicted time periods. Additional details for the modeling methods can be found in Wogan (chapter 2).

Habitat Stability Analyses: It has been shown that habitat stability through time is an important predictor of distribution patterns for low dispersal organisms (Graham et al., 2006). To estimate the stability of each habitat through time, two different approaches were used. First, palaeolayers were overlapped to evaluate which regions have persisted through time by summing the models for each unique habitat type after thresholds have been applied. The number of modeled time slices determines the value range for each

pixel. Three time slices were utilized so each pixel could have a sum of 0, 1, 2, or 3. A value of 0 indicated that the habitat was not predicted to occur at any time slice with a likelihood greater than the applied threshold value. A max value indicated that the habitat was predicted to be present in each time slice at a value greater than the applied thresholds. In the second method, the geometric mean was calculated. The geometric mean minimizes false inferences of suitability across time periods by eliminating pixels predicted to have a zero value at any time period from the stability surface (Bell et al., 2010). To obtain an estimate of area, the geometric mean for each pixel was then classified as present or absent based using the geometric mean of the thresholds derived from each time slice. The area was then calculated by multiplying the total number of pixels predicted as presences by the area of each pixel (approximately 1 km^2). To ensure that model over-prediction was not affecting the interpretation of refugial areas, model overlap was calculated between all three major habitat classifications. It should be noted that this approach assumes that refugia do not move in contrast to recent methods that implement a shifting refuge approach (Graham et al., 2010)

Vertebrate Species Richness: Spatial data for species distributions for mammals, birds, amphibians, and reptiles were obtained from the IUCN Red List and BirdLife International (IUCN, 2011). Species richness was calculated for each km^2 as the total number of overlapping distributions. To test the relationship between species richness and habitat stability, random points were generated within the stable areas of each habitat type and the species richness values were obtained for each vertebrate group independently and combined. To establish a comparative dataset, a set of 5000 random points was generated across all of Asia and species richness values were obtained for each vertebrate group independently and combined. Data were not normally distributed, and so they were log transformed. T-tests were used to evaluate the underlying distributions of species richness drawn from stable areas for each habitat type and the species richness generated from random points. In order to better understand the relationship between species richness and stability, species richness and biome area, and species richness and latitude, a suite of simple linear regressions were implemented. These analyses used the different types of habitat refugia classified into the broad habitat categories of tropical forest, temperate forest, and non-forest refugia. Residuals obtained from regression analyses were utilized to remove the affect of the total predicted area of each habitat when examining the relationship between species richness and stability.

Community Composition: Jaccard and Euclidean distance Indices were calculated for each random point using a presence-absence species matrix. Both measures can be utilized with presence-absence data alone and are not reliant on abundance data to calculate community dissimilarity. Mantel tests and Analyses of Similarity (Anosim) were used to examine the correlative relationship between climate stability (CLS), habitat stability (HIS), species richness (SR), latitude (LAT), and community composition (CC) dissimilarity matrices using the BiodiversityR package (Kindt and Coe, 2005). Mantel tests were performed with 500 permutations using Pearson correlations. The r statistic calculated by the Mantel test can be interpreted as an autocorrelation measure between the distance matrices. Environmental dissimilarity matrices were generated utilizing Euclidean distances. The categorical variable habitat stability (HIS) was evaluated in

relation to other variables using Anosim and Mantel tests. Anosim analyses were conducted with 500 permutations using Pearson correlations. The Mantel test utilized Gower distances for the stability variable. The R statistic from the Anosim analysis can be interpreted as similar to a correlation coefficient (Kindt and Coe, 2005).

RESULTS

Climate: Climate between the LIG and now is relatively stable throughout the tropics, with the greatest degree of change concentrated across China and southeast India (Figure 1A). Climate changes between the LIG and LGM were highly concentrated in the mid-latitude tropics, with Myanmar, India, Bangladesh, and southern China undergoing dramatic climate changes (Figure 1B). These changes were driven by changes in precipitation and the seasonality of precipitation, suggesting changing monsoonal conditions. Between the LGM and now, climate changes were largely concentrated in the same geographic area, and were again related to drastic changes in annual precipitation and precipitation seasonality (Figure 1C). While the most dramatic changes were at higher tropical latitudes, Sulawesi, Borneo, and Peninsular Malaysia were also affected by climate change (Figure 1).

Univariate histograms of annual mean temperature for current and LIG climatic conditions reveal that a larger area of Asia is warmer now (AMT >25 °C) than during the LIG. There are also more very dry areas currently as compared to the LIG (Figure 2). During the LGM, there were more cold areas than now, although a large portion of region maintained warm annual temperatures. The entire region was markedly drier than either contemporary or LIG climates (Figure 2). Two-dimensional histograms reveal that there are several novel climate combinations not in existence during the LIG. These novel climates (present now but not analogous to climate combinations during the LIG) are concentrated in hot dry and cold dry realized climate combinations (Figure 2). A single cold-wet climate combination disappeared between the LIG and now (Figure 2). Overall, wet climates shrank, and dry climates expanded, cool climates shrank, and hot climates dramatically expanded. The LGM saw the emergence of several novel climates, while all of the climates were very dry, here was an expansion of novel cold dry climates in which the annual mean temperature was below –20 °C. At the same time, several LIG climates disappeared in the LGM. Primarily, warm wet climates were lost.

Refugial Reconstructions and Dynamics: Refugia could be identified for most of modeled habitats (Figure 3). There was overlap among all of the habitat classes, with some geographic regions having conflicting predicted habitat types (discussed in detail below) (Figure 4). A little over 3% of the total area of Asia has remained stable since the last interglacial. The habitat type that has by far had the most stability is tropical lowland evergreen broadleaf forest (Figure 5). All habitats except one (alpine grassland) exhibit higher stability between the LGM and present than between the LIG and the LGM. Perhaps this is not surprising given the difference in time span, but this implies that the refugia examined here are primarily determined by the degree of habitat stability during the interglacial-glacial transition rather than the glacial-interglacial transition. Forest habitats are much more stable than non-forest habitats (Figure 5). Tropical forests

habitats, were highly stable between the LIG and LGM with almost 50% of the total stable area occupied by tropical forest habitats, however there was less stability between the LGM and present. Conversely, temperate forests had greater stability between the LGM and present than they did from the LIG to the LGM. Stable non-forest habitats comprise much less total area than forest habitats at all times, but were more stable during the LIG to LGM than the LGM to present (Figure 5). Tropical forest refugia tend to be greater in area than either temperate forest refugia or non-forest refugia (Figure 5).

Tropical Forest Refugia

Five tropical forest types were modeled, and refugia were identified for all. Models suggest that tropical lowland evergreen broadleaf forest refugia existed in Sri Lanka, the Western Ghats, India and the Andaman Islands, India. In Southeast Asia, refugia were identified in central Vietnam and Lao, southern Vietnam, Cambodia, and southern Thailand, peninsular Myanmar and the Myeik archipelago, western peninsular Thailand, the eastern coast of peninsular Malaysia and southern peninsular Thailand, and southern peninsular Malaysia, and in insular SE Asia including Palau Bangka, Biliton, Sumatra, Siberut, Simeulue, Java, much of Borneo, Seram, New Guinea, and New Guinea, and many islands of the Philippines. One final refuge was identified on Taiwan Island.

Mountain evergreen forest is predicted to occur in a series of small refugia across eastern Myanmar, Yunnan China, Vietnam, Laos, and Cambodia. Three small refugia are predicted in peninsular Myanmar and Thailand, and more extensive refugia are predicted for peninsular Malaysia. In insular SE Asia, some mountains in Sumatra, western Java, Borneo, the Philippines, and New Guinea are predicted to have served as refugia, with the largest contiguous areas across the interior mountains of New Guinea, and on Borneo. Taiwan is also predicted to have held mountain evergreen forest refugia.

Tropical broadleaf deciduous forest refugia are identified in South Asia in the Gujarat, Andhra Pradesh, Maharashtra, Karnataka, Kerala and Tamil Nadu States of India, and two refugia are identified in Sri Lanka. In SE Asia, a series of refugia are predicted in Myanmar, northern Thailand, eastern Vietnam, and southern Cambodia. A large refuge is predicted in eastern Thailand, Cambodia, and Laos. The Philippine Islands of Luzon, Mindoro, Panay and Samar are also predicted to have maintained forest refugia. Swamp forest refugia are more locally delineated with three stable refuges found in Borneo, Pulau Bankga, and New Guinea.

Mangrove forest refugia are predicted in eastern and western India and Sri Lanka. In Southeast Asia, they are predicted in Myanmar, Cambodia, peninsular Thailand, and peninsular Malaysia. In insular Asia, they are predicted throughout the smaller Sunda shelf islands, and on Sumatra, one of the Mentawai Islands, Borneo, some of the small offshore islands surrounding Sulawesi, the Moluccas, New Guinea, and the some of the Philippine Islands.

Temperate Forest Refugia

Three temperate forest types were modeled: lowland temperate evergreen broadleaf forests, mountain temperate evergreen broadleaf forests, and coniferous evergreen forests. Lowland temperate evergreen broadleaf forest refugia were restricted to northern Vietnam, eastern China, Taiwan, and southern Japan. Mountain temperate evergreen broadleaf forest refugia were identified across northern Myanmar and Yunnan, with isolated patches through out southeastern China, Taiwan, and southern Japan. Coniferous evergreen forests exhibited dramatic shifts from glacial to interglacial periods, but a narrow band of refugia are predicted along the southern China mountains, and across the southern edge of the Himalayas into Pakistan. Isolated refuges are predicted in Guizhou Province, China, Northern Vietnam, South Korea, and Japan.

Non-Forest Refugia

Nine non-forest habitats were modeled, of these stable habitats were identified for seven habitat types. Alpine grasslands have a wide distribution throughout the Himalayas and mountain ranges of China, and their predicted current distribution closely matches their actual distribution, although is predicted to be more extensive. Stable areas are largely confined to the mountains of Sichuan and Xizang Zizhiqu (Tibet), and comprise a relatively small fraction of their current range. Grasslands are widely distributed in the region, but a very tiny fraction of that total area has remained stable. A few small stable habitat patches are identified in Northern Vietnam, Hainan, and eastern China (Sichuan, Guangxi, New Mongol Zizhiqu). It should be noted that grasslands were very difficult to model in these analyses (Wogan, chapter 2). Temperate meadows are distributed in NE China, palaeo-distributions of this habitat type could not be determined at either the LGM or LIG and thus stable areas were not identifiable. Sparse woods are currently identified primarily from eastern China. A few predicted stable regions are predicted in Hainan, Guangdong, Guangxi, and northern Vietnam. Sparse shrubland habitat has remained stable in Sumatra, western Java, and New Guinea, and some of the smaller Indonesia Islands such as Timor. Stable regions of deciduous shrubland were identified on the Korean peninsula, Fujian, Guangdong, Guangxi, Hainan, Vietnam, and Thailand. Deciduous thorny shrublands have an extensive contemporary distribution, but very little of that distribution has been stable through time. Only in Gujurat India and Pakistan are stable areas identified. Although contemporary deserts are widespread in Asia, palaeo-predictive methods were unable to identify the distribution of this habitat through time, so stable areas could not be identified. However, desert and sparse grassland habitats were successfully modeled and stable regions were identified in Pakistan.

When grouped by major habitat category, overlap between predicted refugia was detected (Figure 4). Tropical forest and temperate forest refugia overlapped in two areas: central Taiwan and the Myanmar-Yunnan border. Overlap between tropical forests and non-forest refugia was more extensive, with overlap predicted in South India, the Myanmar-Yunnan border, Peninsular Myanmar, the Andaman Islands, the Vietnam-Laos border, the Mekong Delta, Sumatra, Borneo, and New Guinea Island. Overlap between temperate forest and non-forest refugia occurred at the Myanmar-Tibet border, Myanmar-Yunnan border, the Vietnam-China border, Taiwan, and in eastern China. All three overlapped on the Myanmar-Yunnan border and in Central Taiwan.

Locations of Refugia in reference to latitude and climate : Refugia and latitude share a strong relationship, with most refugia concentrated in low latitude regions, although some high latitude stable regions were identified (Figure 6). Pearson correlations recovered a significant negative correlation between stable refugia and latitude (R^2= -0.3007146, p-value < .01). Climate stability and latitude are correlated (r=0.169, p-value<.001) (Table 1). Despite the expectation that the changing climate surface would result in a tight relationship between stable climatic areas and stable refugia, Pearson were non-significant (R^2= 0.01116703), however, Mantel tests of these same variables were significant (r= 0.09991, p-value <0.01) (Table 1).

Vertebrate Species Richness: Common patterns of species richness exist among all four vertebrate groups. The most species rich areas for all vertebrate groups are in the tropics and often centered on the Sundaland landmasses, particularly Borneo and Peninsular Malaysia (Figure 7). Not surprisingly, species richness decreases with increasing latitude varies for all groups, and latitude explains a large proportion of the observed variation in species richness across the region, although the relationship is tighter for lower dispersal vertebrates (amphibians) than for highly vagile groups (birds) (Figure 8).

Converse to the expectation that habitat area and species richness would have a tight positive relationship, a negative but non-significant relationship was recovered in these analyses for all habitats combined, and for the three habitat classes for both log transformed and untransformed data (Figure 9). The relationship between habitat area and species richness at the LGM and LIG revealed the same trend (Figure 9).

Refugial regions for all habitats except alpine grasslands, deciduous shrublands, and deciduous thorny shrublands had species richness values that are statistically significantly different than those drawn from random points across Asia (Table 2). Species richness values within stable areas also differ significantly from species richness values in unstable areas for each vertebrate group (Figure 10). Tropical forest refugia had the highest levels of mean species richness for all vertebrate groups combined, followed by non-forest refugia and then temperate forest refugia (Figure 11). Amphibian species richness varied greatly in non-forest habitats, while in temperate and tropical forests species richness values were more clustered. In birds, tropical forests were much more species rich than either temperate forests or non-forest habitats. In mammals, a great deal of variation exists in non-forest habitats, but the mean species richness value was higher in these habitats than in temperate and tropical forests. In reptiles conversely, species richness was highest in tropical forests.

Habitat stability and species richness share a tight positive relationship for all terrestrial vertebrate groups, after removing the affects of area, stability explained a large proportion of the remaining variance (R^2=0.52 for all vertebrate groups), and is particularly important in explaining species richness for amphibians (R^2=0.7474) and mammals (R^2=0.8456). (Table 3, Figure 12). However, species richness values within stable regions vary from 5 to 448 combined vertebrate species (Figure 12), with species richness within tropical forest and temperate forest habitats demonstrating a greater

response to stability than observed in non-forest habitats (Figure 12). The lowest vertebrate species values were found in desert grassland refugia in Pakistan, but stable habitats in lowland New Guinea, central India, and the Korean Peninsula also exhibited low species richness values. While this may reflect actual diversity patterns, meaning that these are simply low diversity habitats, this pattern could also reflect sampling limitations. New Guinea for example has very high diversity levels with a rapid pace of new species discovery, however the lowland stable areas appear to be low diversity based on the available distribution data.

Community Composition: The community composition analyses did not reveal any significant correlation between community dissimilarities that accord with historical habitat stability or climate stability between the LIG and now (Table 1). Interestingly, a slightly negative (though non-significant) correlation was recovered for community composition and climate stability (Table 1).

DISCUSSION

Only a small fraction of the Asian landscape is predicted to have remained stable over the last interglacial-glacial-interglacial cycle. Stable regions that remain stable over an entire cycle of climate change are posited to be of critical importance in structuring the diversity of our contemporary biota (Stewart et al., 2010). Protection of such areas may provide a much needed buffer against future climate change. Despite the emergence of novel climates and disappearance of climates between the LIG and now, stable habitats are predicted to have persisted. The relationship between climatic stability and habitat stability remains unclear with statistical analyses supporting conflicting hypotheses, although the Mantel tests, which support a correlative relationship between climatic and habitat stability, are generally thought to be less subject to spurious correlations than simple Pearson correlations (which did not find a correlative relationship between the two). The finding that climate stability and latitude are correlated is not unexpected, although latitude only explains a small portion of the observed pattern. Also not unexpected is the strong association between latitude and habitat stability since it has been long established that tropical forest systems emerged earlier and have persisted longer non-tropical systems. (Beerling and Woodward, 2001; Morley, 2000; Morley, 2007; Scotese, 2003; Willis and McElwain, 2002; Ziegler et al., 2003). The analyses here demonstrate that the majority of predicted Asian refugia are centered in the tropics, but that there are also highly stable temperate refugia.

The analyses here support the role of stability as an important factor influencing alpha level diversity. Novel to this study is the finding that this is true not only for tropical forest habitats, but also for temperate forest, and non-forest habitats. For all but three of the habitats evaluated here, stable areas had statistically significantly higher levels of species richness than non-stable areas, and this pattern held across all of the terrestrial vertebrate groups examined here. Within the refugia, species richness varied greatly among vertebrate groups with respect to habitat category. Stability explained a large component of species richness for all vertebrate groups across all habitats.

The lack of correlation between community composition dissimilarity matrices and historical habitat stability or historical climate stability may mean that neither of these factors are relevant in community assemblages in Asia, or more likely, suggests that the vertebrate distribution data utilized here are too course to adequately capture this facet of community composition. The vertebrate distribution data are large polygons constructed to encompass the ranges of species, however, within any of those polygons there are areas where habitat is unsuitable. Use of more fine scale point estimates of species presence and absence would likely provide greater insight into the relationship of community composition to historical environmental parameters such as habitat stability and climate stability, although for most Asian species, data at this fine resolution are not currently available.

Of the refugia identified here, the tropical lowland rainforest refugia have received the most research attention. Their locations have been posited using multiple lines of evidence, however the outcomes have often been conflicting and a healthy debate over the location of these forest refugia exists (Cannon et al., 2009; Heaney, 1991; Morley, 2000; Slik et al., 2011). Phylogenetic analyses of termites suggests Quaternary tropical rainforest refugia in the Mentawai Islands, northern Sumatra, along the base of the Barisan Mountains in Sumatra, Northern Sarawak, Brunei, Sabah, and eastern Kalimantan (Gathorne-Hardy et al., 2002). Analyses of neotenic beetles alternatively suggest mid-Tertiary to Quaternary forest refugia in Borneo, the mountains of peninsular Malaysa, northwest Sumatra, and northeastern Indochina (Malohlava and Bocak, 2010). Evidence from ants suggests rainforest refugia in Borneo and on the mountains of Sumatra and peninsular Malaysia (Quek et al., 2007). Palynological evidence suggests the persistence of extensive lowland rainforests on Borneo, Sumatra, and the Mentawai Islands, and across the Sundashelf onto the Malay Peninsula, and a smaller refuge on Java (Morley, 2000).

Forest refugia (not restricted to rainforest) delineated using endemicity analyses of forest birds, are predicted in the Western and Eastern Himalayas, Assam/Burma, Yunnan, North Thailand, North Vietnam, Sichuan. South China, Hainan. Central Vietnam. South Vietnam, Indochina, Isthmus of Kra, Malay Peninsula, the islands of Sumatra and Java, and part of Borneo (Reddy, 2005; Reddy, 2008). Biogeographical analyses of colobine monkeys suggests refugia in Sri Lanka and southwest India, eastern Indochina and China, the northern tip of Sumatra, Siberut Island, the western edge of Java, and northeastern Borneo (Brandon-Jones, 1996). Recent analyses of *Anopheles* species across Asia found support for forest refugia in the southern China mountains and on the Myanmar-India border (Morgan et al., 2011). The models generated here are in general agreement with many of refugia delimited via phylogeographic and population genetic analyses.

Habitats in Asia are suffering from an astounding amount of anthropogenic disturbance, and much of the Earth's biodiversity lies in this region, making conservation efforts a high priority. The models presented here find evidence for refugia that may provide insight for guiding conservation priorities across Asia. Additional validation from phylogeographic and population genetic analyses for additional taxa from tropical

forest, temperate forest, and non-forest Asian refugia is needed to fully characterize the structure of Asian biodiversity and the role of these putative refugia in it's generation and maintenance.

REFERENCES CITED

Achard, F., and C. Estreguil. 1995. Forest classification of Southeast Asia using NOAA AVHRR data. Remote Sensing of Environment.

Achard, F., H. Eva, H.-J. Stibig, P. Mayaux, J. Gallego, T. Richards, and J. Malingreau. 2002. Determination of deforestation rates of the world's humid tropical forests. Science 297:999-1002.

Ackerly, D. D., S. R. Loarie, W. K. Cornwell, S. B. Weiss, H. Hamilton, R. Branciforte, and N. J. B. Kraft. 2010. The geography of climate change: implications for conservation biogeography. Diversity and Distributions 16:476-487.

Beerling, D. J., and F. I. Woodward. 2001. Vegetation and the Terrestrial Carbon Cycle: Modelling the first 400 million years. Cambridge University Press, Cambridge.

Bell, R. C., J. L. Parra, M. Tonione, C. J. Hoskin, J. B. Mackenzie, S. E. Williams, and C. Moritz. 2010. Patterns of persistence and isolation idnicate resilience to cliamte change in montane rainforest lizards. Molecular Ecology 19:2531-2544.

Bennett, K. D., and J. Provan. 2009. What do we mean by 'refugia'? Quaternary Science Reviews 27:2449-2455.

Bingfang, W., X. Wenting, H. Huiping, and Y. Changzhen. 2000. The Land Cover Map for China in the year 2000European Commision Joint Research Center, GLC Database.

Bingfang, W., X. Wenting, H. Huiping, Y. Changzhen, and X. Wenbo. Year. Combining Spot4-Vegetation and Meterological Data Derived Land Cover Map in China *in* Geoscience and Remote Sensing Symposium, IEEE International IGARSS '04. Proceedings. 2004 2004:4.

Brandon-Jones, D. 1996. The Asian Colobinae (Mammalia: Cercopithecidae) as indicators of Quaternary climatic change. Biological Journal of the Linnean Society 59:327-350.

Cannon, C. H., and P. S. Manos. 2003. Phylogeography of the Southeast Asian stone oaks (Lithocarpus). Journal of Biogeography 30:211-226.

Cannon, C. H., R. J. Morley, and A. B. G. Bush. 2009. The current refugial rainforests of Sundaland are unrepresentative of their biogeographic past and highly vulnerable to disturbance. Proceedings of the National Academy of Sciences 106:11188-11193.

Carnaval, A. C., M. Hickerson, C. Haddad, M. T. Rodrigues, and C. Moritz. 2008. Stability predicts genetic diversity in the Brazialian Atlantic Forest Hotspot. Science 323:785-789.

Carnaval, A. C., and C. Moritz. 2008. Historical climate modelling predicts patterns of current biodiversity in the Brazilian Atlantic forest. J Biogeography 35:1187-1201.

Colinvaux, P., P. De Oliveira, and M. Bush. 2000. Amazonian and neotropical plant communities on glacial time-scales: the failure of the aridity and refuge hypotheses. Quaternary Science Reviews 19:141-169.

Darwin, C. 1859. On the origin of species by means of natural selection, or the preservation of favoured races in the struggle for life, first edition. John Murray, London.

Gathorne-Hardy, F., R. Davies, P. Eggleton, and D. Jones. 2002. Quaternary rainforest refugia in south-east Asia: using termites (Isoptera) as indicators. Biological Journal of the Linnean Society.

Graham, C., C. Moritz, and S. Williams. 2006. Habitat history improves prediction of biodiversity in rainforest fauna. Proc. Natl. Acad. Sci. U.S.A. 103:632-6.

Graham, C. H., J. VanDerWal, S. J. Phillips, C. Moritz, and S. E. Williams. 2010. Dynamic refugia and species persistence: tracking spatial shifts in habitat through time. Ecography 33:1062-1069.

Haffer, J. 1969. Speciation in Amazonian forest birds. Science 165:131-137.

Heaney, L. 1991. A synopsis of climatic and vegetational change in Southeast Asia. Climatic Change.

Hewitt, G. 2000. The genetic legacy of the Quaternary ice ages. Nature 405:907-913.

Hewitt, G. 2004. Genetic consequences of climatic oscillations in the Quaternary. Philos T Roy Soc B 359:183-195.

Hoorn, C., F. P. Wesselingh, H. Ter Steege, M. A. Bermudez, A. Mora, J. Sevink, I. Sanmartin, A. Sanchez-Meseguer, C. L. Anderson, J. P. Figueiredo, C. Jaramillo, D. Riff, F. R. Negri, H. Hooghiemstra, J. Lundberg, T. Stadler, T. Sarkinen, and A. Antonelli. 2010. Amazonia Through Time: Andean Uplift, Climate Change, Landscape Evolution, and Biodiversity. Science 330:927-931.

Hugall, A., C. Moritz, A. Moussalli, and J. Stanisic. 2002. Reconciling paleodistribution models and comparative phylogeography in the Wet Tropics rainforest land snail Gnarosophia bellendenkerensis (Brazier 1875). Proc. Natl. Acad. Sci. U.S.A. 99:6112-7.

IUCN. 2011. International Union for Conservation of Nature Red List IUCN.

Jackson, S., and J. Overpeck. 2000. Responses of plant populations and communities to environmental changes of the late Quaternary. Paleobiology 26:194-220.

Jansson, R., and M. Dynesius. 2002. The fate of clades in a world of recurrent climatic change: Milankovitch oscillations and evolution. Annual Reviews in Ecology and Systematics 33:741-777.

Kindt, R., and R. Coe. 2005. Tree diversity analyses. A manual and software for common statstical methods for ecological and biodiversity studies. World Agroforestry Centre (ICRAF).

Lessa, E. P., J. A. Cook, and J. L. Patton. 2003. Genetic footprints of demographic expansion in North America, but not Amazonia, during the Late Quaternary. Proceedings of the National Academy of Science 100:10331-10334.

Loarie, S. R., P. B. Duffy, H. Hamilton, G. P. Asner, C. B. Field, and D. D. Ackerly. 2009. The velocity of climate change. Nature 462:1052-1055.

Mackinnon, K. 2005. Parks, people, and policies: conflicting agendas for forests in Southeast Asia. Pages 558-582 *in* Tropical Rainforests Past, Present, and Future (E. Bermingham, C. Dick, and C. Moritz, eds.). The University of Chicago Press, Chicago & London.

Malohlava, V., and L. Bocak. 2010. Evidence of extreme habitat stability in a Southeast Asian biodiversity hotspot based on the evolutionary analysis of neotenic net-winged beetles. Molecular Ecology 19:4800-4811.

Marshall, C. R. 2007. Explaining latitudinal diversity gradients. Science 317:451-3; author reply 451-3.

Mayle, F. E., D. J. Beerling, W. D. Gosling, and M. B. Bush. 2004. Responses of Amazonian ecosystems to climatic and atmospheric carbon dioxide changes since the last glacial maximum. Philosophical Transactions of the Royal Society B: Biological Sciences 359:499-514.

Mayr, E., and R. J. O'Hara. 1986. The biogegraphic evidence supporting the Pleistocene forest refuge hypothesis. Evolution 40:55-67.

Mittermeier, R. A., P. R. Gil, M. Hoffmann, J. Pilgrim, T. Brooks, C. Mittermeier, J. Lamoreux, and G. da Fonseca. 2004. Hotspots Revisited. CEMEX Conservation International, Mexico City.

Mittermeier, R. A., N. Myers, and C. Mittermeier. 1999. Hotspots. Earth's biologically richest and most endangered terrestrial ecoregions, English edition. CEMEX Conservation International, Mexico City.

Morgan, K., S. M. O'Loughlin, C. Bin, Y.-M. Linton, D. Thongwat, P. Somboon, M. Y. Fong, R. K. Butlin, R. Verity, A. Prakash, P. T. Htun, T. Hlaing, S. Nambanya, D. Socheat, T. H. Dinh, and C. Walton. 2011. Comparative phylogeography reveals a shared impact of plesitocene environmental change in shaping genetic diversity

within nine Anopheles mosquito species acros the Indo-Burma biodiversity hotspot. Molecular Ecology 20:4533-4549.

Moritz, C., C. J. Hoskin, J. B. MacKenzie, B. L. Phillips, M. Tonione, N. Silva, J. VanDerWal, S. E. Williams, and C. H. Graham. 2009. Identification and dynamics of a cryptic suture zone in tropical rainforest. P R Soc B 276:1235-1244.

Morley, R. J. 2000. Origin and evolution of tropical rain forests. John Wiley & Sons, LTD, Chichester.

Morley, R. J. 2007. Cretaceous and Tertiary climate change and the past distribution of megathermal rainforests. Pages 1-31 *in* Tropical rainforest responses to climatic change (M. B. Bush, and J. R. Flenley, eds.). Springer Praxis Publishing, Chichester.

Myers, N., R. A. Mittermeier, C. Mittermeier, G. da Fonseca, and J. Kent. 2000. Biodiversity hotspots for conservation priorities. Nature 403:853-858.

Phillips, B. L., S. J. E. Baird, and C. Moritz. 2004a. When vicars meet: a narrow contact zone between morphologically cryptic phylogeographic lineages of the rainforest skink, Carlia rubrigularis. Evolution 58:1536-48.

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

Phillips, S. J., M. Dudık, and R. E. Schapire. 2004b. A Maximum Entropy Approach to Species Distribution Modeling Proceedings of the 21st International Conference on Machine Learning:8.

Quek, S., S. Davies, P. Ashton, T. Itino, and N. Pierce. 2007. The geography of diversification in mutualistic ants: a gene's-eye view into the Neogene history of Sundaland rain forests. Mol. Ecol. 16:2045-62.

Reddy, S. 2005. Historical biogeography of southern Asia: investigating patterns of biotic assemblages and their relationships using endemic avian taxaColumbia University, New York.

Reddy, S. 2008. Systematics and biogeography of shrike-babblers (Pteruthius): Species limits, molecular phylogenetics, and diversification patterns across southern Asia. Molecular Phylogenetics and Evolution 47:54-72.

Sandel, B., L. Arge, B. Dalsgaard, R. G. Davies, K. J. Gaston, W. J. Sutherland, and J.-C. Svenning. 2011. The influence of Late Quaternary climate-change velocity on species endemism. Science early online October 6, 2011:6.

Scotese, C. R. 2003. PALEOMAP project http://www.scotese.com.

Shafer, A. B. A., C. I. Cullingham, S. D. Côté, and D. W. Coltman. 2010. Of glaciers and refugia: a decade of study sheds new light on the phylogeography of northwestern North America. Molecular Ecology 19:4589-4621.

Slik, J. W. F., S.-I. Aiba, M. Bastian, F. Q. Breasley, C. H. Cannon, K. A. O. Eichhorn, G. Fredriksson, K. Kartawintata, Y. Laumonier, A. Mansor, A. Marjokorpi, E. Meijaard, R. J. Morley, H. Nagamasu, R. Nilus, E. Nurtjahya, J. Payne, A. Permana, A. D. Poulsen, N. Raes, S. Riswan, C. P. van Schaik, D. Sheil, K. Sidiyasa, E. Suzuki, J. C. C. H. van Valkenburg, C. O. Webb, S. Wich, T. Yoneda, R. Zakaria, and N. Zweifel. 2011. Soils on exposed Sunda Shelf shaped biogeographic patterns in the equatorial forests of Southeast Asia. Proceedings of the National Academy of Science 108:12343-12347.

Sodhi, N., L. Koh, B. Brook, and P. Ng. 2004. Southeast Asian biodiversity: an impending disaster. Trends Ecol. Evol. (Amst.) 19:654-60.

Stewart, J. R., A. M. Lister, I. Barnes, and L. Dalen. 2010. Refugia revisited: individualistic responses of species in space and time. Proceedings of the Royal Society B: Biological Sciences 277:661-671.

Stibig, H.-J., A. S. Belward, P. S. Roy, U. Rosalina-Wasrin, S. Agrawal, P. K. Joshi, R. Beuchle, S. Fritz, S. Mubareka, and C. Giri. 2007. A land-cover map for South and Southeast Asia derived from SPOT-VEGETATION data. J Biogeography 34:625-637.

VanDerWal, J., L. P. Shoo, and S. E. Williams. 2009. New approaches to understanding late Quaternary climate fluctuations and refugial dynamics in Australian wet tropical rain forests. Journal of Biogeography 36:291-301.

Weir, J. T., and D. Schluter. 2007. The Latitudinal Gradient in Recent Speciation and Extinction Rates of Birds and Mammals. Science 315:1574-1576.

Williams, J. W., and S. T. Jackson. 2007. Novel climates, no-analog communities, and ecological surprises. Frontiers in Ecology and the Environment 5:475-482.

Williams, J. W., S. T. Jackson, and J. E. Kutzbach. 2007. Projected distributions of novel and disappearing climates by 2100 AD. P Natl Acad Sci Usa 104:5738-42.

Williams, J. W., B. N. Shuman, and T. Webb. 2001. Dissimilarity analyses of Late-Quaternary vegetation and climate in Eastern North America. Ecology 82:3346-3362.

Williams, S. E., and R. G. Pearson. 1997. Historical rainforest contractions, localized extinctions and patterns of vertebrate endemism in the rainforests of Australia's wet tropics. Proc. Biol. Sci. 264:709-16.

Willis, K. J., and J. C. McElwain. 2002. The evolution of plants. Oxford University Press, Oxford.

Wogan, G. O. U. 2011. Chapter 2. Shifting habitats across glacial and interglacial Asia. Pages 28 + figures *in* Integrative Biology University of California, Berkeley, Berkeley.

Ziegler, A. M., G. Eshel, P. M. Rees, T. A. Rothfus, D. B. Rowley, and D. Sunderlin. 2003. Tracing the tropics across land and sea: Permian to present. Lethaia 36:227-254.

Figure 1. The Standard Euclidean distance between four climate variables (mean temperature, annual precipitation, temperature seasonality, precipitation seasonality). Red areas depict regions where climates differ the most, blue areas depict regions where climate has been relatively stable as measured between those two points in time. Panel A depicts SED from the LIG to now. Panel B depicts SED from the LIG to the LGM. Panel C depicts SED from the LGM to now.

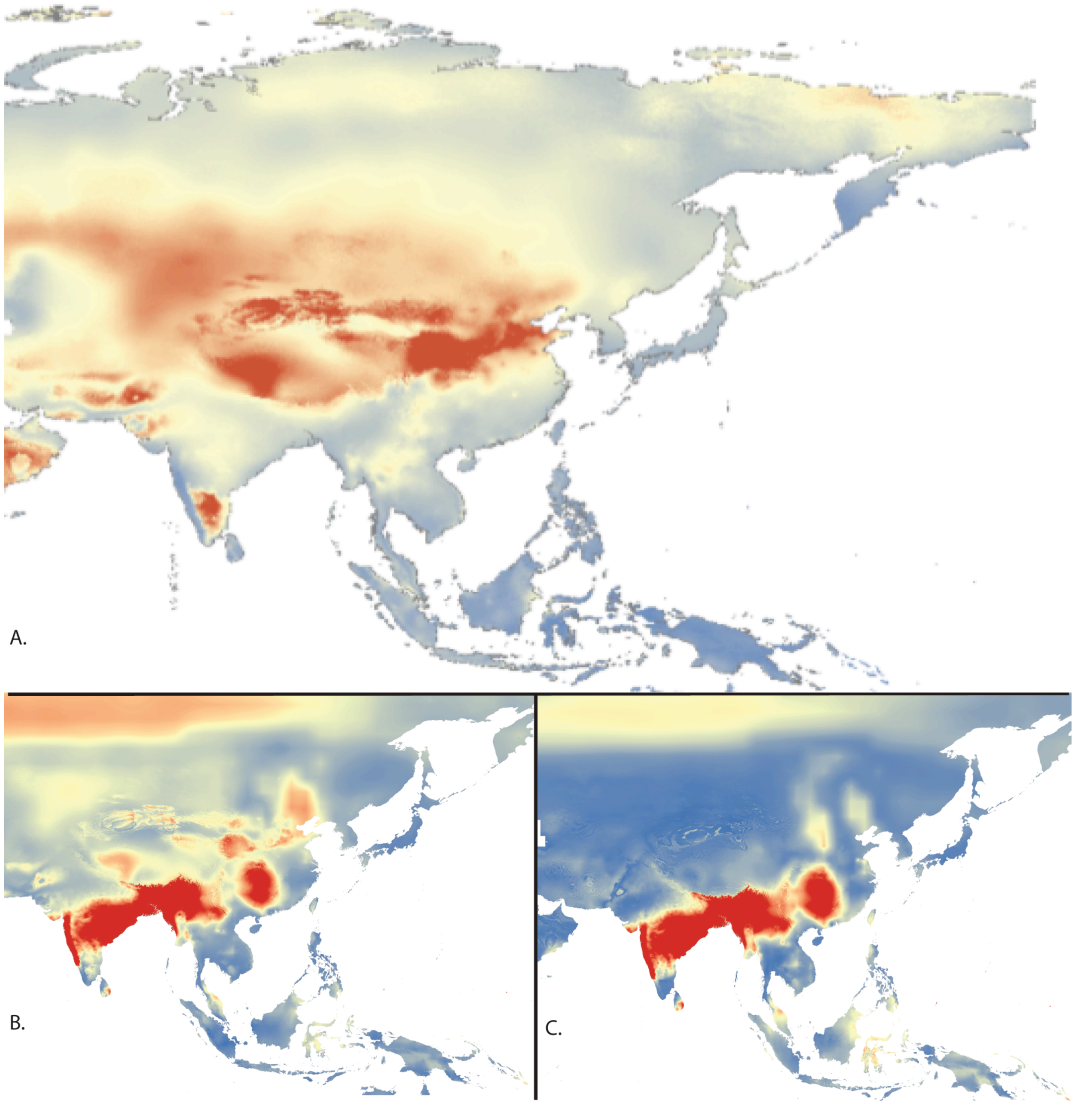

A.
B.
C.

Figure 2. Climate change between the Last Interglacial and now. Panel A depicts univariate histograms for annual mean temperature and annual precipitation for current climatic conditions, Last Glacial Maximum climate conditions, and for Last Interglacial climate conditions. Panel B is a two-dimensional histogram for the same climate variables. It demonstrates how the realized combination of climatic variables differs between the time-periods.

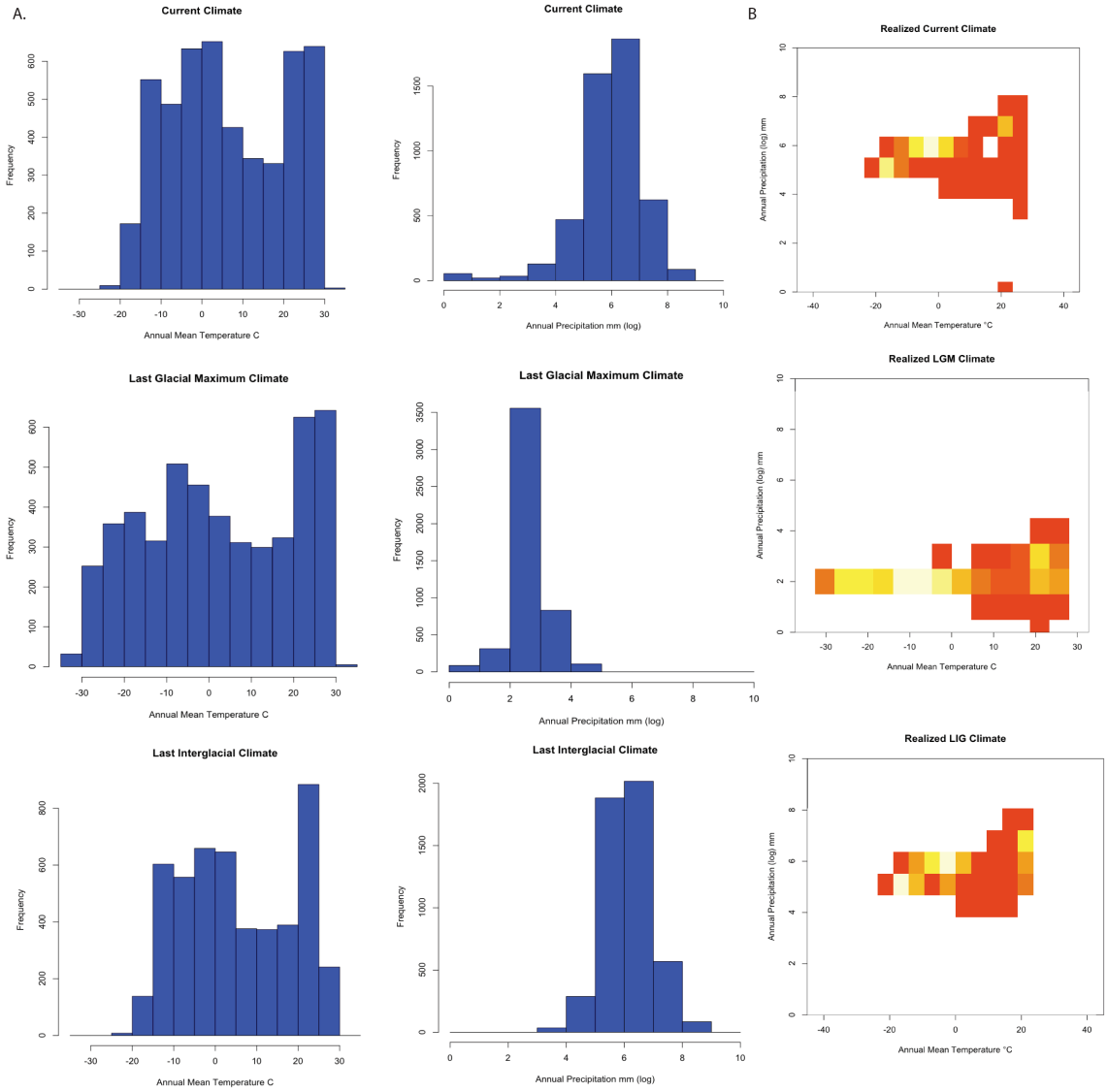

A.
Current Climate
Frequency
Annual Mean Temperature C
Current Climate
Frequency
Annual Precipitation mm (log)
B
Realized Current Climate
Annual Precipitation (log) mm
Annual Mean Temperature °C
Last Glacial Maximum Climate
Frequency
Annual Mean Temperature C
Last Glacial Maximum Climate
Frequency
Annual Precipitation mm (log)
Realized LGM Climate
Annual Precipitation (log) mm
Annual Mean Temperature C
Last Interglacial Climate
Frequency
Annual Mean Temperature C
Last Interglacial Climate
Frequency
Annual Precipitation mm (log)
Realized LIG Climate
Annual Precipitation (log) mm
Annual Mean Temperature °C

Figure 3. Asian Habitat Refugia determined from predictive models and stability analyses from the Last Interglacial to present.


Alpine Grassland
Coniferous Mountain Evergreen Forest
Deciduous Shrubland/Bush
Deciduous Thorny Shrubland
Sparse Desert Grassland
Grasslands on plains and slopes
Mangrove Forest
Sparse Shrubland
Sparse Woods
Temperate Broadleaf Evergreen Lowland Forest
Temperate Broadleaf Evergreen Mountain Forest
Tropical Subtropical Broadleaf Mountain Evergreen Forest
Tropical Broadleaf Deciduous Forest
Tropical Broadleaf Evergreen Lowland Forest

Figure 4. Asian Habitat Refugia, grouped into tropical forest (green), temperate forest (blue), and non-forest categories (cream). Overlap between tropical forest and non-forest refugia (orange), between tropical and temperate forests (dark green), between temperate forests and non-forests (dark blue) are depicted. Overlap among all three habitat classes are in red.

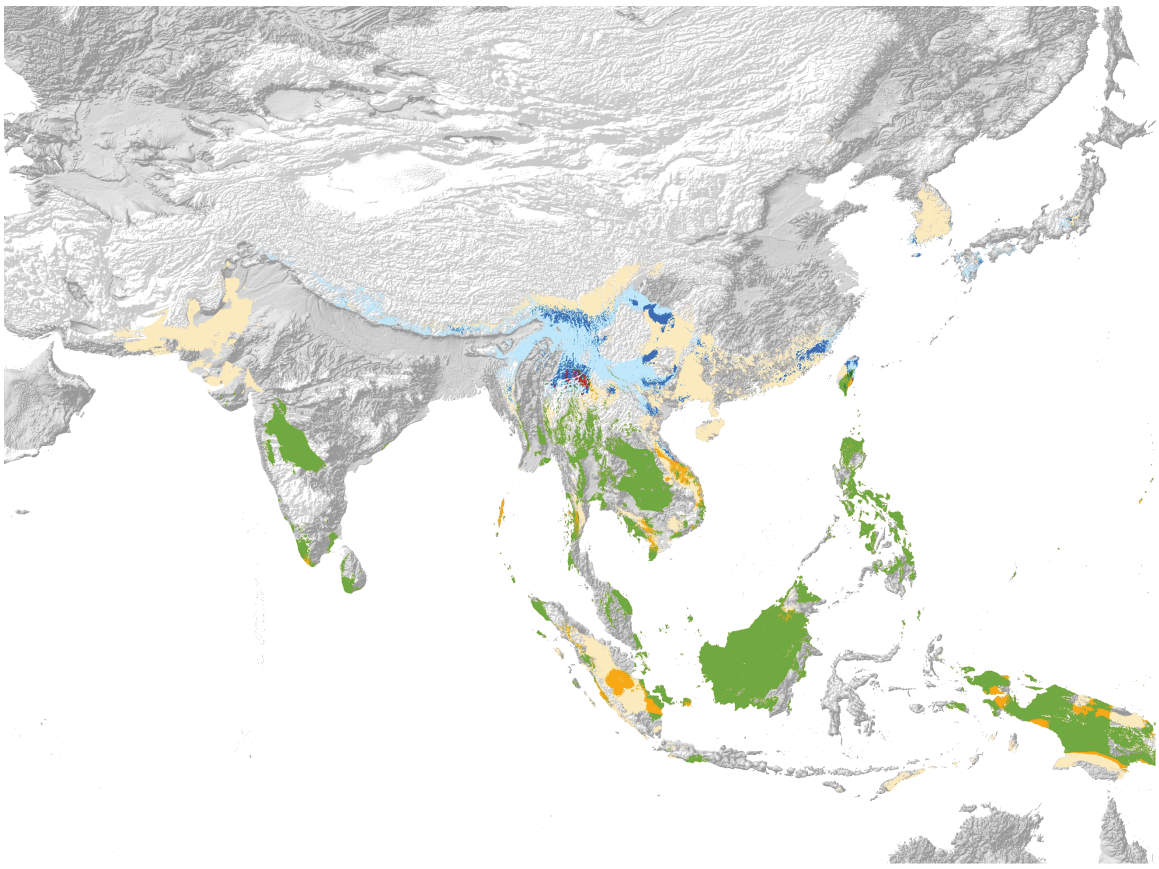


Figure 5. Dynamics of forest change. (A) The relative proportion of each stable habitat from the Last Interglacial to the Last Glacial Maxima, from the Last Glacial Maxima to now, from the Last Interglacial to now, and refugia calculated using the geometric mean (GM). (B) The relative proportion of stable habitats through time depicted by habitat classification (C) The mean area of stable refugia identified for each habitat class.

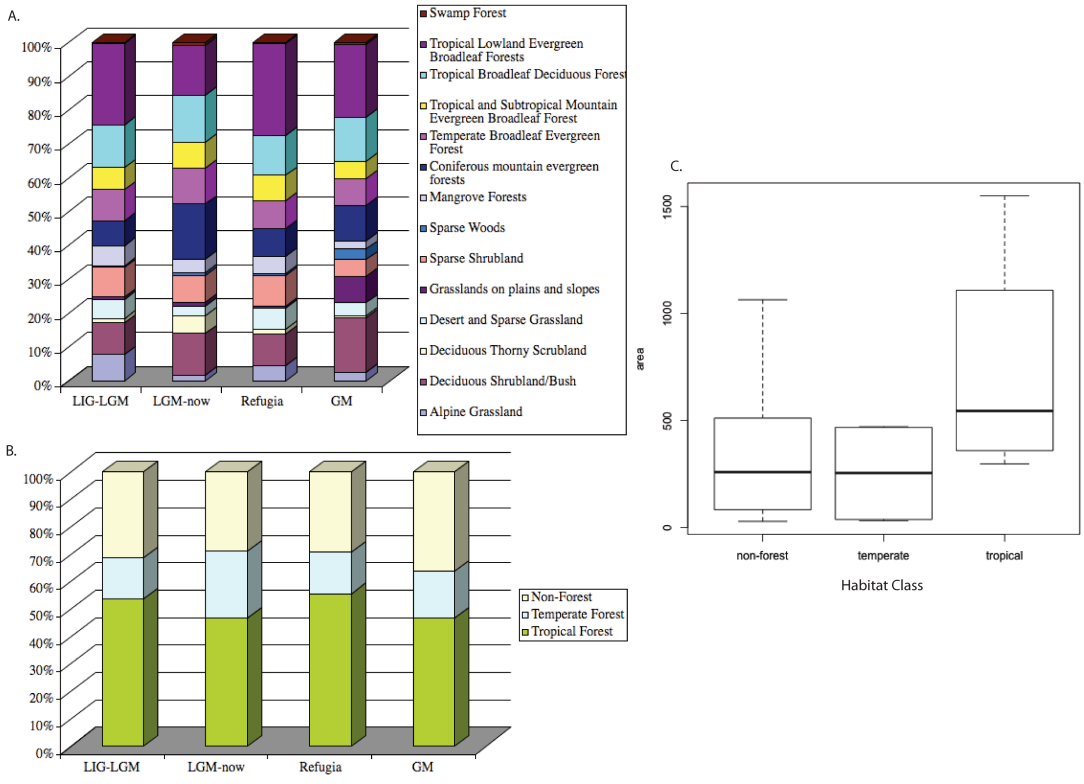

A.
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
LIG-LGM
LGM-now
Refugia
GM
Swamp Forest
Tropical Lowland Evergreen Broadleaf Forests
Tropical Broadleaf Deciduous Forest
Tropical and Subtropical Mountain Evergreen Broadleaf Forest
Temperate Broadleaf Evergreen Forest
Coniferous mountain evergreen forests
Mangrove Forests
Sparse Woods
Sparse Shrubland
Grasslands on plains and slopes
Desert and Sparse Grassland
Deciduous Thorny Scrubland
Deciduous Shrubland/Bush
Alpine Grassland
B.
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
LIG-LGM
LGM-now
Refugia
GM
Non-Forest
Temperate Forest
Tropical Forest
C.
area
1500
1000
500
0
non-forest
temperate
tropical
Habitat Class

Figure 6. Box plot depicting the latitudinal distribution of stable and unstable areas in Asia. As evidenced here, stable areas tend to be found at lower latitudes while unstable habitats are generally found at higher latitudes.

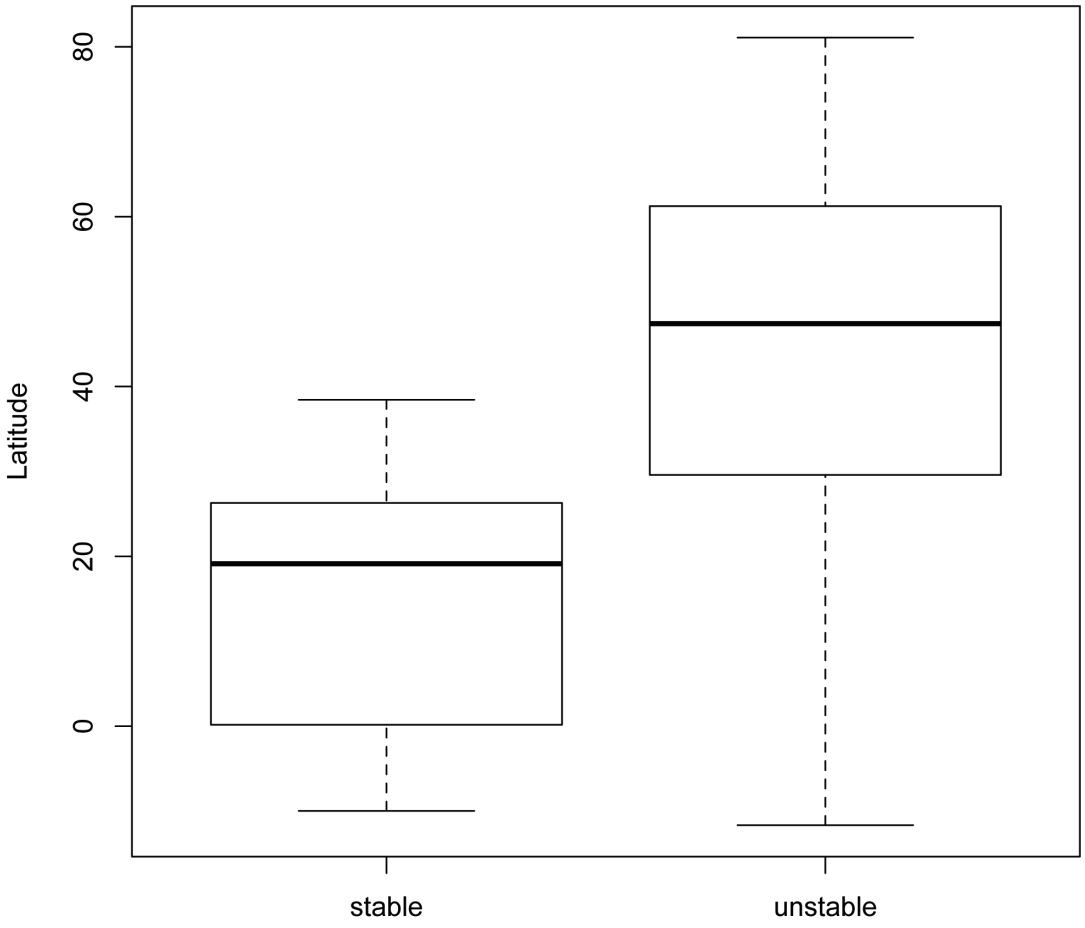
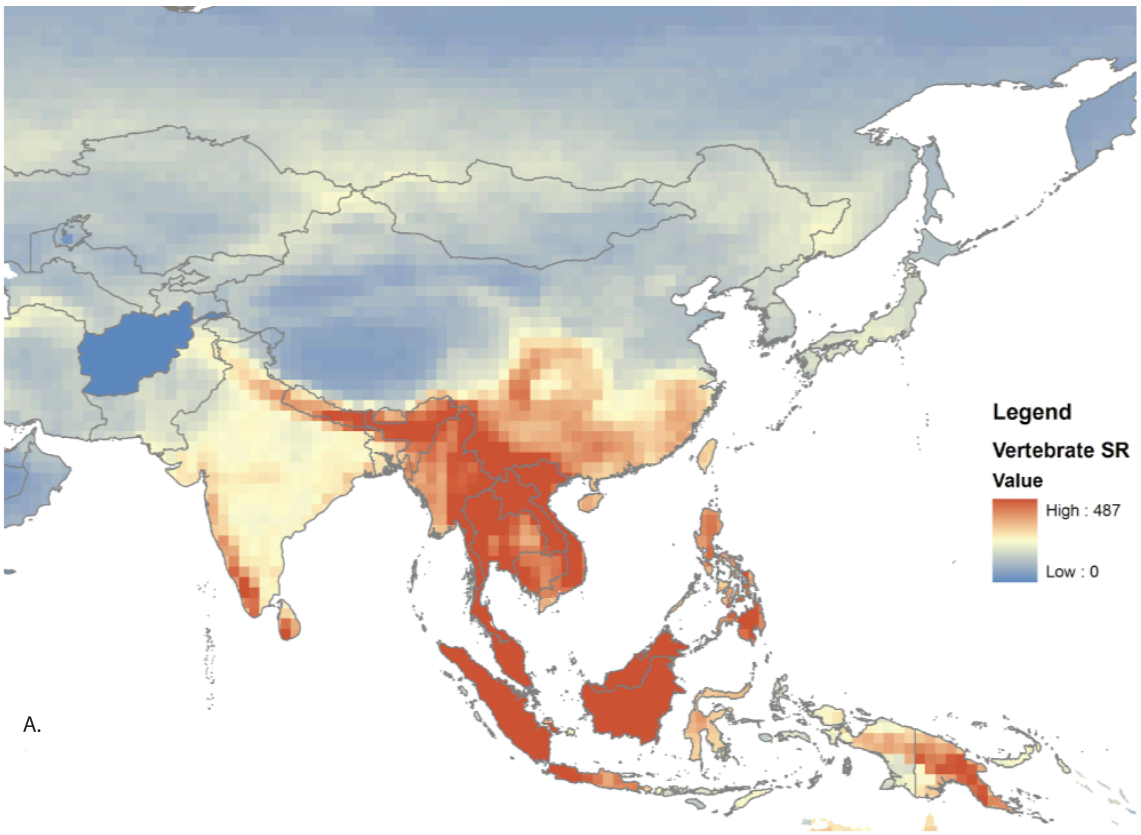
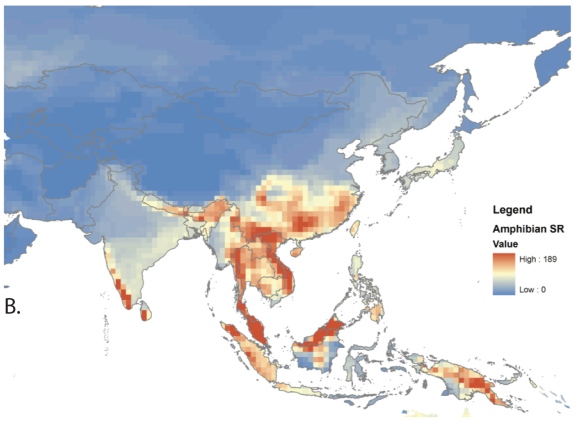
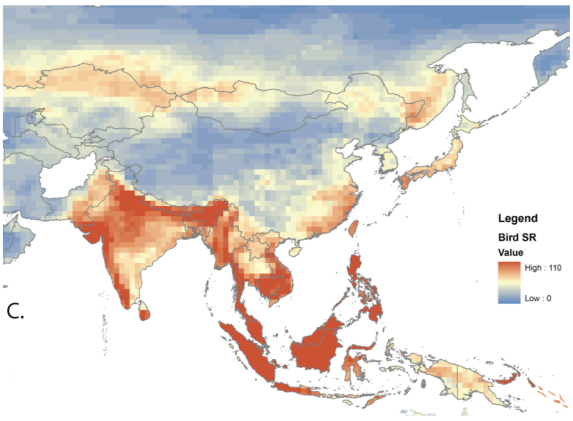
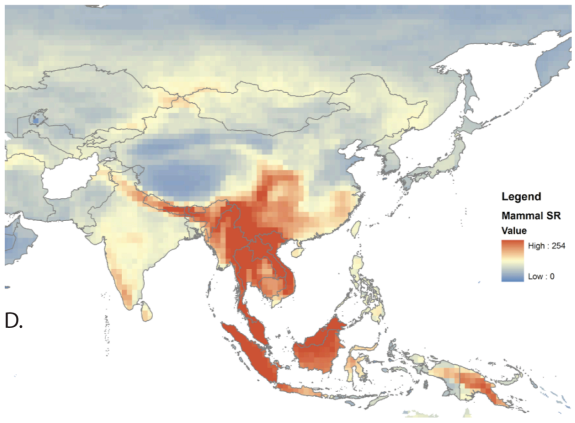

Latitude
0
20
40
60
80
stable
unstable

Figure 7. Species Richness Maps (A) All vertebrates combined (B) amphibians (C) birds (D) mammals (E) reptiles. The color scales from high diversity (red) to low diversity (blue).


Legend
Vertebrate SR
Value
High : 487
Low : 0
A.


Legend
Amphibian SR
Value
High : 189
Low : 0
B.


Legend
Bird SR
Value
High : 110
Low : 0
C.


Legend
Mammal SR
Value
High : 254
Low : 0
D.

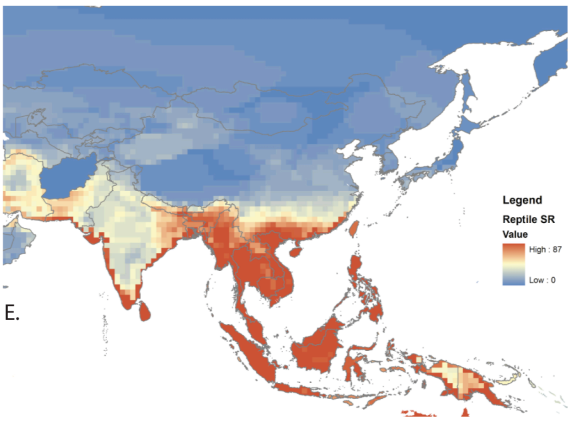
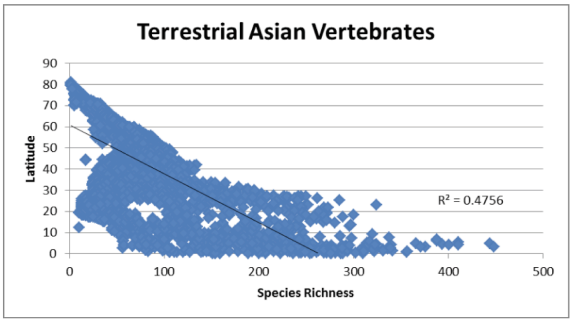
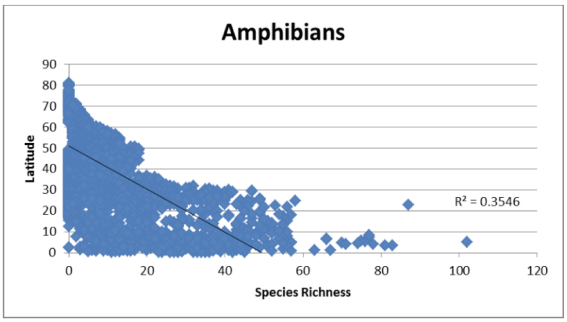
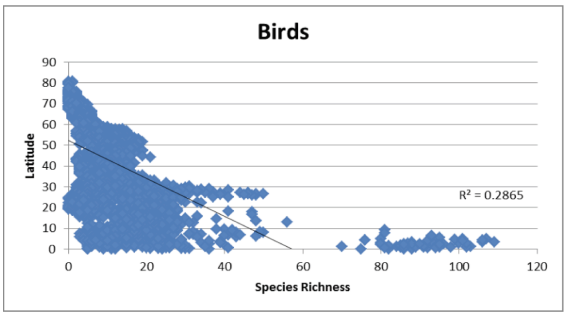
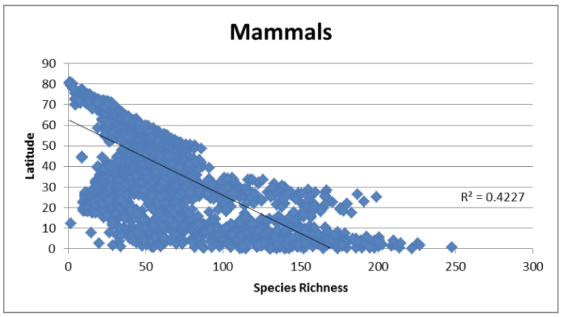

Legend
Reptile SR
Value
High : 87
Low : 0
E.

Figure 8. The relationship between Species Richness and Latitude. Latitude explains a large part of the variance in species richness for all Asian vertebrates, and is particularly important in structuring mammal species richness.


Terrestrial Asian Vertebrates
Latitude
Species Richness
R² = 0.4756


Amphibians
Latitude
Species Richness
R² = 0.3546


Birds
Latitude
Species Richness
R² = 0.2865


Mammals
Latitude
Species Richness
R² = 0.4227

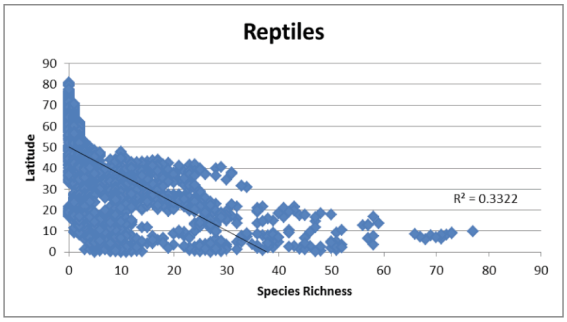

Reptiles
Latitude
Species Richness
R² = 0.3322

Figure 9. The relationship between Species Richness and Habitat Area. The top row depicts the untransformed values, the bottom row the log transformed values. The points used in these analyses are the refugia identified for each habitat type and classified into the three major habitat categories tropical forest (**+**), temperate forest (**o**), and non-forest (△).

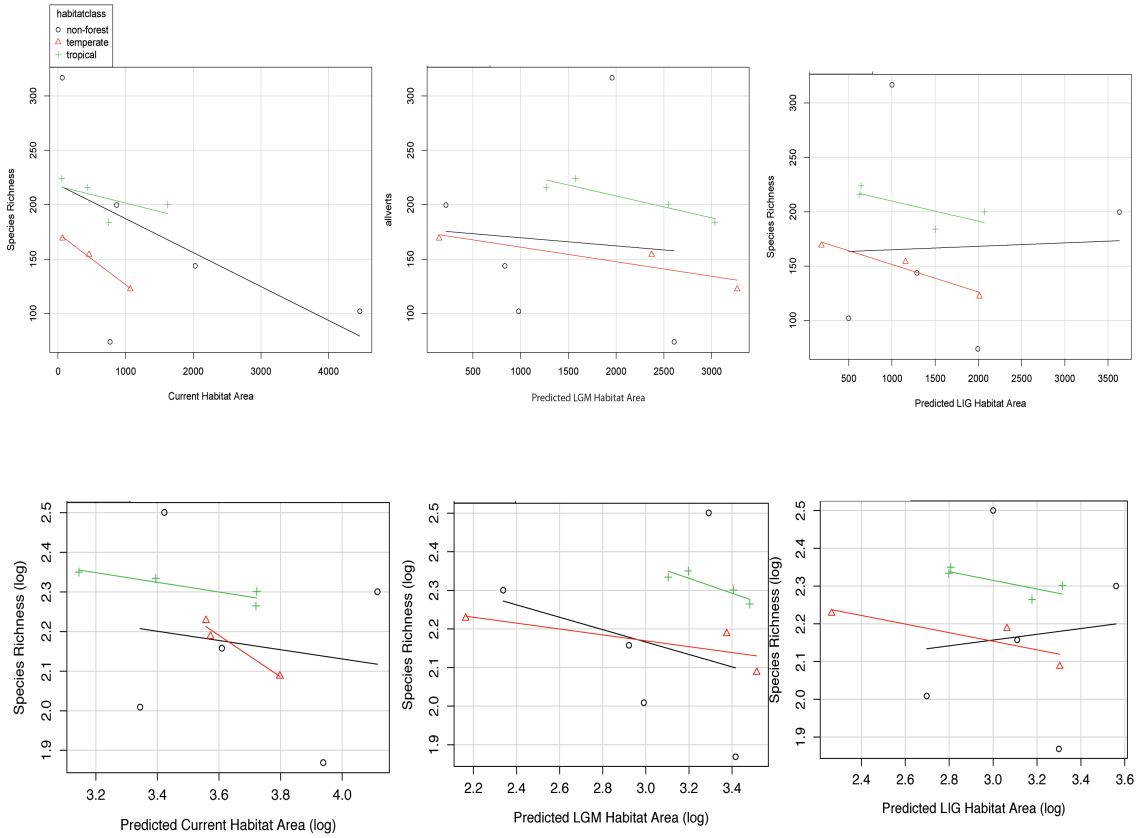

habitatclass
non-forest
temperate
tropical
Species Richness
Current Habitat Area
allverts
Predicted LGM Habitat Area
Species Richness
Predicted LIG Habitat Area
Species Richness (log)
Predicted Current Habitat Area (log)
Species Richness (log)
Predicted LGM Habitat Area (log)
Species Richness (log)
Predicted LIG Habitat Area (log)

Figure 10. Species richness in stable versus unstable areas. Boxplots of species richness values in stable and unstable habitats for all vertebrates combined and for each group independently.

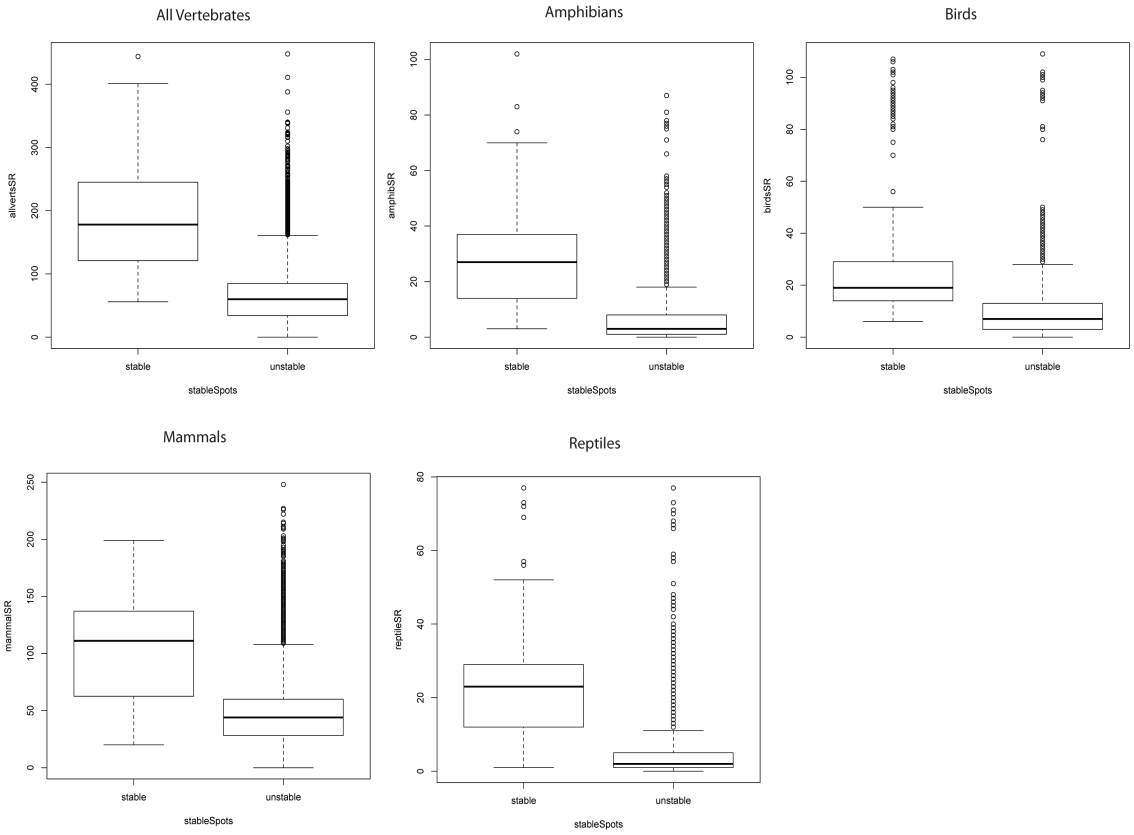

All Vertebrates
allvertsSR
0
100
200
300
400
stable
unstable
stableSpots
Amphibians
amphibSR
0
20
40
60
80
100
stable
unstable
stableSpots
Birds
birdsSR
0
20
40
60
80
100
stable
unstable
stableSpots
Mammals
mammalSR
0
50
100
150
200
250
stable
unstable
stableSpots
Reptiles
reptilesSR
0
20
40
60
80
stable
unstable
stableSpots

Figure 11. Species Richness by habitat classification.

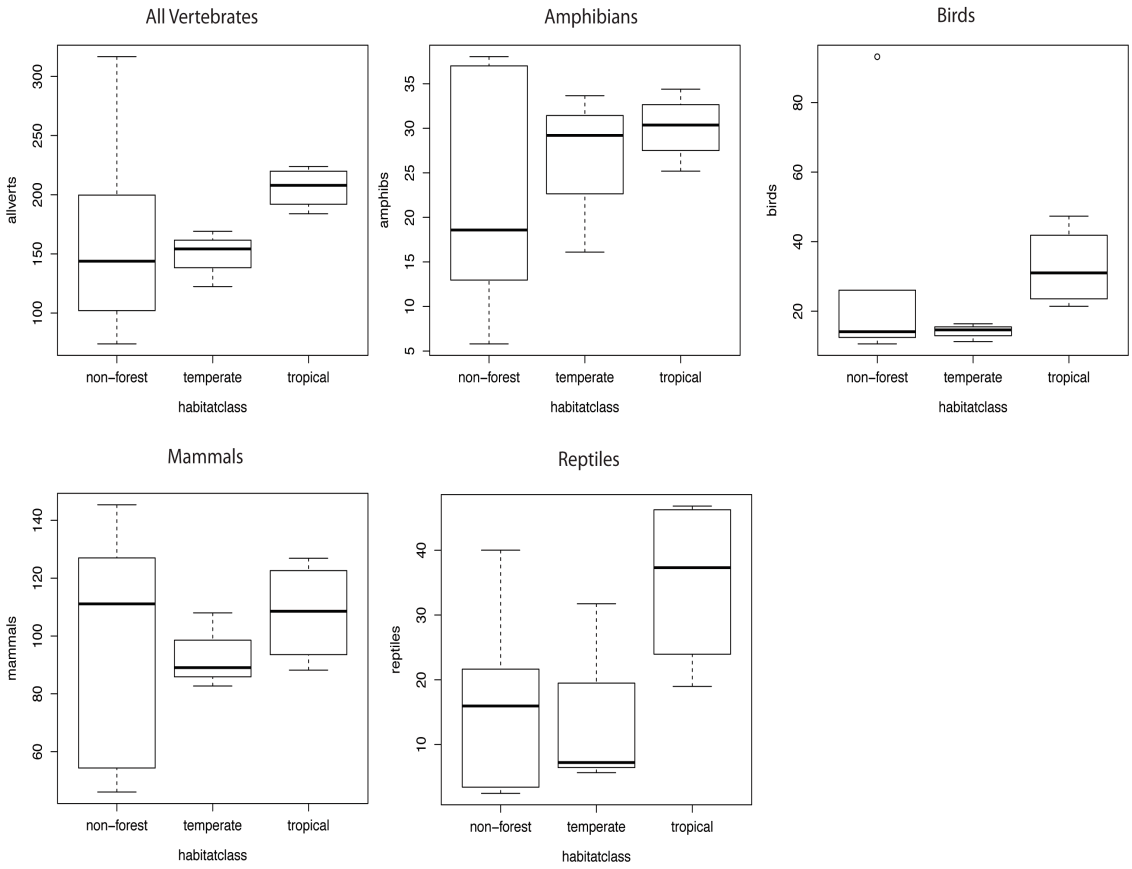

All Vertebrates
allverts
100
150
200
250
300
non–forest
temperate
tropical
habitatclass
Amphibians
amphibs
5
10
15
20
25
30
35
non–forest
temperate
tropical
habitatclass
Birds
birds
20
40
60
80
non–forest
temperate
tropical
habitatclass
Mammals
mammals
60
80
100
120
140
non–forest
temperate
tropical
habitatclass
Reptiles
reptiles
10
20
30
40
non–forest
temperate
tropical
habitatclass

Figure 12. The relationship between Species Richness and habitat stability. (A) species richness values overlaid with identified refugial areas (B) plot of regression analyses of species richness and stability depicted by habitat category (tropical forest, temperate forest, non-forest) (C) Species richness and stability for each of the vertebrate groups.

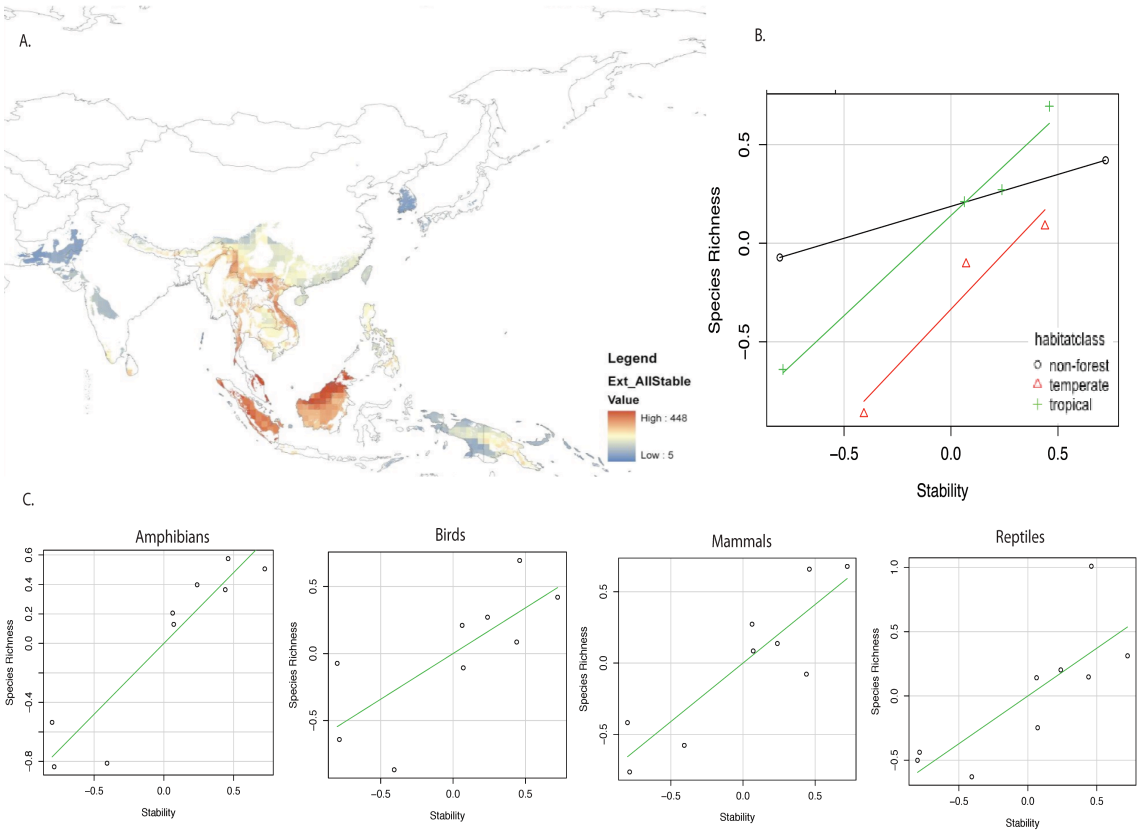

A.
Legend
Ext_AllStable
Value
High : 448
Low : 5
B.
Species Richness
Stability
habitatclass
non-forest
temperate
tropical
C.
Amphibians
Birds
Mammals
Reptiles
Species Richness
Stability

Table 1. Results of Mantel and Anosim Matrix analyses. Significant results are demarcated with **.

Analyses	Matrix 1	Matrix 2	R_(anosim) r_(mantel)	P-value	95% CI
Anosim	CC-Amphibians_Jaccard	HIS	0.0232	0.093812	0.0291
Anosim	CC-Amphibians_Euclidean	HIS	0.005169	0.27545	0.0146
Anosim	CC-Birds_Jaccard	HIS	0.02483	0.077844	0.0292
Anosim	CC-Birds_Euclidean	HIS	0.005087	0.30339	0.0164
Anosim	CC-Mammals_Jaccard	HIS	0.02509	0.11881	0.0316
Anosim	CC-Mammals_Euclidean	HIS	0.005087	0.31683	0.0163
Anosim	CC-Reptiles_Jaccard	HIS	0.02179	0.093812	0.0289
Anosim	CC-Reptiles_Euclidean	HIS	0.005102	0.26747	0.0157
Mantel	CC-Amphibians_Jaccard	HIS	0.01201	0.095808	0.0148
Mantel	CC-Amphibians_Euclidean	HIS	0.003241	0.31936	0.00947
Mantel	CC-Birds_Jaccard	HIS	0.01284	0.13972	0.0168
Mantel	CC-Birds_Euclidean	HIS	0.003228	0.26747	0.00973
Mantel	CC-Mammals_Jaccard	HIS	0.01298	0.06986	0.0140
Mantel	CC-Mammals_Euclidean	HIS	0.003228	0.33733	0.00989
Mantel	CC-Reptiles_Jaccard	HIS	0.0128	0.10978	0.0148
Mantel	CC-Reptiles_Euclidean	HIS	0.003199	0.28343	0.0143
Mantel	CC-Amphibians_Jaccard	CLS	-0.01274	0.92216	0.0158
Mantel	CC-Amphibians_Euclidean	CLS	-0.007046	0.85429	0.0114
Mantel	CC-Birds_Jaccard	CLS	-0.01277	0.91617	0.0182
Mantel	CC-Birds_Euclidean	CLS	-0.007049	0.87226	0.01092
Mantel	CC-Mammals_Jaccard	CLS	-0.01272	0.90419	0.0157
Mantel	CC-Mammals_Euclidean	CLS	-0.007044	0.87026	0.01117
Mantel	CC-Reptiles_Jaccard	CLS	-0.01272	0.91816	0.0178
Mantel	CC-Reptiles_Euclidean	CLS	-0.007044	0.87824	0.01148
Mantel	CLS	LAT	0.169	0.001996**	0.00591
Mantel	LAT	HIS	0.00324	0.30539	0.01043
Mantel	CLS	HIS	0.09991	0.001996**	0.01046

Table 2. Comparison of the Species Richness values for amphibians, birds, mammals, reptiles, and combined vertebrates between identified regions of habitat stability and a pool of random SR values generated from across Asia. All analyses had 5 degrees of freedom. Significance values at 95% confidence are marked with a single *. Significance values at greater than 99% are designated with **.

Habitat	t value	p-value
Alpine grasslands	1.8889	0.1175
Coniferous evergreen forest	8.5356	0.0003633**
Deciduous schrubland	-0.3786	0.7205
Deciduous thorny schrubland	1.57	0.1772
Grasslands	5.5267	0.002658**
Mangrove forest	5.6658	0.002382**
Sparse shrubland	8.7025	0.0003315**
Sparse woods	3.9251	0.01112*
Temperate broadleaf evergreen forest (lowland)	5.4689	0.002783**
Temperate broadleaf evergreen forest	3.318	0.02902*
Tropical and Subtropical Mountain Evergreen Broadleaf Forest	8.2436	0.0004281**
Tropical Broadleaf Deciduous Forest	10.5858	0.00013**
Tropical Lowland Evergreen Broadleaf Forests	4.98008	0.004471**

Table 3. Results from linear regression of species richness and stability with the effects of total predicted habitat area removed. Significance levels demarcated with either * (95% confidence) or ** (99% confidence).

	R^2 (adjusted)	p value
amphibians	0.7474	0.001624**
birds	0.4256	0.03381*
mammals	0.8456	0.0002787**
reptiles	0.5839	0.01004*
combined vertebrates	0.51	0.01849*

CHAPTER FOUR

RECONSTRUCTING ANCIENT PATHWAYS: THE HABITATS OF SUNDALAND DURING GLACIALS AND INTERGLACIALS

INTRODUCTION

The Sundaland biodiversity hotspot harbors an amazing array of biodiversity with an intriguing combination of endemic and widespread shared fauna and flora. Sundaland is a continental shelf that presently lies under the Gulf of Thailand and South China Sea. During glacial periods when ocean levels drop, much of the shelf is exposed creating additional land positive area. The continuous land across the region connects mainland SE Asia with many of the large islands of SE Asia such as Java, Sumatra and Borneo, as well the many small islands dotting the Gulf of Thailand. Borneo and some of the Philippine Islands also become connected at this time. When land positive, this region provides a potential migration corridor, which could connect faunas isolated for long stretches of geologic time. Understanding the timing and degree of exposure, the topography of the region, and the associated palaeohabitats across the shelf is essential to understanding the abiotic forces structuring the biodiversity in mainland and insular SE Asia. Inger and Voris (2001) and Woodruff and Turner (2009) using different eustatic curves calculated that over the deep geologic history, concluded that the area is land positive more often than not. Marine transgressions occur along with interglacial periods when sea levels are high; the extent of those marine transgressions however remains debated. Woodruff and Turner (2009) suggested that at least 58 marine transgressions occurred over the past 5 million years which cut the Malay Peninsula in to a series of isolated islands. Over the last 1.8-2 million years, roughly 18-20 Pleistocene glacial-interglacial cycles have occurred (Gibbard and Cohen, 2008). For the past million years, the tempo of the cycles has been on 100 thousand year glacials, and 10 thousand year interglacials, therefore in recent geologic history the Sundaic region is predicted to have been contiguous much more than not. This time scale is relevant to most of the extant vertebrate fauna as species durations are estimated on the order of 2-3 million years for mammals (Alroy, 2000), and 2 million years for birds (Zink and Slowinski, 1995). Species durations are undetermined for reptiles and for amphibians but are potentially congruent with estimates for other terrestrial vertebrates (Avise et al., 1998).

Several analyses have focused on reconstructing the paleo-ecology of SE Asia and Sundaland. Heaney (1991) suggested a general pattern in which as sea levels decreased, there was a corresponding decrease in rainfall and a subsequent decrease in rainforest and forest areas coupled with an increase in savanna habitats. Ray and Adams (2001) suggest that during the LGM most of SE Asia was drier than present with open grassland and dry forest habitats dominating. Multiple studies suggest the presence of a Sundaic Savannah Corridor traversing Java, the Malay Peninsula, and Indochina lasting well into the Pleistocene (Bird et al., 2005; Heaney, 1991; Wurster et al., 2010). The savannah corridor has been inferred from palynology (Stuijts et al., 1988; Wogan, 2011b), vertebrate fossils (Cranbrook, 2010; Tougard, 2001; van den Bergh et al., 2001), vertebrate distributions (Meijaard, 2003), and most recently isotopic evidence (Wurster et al., 2010). Alternatively, Cannon et al., (2009) point out that the palynological evidence

for the corridor is from a poorly dated site, and suggest that Sundaland lowland forests are currently in a refugial state, having a wider distribution when sea levels are low during glacial maxima. Evidence supporting this hypothesis comes from palynology (Sun et al., 2000) and phylogeography (Cannon and Manos, 2003). More recently, a third study suggests that the much of the Sunda Shelf was covered by kerapah swamps and heath forests due to the presence of course textured sands and poor drainage soil types on the exposed shelf (Slik et al., 2011). They suggest that sandy soils would support heath forest, and poor drainage soils kerapah peat forest.

Here these alternative hypotheses are tested using a series of environmental niche models of habitats across the region. I ask (1) Is potential forest area predicted to increase or decrease during glacial periods as compared to interglacial periods? (2) Are dry savannah/open corridors predicted across Sundaland during glacial maxima? and (3) What is the potential distribution of habitats across Sundaland during interglacial versus glacial periods?

METHODS

Bathymetric data: Using the Etopo2 global DEM (NOAA, 2006) in combination with estimates of ocean level from global sea-level estimates from the Phanerozoic (Horton et al., 2005; Miller et al., 2005; Müller et al., 2008), the land positive area of the Sunda shelf was estimated for the past 3 million years (Table 1).

Habitat Modeling: Using a GIS approach, I modeled the distribution of habitats across Sundaland at the Last Glacial Maximum, when the maximum amount of land was above water, and at the last interglacial (LIG) when much less land was exposed. To do this, prior modeling efforts were utilized to determine which habitats were predicted to occur anywhere in Sundaland (Wogan, chapter 2). Seven habitats were predicted to occur within the Sunda region (Wogan, chapter 2), so the current analyses were limited to those habitats. Analyses were restricted to an area encompassing Sundaland. The input data consisted of habitat points categorized from satellite imagery into habitat categories derived from Stibig et al. (2007) (detailed in Wogan, chapter 2). MaxEnt v. 3.3.2 was utilized to generate a suitability surface for each habitat category (Phillips et al., 2006; Phillips et al., 2004). For current climate, I used the WorldClim 1.4 GCM (Hijmans et al., 2005) downscaled to 30 second arc resolution. Previous modeling has shown that for predicting the distribution of habitat types, measures that incorporate seasonality are particularly important in SE Asia (Wogan, chapter 2), thus 19 bioclimatic variables were utilized for modeling (Table 2). Current climate predictions were generated and then projected back in time onto paleoclimatic GCMs: Miroc 3.2 at the LGM, and the Otto-Bliesner et al. (2006) GCM for the LIG. Paleoclimatic GCMs were downscaled to 30 arc second resolution following standard protocols (Hijmans et al., 2005; Richards et al., 2007). The land positive areas at both time periods were utilized to determine the appropriate landmask. AUC statistics were utilized to determine the performance of the models under current climate conditions. Thresholds were applied that equalize sensitivity and selectivity. Since the area of exposed land changes through time, the total

area of each habitat is reported as the total area, but is also standardized against the total available land area.

Next, I examined the congruence of the GIS models with other independent datasets such as isotopes (Wurster et al., 2010) and fossil pollen (Anshari et al., 2001; Newsome and Flenley, 1988; Stuijts et al., 1988; Sun et al., 2000; Van Der Kaars and Dam, 1995; Wang et al., 2009). Lastly, I make some predictions regarding the structure of genetic data that would validate the habitat models, and compare these predictions with a few phylogeographic studies in the region.

RESULTS

During the LGM sea levels are believed to have been about 120 m lower than present (Miller et al., 2005; Sathiamurthy and Voris, 2006; Voris, 2000). At the LIG (120k y.b.p.) sea levels are predicted to have been maximally 24 m higher than present (Miller et al., 2005). The high stand lasted less than 5 thousand years and was one of the highest recorded sea level estimates from the past 3 million years; one higher highstand 24.8 m greater than present was reported from 2.38 million years ago (Miller et al., 2005). Woodruff and Turner (2009) however mention that the LIG highstand may have been erroneous. Sea levels would have needed to rise to +50 m to have caused the Malay Peninsula to break into isolated islands (Figure 1 panels A, B), a highstand of +48 m occurred 5.33 mya lasting for about 5 ky, and highstands up to +40 m occurred between 8-9 mya, suggesting that complete marine transgressions are not the most likely cause of recent phylogeographic breaks across the region, but could be involved in deeper phylogenetic divergences. Based on these analyses, land area during the LIG was reduced to 456,445 km^2, while at the LGM almost three times as much land was above sea level, with a total area of 1,306,086 km^2 (Figure 1). However, for the majority of the past three million years, sea levels were between 0 to –50 m relative to current sea level (Figure 2). Sea-levels during the last 3 million years were between –20 to –30 m below present for the longest duration of time (Figure 2, Table 1), thus the most common land form corresponds with panels I and J of Figure 1. Sea levels would have been high enough to isolate Borneo from the rest of the Sundaic islands, while Java, Sumatra and the Malay Peninsula would have been connected.

It should be noted that these analyses do not take into account orogenic effects. Ongoing orogeny may be particularly important to Borneo, where dynamic geomorphologies have been demonstrated (Liechti et al., 1960). However, other regions of Sundaland such as Peninsular Malaysia are thought to have remained stable since the Paleocene (Gobbett and Hutchison, 1973). Since the focus of marine transgressions has centered on Peninsula Malaysia, and the geomorphology there is thought to be stable, the results obtained here should not be affected. Closely tied to orogeny is orography which could also be important across the region (Quek et al., 2007). The orographic affects in Peninsular Malaysia should have remained relatively stable, while these may be more dynamic in Borneo. Subsidence could also affect interpretation of sea-levels in relation to land. Subsidence on Sundaland could be a major geologic force (Hall and Morley, 2004), but the degree to which it impacts interpretation at the time-scale under consideration

here is uncertain. Heine et al. (2008) calculated total tectonic subsidence and anomalous tectonic subsidence of the Sundaland basin and found that while total tectonic subsidence was relatively low (less than or close to 1 km), anomalous tectonic subsidence was high indicating a high sediment accumulation.

In total, five forest habitat types were modeled across Sundaland: rainforest, montane rainforest, deciduous tropical and subtropical forest, mangrove forest, and swamp forest. Two non-forest habitats were modeled: deciduous shrubland and mixed sparse shrub and grass (Figure 3). Models performed well with AUC scores above 0.9 in all instances, and clamping was not an issue in any of the analyses.

The distribution of Sundaland habitats during glacials (Figure 4)

During the LGM, an extensive patch of lowland rainforest connecting Borneo, Sumatra, and Peninsular Malaysia is predicted. Three separate rainforest patches are predicted on Borneo Island, one in NE Sabah, the second in East Kalimantan, and the third encompassing much of Borneo and Sundaland, connecting with Sumatra (South Sumatra) and Peninsular Malaysia (Johor Province). Smaller rainforest patches are predicted on Sumatra (Bengkulu and Jambi regions) and some of the Mentawai Islands (Enggano, Siberut, Sipora), Java (at the West Java, Central Java border), as well as a few of the small islands north of Java (Gresik, Jepara), a single patch in Malaysia (Terengannu), and another extensive patch through what is now the Myeik Archipelago and Peninsular Myanmar region. Extensive mountain evergreen rainforest is predicted across the mountains of Borneo under the Miroc GCM. Most of the Javan and Sumatran mountains and some of the mountains on the Lesser Sunda Islands (Bali, Lombok, Sumbawa, Flores) are also predicted to have been mountain evergreen forest. Two separate patches are predicted in Peninsular Malaysa in Terrenganu and Perak. Deciduous broadleaf forest existed on northern Sundaland on continental Asia, with a large forest patch predicted near the coast of eastern Peninsular Malaysa. Much of the emergent Sunda shelf is predicted to be covered in dry shrub growth. Swamp forests at the LGM are predicted to have had a highly restricted distribution with a single patch on the southern coasts of Java. Mangrove forests at the LGM are predicted along the northwest coast of Sundaland.

Deciduous shrubland was restricted in Sundaland, existing only in the vicinity of what is now the Isthmus of Kra. Conversely, [non-deciduous] shrubland habitat was extensive across Sundaland. A continuous corridor stretched from Java to continental Asia (Figure 4).

The distribution of Sundaland habitats during inter-glacials (Figure 4)

During the LIG, rainforest contracted. Although lowland evergreen rainforest covered much of Borneo, in Java, it was restricted to a single isolated patch, and in Sumatra, it was only found in the Aceh, Jambi, and Bengkulu Provinces as well as some of the Mentawai Islands. In Peninsular Malaysia, the Johor and Terengganu regions contained rainforest, as well as peninsular Myanmar (Tanintharyi region). Mountain evergreen rainforest is predicted on the mountains of Borneo and Sumatra, and on the Peninsular Malaysian Mountains. Interestingly, deciduous forests are predicted farther

south than their current distribution with forest patches predicted for the Malaysian province of Terengannu, and the Indonesian regions of Aceh, Bangka-Belitung, and Kalimantan, as well as peninsular Myanmar and Thailand. Swamp forests are predicted in Malaysian Borneo, Kalimantan, western Java, Sumatra, and Peninsular Malaysia. Extensive mangrove forests are predicted along the northern and western coasts of Borneo, and to a lesser degree in Sumatra and Peninsular Malaysia, with a single mangrove forest patch predicted on Java.

Deciduous shrubland had a restricted distribution mainly confined to mainland Asia, north of the Isthmus of Kra. Shrubland habitat however, was extensive in the lowlands of Sumatra, Java, Borneo, and some of the Mentawai Islands but altogether absent in Peninsular Malaysia.

The absolute proportion of rainforest habitat increased at the LGM, although the relative proportion (scaled by available area) decreased (Figure 4). A very large expansion of shrubland habitat is observed between the LIG and LGM.

The glacial model agrees with the isotope evidence from four cave sites in the Sunda region (Wurster et al., 2010). The Batu caves (Peninsular Malaysia) and Gangub and Makangit caves (Palawan, Philippines) were thought to be surrounded by savannah/grassland areas, while the Niah cave site (Borneo) was forested (Wurster et al., 2010) (Figure 5), which is consistent with the glacial models. Palynological samples from throughout the Sunda region also concord well with the glacial model (Figure 5). Of the ten cores available, seven are in agreement with the predicted distribution of habitats during the LGM, and these are described below. Both of the Javan cores evidenced mountain evergreen forest at the LGM (Stujits et al., 1988; Van Der Kaars and Dam, 1995), this accords with our predictions. The Lake Permarak core on Borneo provided evidence of lowland rainforest (Anshari et al., 2001). The Subung core in southern Peninsular Malaysia reflected open habitat consisting of primarily of conifers, grasses and ferns (Morley and Flenley, 1987). An oceanic core near the mouth of the ancient North Sunda/Molengraaff River, which drained from Borneo, suggests that Borneo was covered in mountain and lowland rainforest (Wang et al., 2007). An oceanic core near the northeastern Borneo coast also provides evidence of lowland and mountain rainforest on Borneo (Sun et al., 2000). Although there are no conflicts between the glacial model and the pollen data, there are four ambiguous locations where the pollen core does not fall into an identified habitat in the model (Figure 5). Two of these are located in the South China Sea off of the Borneo coast and provide evidence for lowland rainforest and mangrove forest near Natuna Island during the LGM (Wang et al., 2009), whereas the glacial models did not predict any of the six habitats to occur there. Two pollen cores on Sumatra come from Danau di Atas a mid-elevation site where clear depression of elevational limits of mountain rainforest has been demonstrated, suggesting that this region was undergoing dramatic climate related changes at this time (Newsome and Flenley, 1988; Stujits et al., 1988). Our model was not able to predict the occurrence of any habitat here. Additional palynological cores from the Indian Ocean, the South China Sea, and from current Sundaic landmasses would help refine and validate understanding of vegetation history in the region. However, given the good match between those cores

that do exist, the isotopic data, and the predictive models generated here, it seems that the approach utilized in this study provides a robust means of generating estimates of potential palaeo-habitats.

DISCUSSION

The existence and nature of the Sundaland savannah corridor has been a matter of debate since proposed. Conflicting evidence, uneven sampling, and ocean inundation across the region have all contributed to a poor understanding of the environmental history of Sundaland. The models generated for this study, in combination with sea-level estimates for the past three million years, reveal a complex interaction between habitat distributions, exposed land area, and connectivity among Sundaic land masses. The models support the existence of a large dry corridor connecting southern Borneo, Java, Sumatra, and Peninsular Malaysia with southern Vietnam, and Cambodia during glacial maxima (Figure 3). However, the extent and position of the corridor and hence the connectivity among landmasses would be altered with rising sea levels. Based on the eustatic curve for the last 3 million years, there have been very few low-stands that would have allowed the most extensive formation of the corridor, these occurred at the LGM (~21ky bp), 145ky, 360ky, 460ky, 625ky, 875ky bp (based on Miller et al. (2005)).

Other studies have predicted extensive lowland evergreen rainforest during glacial periods (Cannon et al., 2009). Our models demonstrate that while lowland rainforest is greater in absolute area during glacial periods, there was not extensive lowland forest across Sundaland. Instead, a few large lowland forest habitat patches are predicted that correspond with refugia delineated through analyses of forest termites (Gathorne-Hardy et al., 2002) (Figure 5). Major lowland forest patches centered on western Borneo, Java, Sumatra, the Mentawai Islands, southern Cambodia, and northwest Sundaland (off the coast of peninsular Myanmar) are predicted during the LGM. The Borneo forest patch is greatest in extent. The size of the lowland forest patches cover a greater absolute area than either contemporary rainforests or rainforest from the previous interglacial period, however, the LGM lowland rainforests are not widely distributed but instead are concentrated into large forest blocks. Relative to the entire area of exposed landmass at the LGM, there was less forest during the LGM.

Although the predictive models support the presence of a large forest tract across the Sunda shelf connecting Borneo with Peninsular Malaysia, the realized forest extent could have been limited by barriers other than climate, in particular, sandy soils and poor drainage soils could have presented a dramatic limitation on lowland rainforest expansion. Evidence of soils as a limiting factor come from recent analyses of forest tree community structure in Peninsular Malaysia, Borneo, and Sumatra (Slik et al., 2011).

Do these models support the hypothesis that modern lowland rainforests are in a refugial state (Cannon et al., 2009)? Refugia generally are defined through their reduced spatial distributions during some specific category of climatic conditions, e.g. glacial or interglacial refugia (Ashcroft, 2010; Bennett and Provan, 2009). The reduction of absolute area of contemporary and interglacial lowland rainforests relative to glacial

lowland rainforests suggests that yes, modern lowland rainforests are refugial, while the reduction of the relative proportion of glacial lowland rainforest to available land suggests the opposite. The very different distributions of contemporary lowland rainforests versus glacial lowland rainforests, suggest that glacial lowland rainforests are refugial, while the more widely distributed contemporary lowland forests are non-refugial.

It is unique that Asian rainforests expand in absolute area at glacial times, since Afrotropical, some Neotropical, and Australian tropical rainforests undergo range restrictions and contractions during glacial periods (Bonnefille, 2007; Carnaval et al., 2008; Carnaval and Moritz, 2008; VanDerWal et al., 2009). Haffer (1969) suggested that orographic effects were particularly important in the Amazon, although it should be noted that Amazonian refugial dynamics remain an active topic of debate (Bush and de Oliveira, 2006; Bush et al., 2007; Colinvaux and De Oliveira, 2001; Colinvaux et al., 2000; Colinvaux et al., 1996; Haffer, 1969; Haffer and Prance, 2001).

Based on the good accord between these models and the independent palaeo-ecological data discussed above, and thus assuming that the models represent a good proxy for the distribution of palaeo-habitats, some predictions regarding expected genetic structure can be generated. First, it is a basic tenet of population genetic theory that ancestral geographic regions will house high genetic diversity. Based on this expectation, we would expect to see the highest levels of genetic diversity associated with regions that are predicted to have maintained stable habitat over time. In this instance, for forest endemics, high genetic diversity would be expected in Borneo, the mountains of Peninsular Malaysia and Sumatra, and some of the Mentawai Islands and central Java. Signatures of recent expansion (e.g. low genetic diversity) would be expected in the forests of southern Malaysia as well as some of the mountainous regions of Sumatra and Java. For open habitat inhabitants, it might be expected that there would be evidence for recent gene flow among mainland SE Asia, Peninsular Malaysia, Sumatra, and Java.

Quek et al. (2007) examined the genetic structure of forest dependent ants, and found that genetic diversity was by far the highest in Borneo (both in lowland and mountain rainforest), although some populations in the mountains of Peninsular Malaysia and Sumatra also had high genetic diversity. Signatures of Pleistocene demographic expansions were evident among many of the lineages. Cannon and Manos (2003) demonstrated very high diversity among Stone Oaks in Borneo compared to populations in mainland Asia, unfortunately, their sampling did not include populations from other Sundaic landmasses. Both of these studies support the suggestion that Borneo has been a stable refugial area, and accord well with palaeo-ecological models, which predict constant forest on the island.

The complex relationship between rainforest expansion, available land area, and glacial-interglacial cycles in Sundaland presents a unique biogeographic scenario and may mean that patterns observed in other tropical areas may be very different from those in tropical Asia. Studies examining how these unique rainforest dynamics affect population structure, phylogeographic patterns, and speciation dynamics are required to

elucidate the complex biogeographic history of this region. Regardless of whether or not contemporary rainforests are refugial or not, the models definitely support the idea that modern lowland rainforests are particularly vulnerable to perturbation since their total area is greatly reduced as compared to their recent past.

Literature Cited

Alroy, J. 2000. New methods for quantifying macroevolutioanry patterns and processes. Paleobiology 26:707-733.

Anshari, G., A. P. Kershaw, and S. Van Der Kaars. 2001. A Late Pleistocene and Holocene pollen and charcoal record from peat swamp forest, Lake Sentarum Wildlife Reserve, West Kalimantan, Indonesia Palaeogeography, Palaeoclimatology, Palaeoecology 171:213-228.

Ashcroft, M. B. 2010. Identifying refugia from climate change. Journal of Biogeography:1407-1413.

Avise, J. C., D. Walker, and G. C. Johns. 1998. Speciation durations and Pleistocene effects on vertebrate phylogeography. Proceedings of the Royal Society B: Biological Sciences 265:1707-1712.

Bennett, K. D., and J. Provan. 2009. What do we mean by 'refugia'? Quaternary Science Reviews 27:2449-2455.

Bird, M., D. Taylor, and C. Hunt. 2005. Palaeoenvironments of insular Southeast Asia during the Last Glacial Period: a savanna corridor in Sundaland? Quaternary Science Reviews 24:2228-2242.

Bonnefille, R. 2007. Rainforest responses to past climatic changes in tropical Africa. Pages 117-170 *in* Tropical Rainforest responses to climatic change (M. B. Bush, and J. Flenley, eds.). Springer-Praxis Books in Environmntal Sciences, Berlin.

Bush, M. B., and P. E. de Oliveira. 2006. The rise and fall of the refugial hypothesis of Amazonian speciation: a paleo-ecological perspective. Biota Neotropica 6.

Bush, M. B., W. D. Gosling, and P. Colinvaux. 2007. Climate change in the lowlands of the Amazon basin. Pages 55-76 *in* Tropical Rainforest Responses to Climatic Change (M. B. Bush, and J. R. Flenley, eds.). Springer-Praxis Books in Environmental Sciences, Berlin.

Cannon, C. H., and P. S. Manos. 2003. Phylogeography of the Southeast Asian stone oaks (Lithocarpus). Journal of Biogeography 30:211-226.

Cannon, C. H., R. J. Morley, and A. B. G. Bush. 2009. The current refugial rainforests of Sundaland are unrepresentative of their biogeographic past and highly vulnerable to disturbance. Proceedings of the National Academy of Sciences 106:11188-11193.

Carnaval, A. C., M. Hickerson, C. Haddad, M. T. Rodrigues, and C. Moritz. 2008. Stability predicts genetic diversity in the Brazialian Atlantic Forest Hotspot. Science 323:785-789.

Carnaval, A. C., and C. Moritz. 2008. Historical climate modelling predicts patterns of current biodiversity in the Brazilian Atlantic forest. J Biogeography 35:1187-1201.

Colinvaux, P., and P. De Oliveira. 2001. Amazon plant diversity and climate through the Cenozoic. Palaeogeography, Palaeoclimatology, Palaeoecology 166:51-63.

Colinvaux, P., P. De Oliveira, and M. Bush. 2000. Amazonian and neotropical plant communities on glacial time-scales: the failure of the aridity and refuge hypotheses. Quaternary Science Reviews 19:141-169.

Colinvaux, P., P. De Oliveira, J. Moreno, M. Miller, and M. Bush. 1996. A long pollen record from lowland Amazonia: forest and cooling in glacial times. Science 274:85.

Cranbrook, E. O. 2010. Late quaternary turnover of mammals in Borneo: the zooarchaeological record. Biodivers Conserv 19:373-391.

Gathorne-Hardy, F., R. Davies, P. Eggleton, and D. Jones. 2002. Quaternary rainforest refugia in south-east Asia: using termites (Isoptera) as indicators. Biological Journal of the Linnean Society.

Gibbard, P., and K. M. Cohen. 2008. Global chronostratigraphical correlation table for the last 2.7 million years. Episodes 31:243-247.

Gobbett, D., and C. Hutchison. 1973. Geology of the Malay Peninsula (West Malaysia and Singapore). Pages 438 *in* Regional Geology Series (L. DeSitter, ed.) Wiley-Interscience, New York.

Haffer, J. 1969. Speciation in Amazonian forest birds. Science 165:131-137.

Haffer, J., and G. T. Prance. 2001. Climatic forcing of evolution in Amazonia during the Cenozoic: on the refuge theory of biotic differentiation. Amazoniana 16:609-646.

Hall, R. K., and C. K. Morley. 2004. Sundaland Basins. Pages 55-85 *in* Continent-ocean interactions within East Asian marginal seas (P. D. Clift, P. Wang, W. Kuhnt, and D. Hayes, eds.). American Geophysical Union, Washington DC.

Heaney, L. 1991. A synopsis of climatic and vegetational change in Southeast Asia. Climatic Change.

Heine, C., R. D. Muller, B. Steinberger, and T. H. Torsvik. 2008. Subsidence in intracontinental basins due to dynamic topography. Physics of the Earth and Planetary Interiors 171:252-264.

Hijmans, R., S. E. Cameron, J. L. Parra, P. G. Jones, and A. Jarvis. 2005. Very high resolution interpolated climate surfaces for global land areas. International Journal of Climatology 25:1965-1978.

Horton, B. P., P. L. Gibbard, G. M. Milne, R. J. Morley, C. Purintavaragul, and J. M. Stargardt. 2005. Holocene sea levels and palaeoenvironments, Malay-Thai Peninsula, southeast Asia. The Holocene 15:1199-1213.

Inger, R., and H. K. Voris. 2001. The biogeographical relations of the frogs and snakes of Sundaland. J Biogeography.

Liechti, P., F. Roe, and N. Haile. 1960. The Geology of Sarawak, Brunaei, and the Western Part of North BorneoGeologial Survey Department, British Territories in Borneo, Kuching, Sarawak.

Meijaard, E. 2003. Mammals of south-east Asian islands and their late Pleistocene environments. Journal of Biogeography 30:1245-1257.

Miller, K. G., M. A. Kominz, J. V. Browning, J. D. Wright, G. S. Mountain, M. E. Katz, P. J. Sugarman, B. S. Cramer, N. Christie-Blick, and S. F. Pekar. 2005. The Phanerozoic record of global sea-level change. Science 310:1293-8.

Morley, R. J., and J. R. Flenley. 1987. Late Cainozoic vegetational and environmental changes in the Malay archipelago. Pages 50-59 *in* Biogeographical evolution of the Malay archipelago (T. C. Whitemore, ed.) Clarendon Press, Oxford.

Müller, R. D., M. Sdrolias, C. Gaina, B. Steinberger, and C. Heine. 2008. Long-term sea-level fluctuations driven by ocean basin dynamics. Science 319:1357-62.

Newsome, J., and J. Flenley. 1988. Late Quaternary Vegetational History of the Central Highlands of Sumatra. II. Palaeopalynology and Vegetational History. J Biogeography 15:555-578.

NOAA. 2006. ETOPO2v2 Global Gridded 2-minute database(N. G. D. Center, ed.) National Oceanic and Atmospheric Administration, US Department of Commerce.

Otto-Bliesner, B. L., S. J. Marshall, J. T. Overpeck, G. H. Miller, and A. Hu. 2006. Simulating Arctic climate warmth and icefield retreat in the last interglaciation. Science 311:1751-3.

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

Phillips, S. J., M. Dudık, and R. E. Schapire. 2004. A Maximum Entropy Approach to Species Distribution Modeling Proceedings of the 21st International Conference on Machine Learning:8.

Quek, S., S. Davies, P. Ashton, T. Itino, and N. Pierce. 2007. The geography of diversification in mutualistic ants: a gene's-eye view into the Neogene history of Sundaland rain forests. Mol. Ecol. 16:2045-62.

Ray, N., and J. M. Adams. 2001. A GIS-based vegetation map of the world at the last glacial maximu, (25,000-15,000bp). Internet Archaeology 11:1-43.

Richards, C. L., B. C. Carstens, and L. Lacey Knowles. 2007. Distribution modelling and statistical phylogeography: an integrative framework for generating and testing alternative biogeographical hypotheses. J Biogeography 34:1833-1845.

Sathiamurthy, E., and H. K. Voris. 2006. Maps of Holocene Sea Level Transgression and Submerged Lakes on the Sunda Shelf. Natural History.

Slik, J. W. F., S.-I. Aiba, M. Bastian, F. Q. Breasley, C. H. Cannon, K. A. O. Eichhorn, G. Fredriksson, K. Kartawintata, Y. Laumonier, A. Mansor, A. Marjokorpi, E. Meijaard, R. J. Morley, H. Nagamasu, R. Nilus, E. Nurtjahya, J. Payne, A. Permana, A. D. Poulsen, N. Raes, S. Riswan, C. P. van Schaik, D. Sheil, K. Sidiyasa, E. Suzuki, J. C. C. H. van Valkenburg, C. O. Webb, S. Wich, T. Yoneda, R. Zakaria, and N. Zweifel. 2011. Soils on exposed Sunda Shelf shaped biogeographic patterns in the equatorial forests of Southeast Asia. Proceedings of the National Academy of Science 108:12343-12347.

Stibig, H.-J., A. S. Belward, P. S. Roy, U. Rosalina-Wasrin, S. Agrawal, P. K. Joshi, R. Beuchle, S. Fritz, S. Mubareka, and C. Giri. 2007. A land-cover map for South and Southeast Asia derived from SPOT-VEGETATION data. J Biogeography 34:625-637.

Stuijts, I., J. C. Newsome, and J. R. Flenley. 1988. Evidence for late quaternary vegetational change in the Sumatran and Javan highlands. Review of Palaeobotany and Palynology 55:207-216.

Stujits, I., J. C. Newsome, and J. R. Flenley. 1988. Evidence for late quaternary vegetational change in the Sumatran and Javan highlands. Review of Palaeobotany and Palynology 55:207-216.

Sun, X., X. Li, Y. Luo, and X. Chen. 2000. The vegetation and climate at the last glaciation on the emerged continental shelf of the South China Sea. Palaeogeography, Palaeoclimatology, Palaeoecology 160:301-316.

Tougard, C. 2001. Biogeography and migration routes of large mammal faunas in South-East Asia during the Late Middle Pleistocene: focus on the fossil and extant faunas from Thailand. Palaeogeography, Palaeoclimatology, Palaeoecology 168:337-358.

van den Bergh, G. D., J. de Vos, and P. Sondaar. 2001. The Late Quaternary palaeogeography of mammal evolution in the Indonesian Archipelago. Palaeogeography.

Van Der Kaars, S., and M. A. C. Dam. 1995. A 135,000-year record of vegetational and climatic change from the Bandung area, West Java, Indonesia. Palaeogeography, Palaeoclimatology, Palaeoecology 117:55-72.

VanDerWal, J., L. P. Shoo, and S. E. Williams. 2009. New approaches to understanding late Quaternary climate fluctuations and refugial dynamics in Australian wet tropical rain forests. Journal of Biogeography 36:291-301.

Voris, H. K. 2000. Maps of Pleistocene sea levels in southeast asia: shorelines, river systems and time durations. Journal of Biogeography 27:1153-1167.

Wang, X., X. Sun, P. Wang, and K. Stattegger. 2007. A high-resolution history of vegetation and climate history on Sunda Shelf since the last glaciation. Science in China Series D: Earth Sciences 50:75-80.

Wang, X., X. Sun, P. Wang, and K. Stattegger. 2009. Vegetation on the Sunda Shelf, South China Sea, during the Last Glacial Maximum. Palaeogeography, Palaeoclimatology, Palaeoecology 278:88-97.

Wogan, G. O. U. 2011a. Chapter 2. Shifting habitats across glacial and interglacial Asia. Pages 28 + figures *in* Integrative Biology University of California, Berkeley, Berkeley.

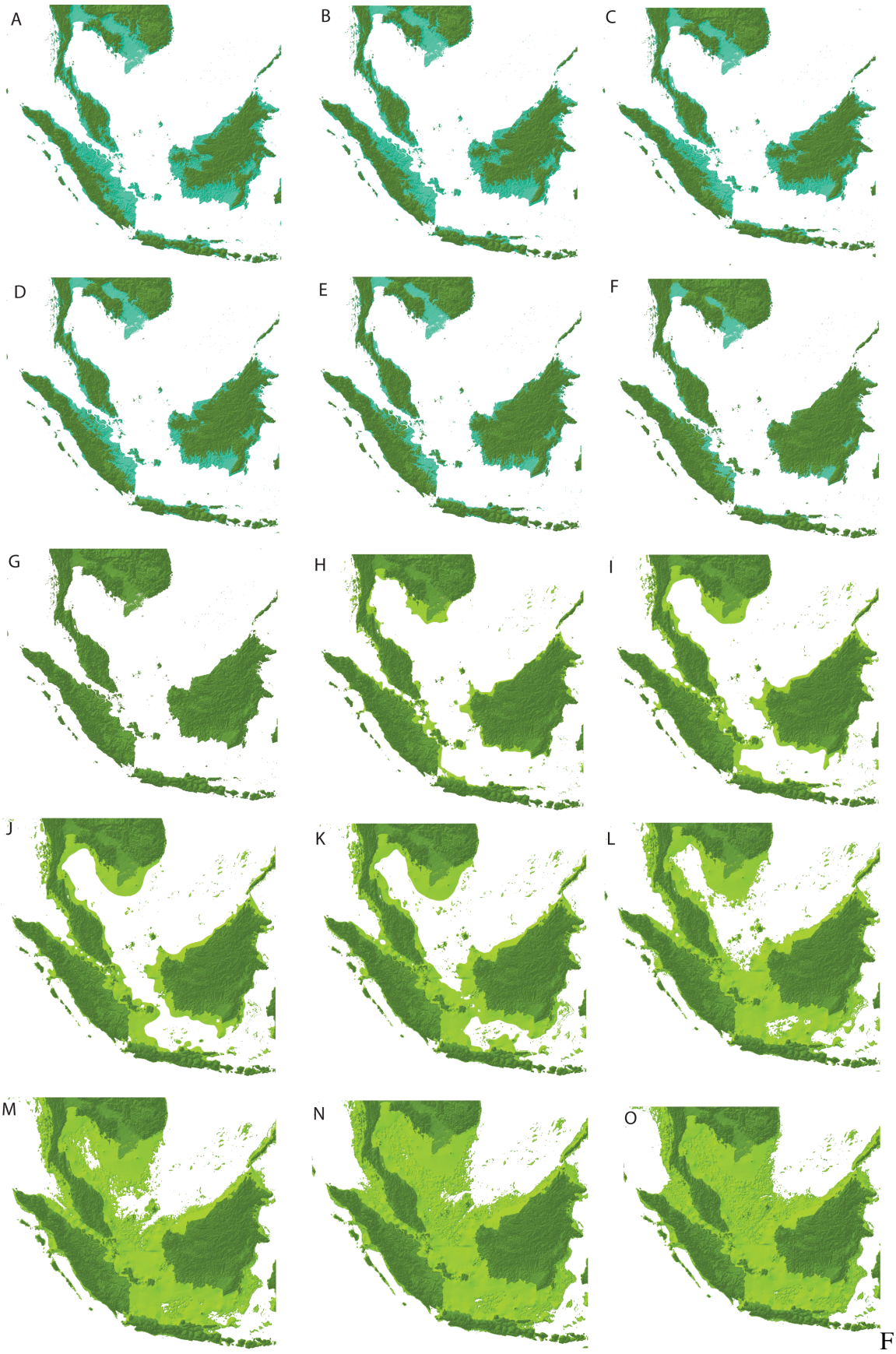
Wogan, G. O. U. 2011b. Chapter 4. Reconstructing ancient pathways: The Habitats of Sundaland during Glacials and Interglacials*in* Integrative Biology University of California, Berkeley, Berkeley.

Woodruff, D. S., and L. M. Turner. 2009. The Indochinese-Sundaic zoogeographic transition: a description and analysis of terrestrial mammal species distributions. J Biogeography 36:803-821.

Wurster, C. M., M. I. Bird, I. D. Bull, F. Creed, C. Bryant, J. A. J. Dungait, and V. Paz. 2010. Forest contraction in north equatorial Southeast Asia during the Last Glacial Period. Proceedings of the National Academy of Sciences:1-4.

Zink, R., and J. Slowinski. 1995. Evidence from molecular systematics for decreased avian diversification in the pleistocene Epoch. Proceedings of the National Academy of Science 92:5832-5.

Figure 1. Changing Sea levels have greatly affected the topography of Sunda region. Maps depicting Sunda with sea levels ranging from +60 to –120 m. are shown. Maps A-F depict changes with sea levels above current sea levels, blue indicates regions that would be inundated with rising sea levels during interglacial times. G depicts current sea level. H-O depicts the region with sea levels lower than present, light green highlights areas currently below sea level that become land positive as sea levels decrease during glacial times. A) +60m B) +50m C) +40m D) +30m E) +20m F) +10m G) 0m (now) H) -10m I) -20m J) -30m K) -40m L) –50m M) –70m N) –90m O) -120m



F

Figure 2. Simple bar graph depicting the total amount of time (in thousands of years) at different sea levels (relative to contemporary sea levels) over the past 3 million years. As can be observed, for the majority of the last 3 million years, sea levels were between 0 and –50 m, with the most time at –20 m below present.

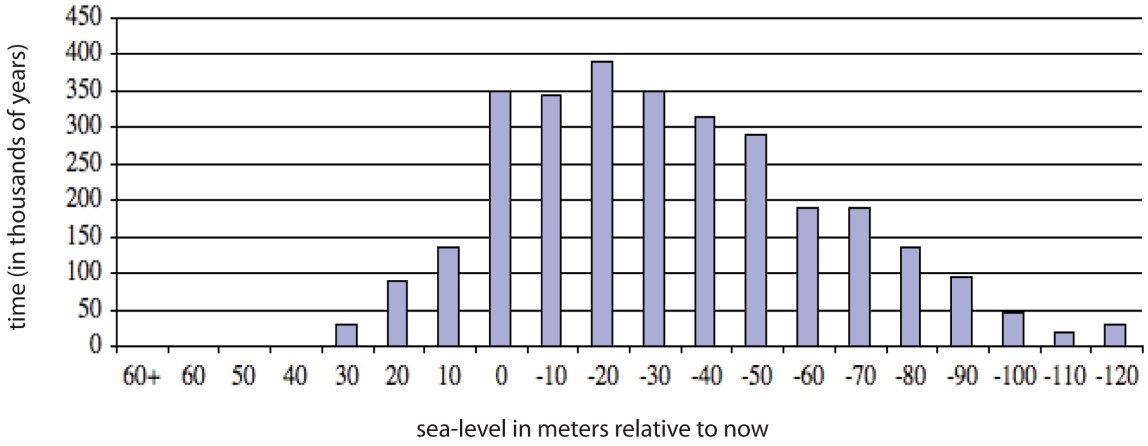
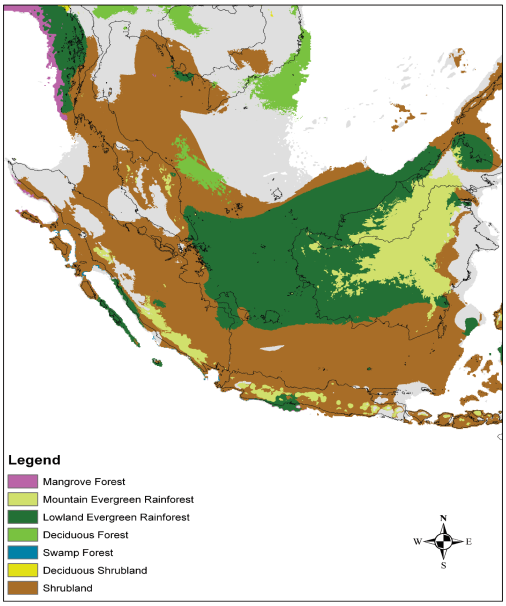
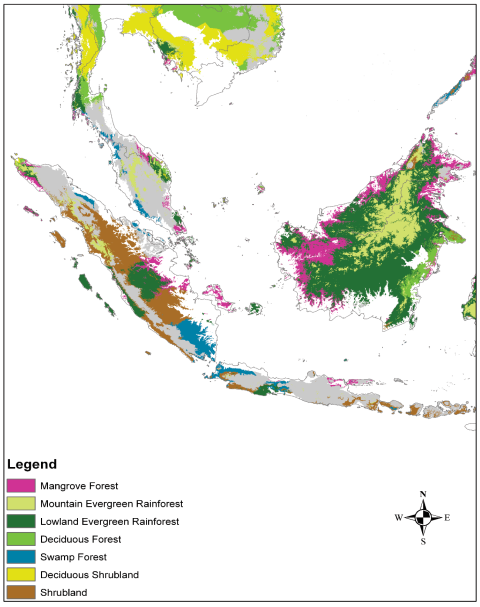
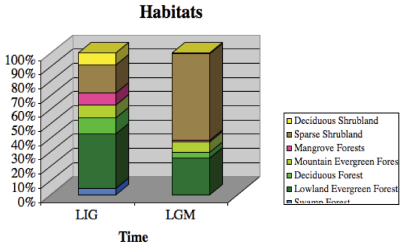

time (in thousands of years)
450
400
350
300
250
200
150
100
50
0
60+ 60 50 40 30 20 10 0 -10 -20 -30 -40 -50 -60 -70 -80 -90 -100 -110 -120
sea-level in meters relative to now

Figure 3. Contemporary distribution of natural habitats in the Sunda region (after Stibig et al. 2007). White areas are human dominated habitats (cities, agriculture, plantations).

Figure 4. Distribution of habitats at the last interglacial (120 ky) and the last glacial maximum (21ky). Light gray areas depict land positive portions of the Sundaland continent, but where models failed to predict any of the seven habitats modeled here (above the prescribed threshold). Black lines indicate current geopolitical boundaries for reference. The bar graph depicts the relative proportions of habitats at the last interglacial and last glacial maximum. The donut depicts the proportion of forest versus non-forest habitats during the last interglacial (inner circle) and the last glacial maximum (outer circle).


Legend
Mangrove Forest
Mountain Evergreen Rainforest
Lowland Evergreen Rainforest
Deciduous Forest
Swamp Forest
Deciduous Shrubland
Shrubland
N
W
E
S


Legend
Mangrove Forest
Mountain Evergreen Rainforest
Lowland Evergreen Rainforest
Deciduous Forest
Swamp Forest
Deciduous Shrubland
Shrubland
N
W
E
S


Habitats
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
LIG
LGM
Time
Deciduous Shrubland
Sparse Shrubland
Mangrove Forests
Mountain Evergreen Fores
Deciduous Forest
Lowland Evergreen Forest

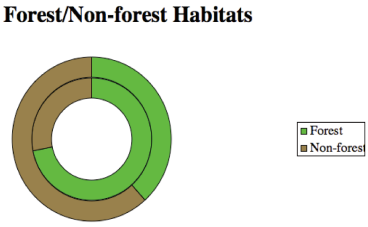

Forest/Non-forest Habitats
Forest
Non-forest

Figure 5. Models of forest/ non-forest habitat distribution during the last glacial maximum with independent data overlaid. Blue symbols indicate agreement with the predictive models, while grey symbols indicate ambiguity with regard to the models: (Δ) isotope inference from cave guano (Wurster et al., 2010); (o) palynology core samples (Anshari et al., 2001; Morley and Flenley, 1987; Newsome and Flenley, 1988; Stuijts et al., 1988; Sun et al., 2000; Van Der Kaars and Dam, 1995; Wang et al., 2007; Wang et al., 2009). Black hatched area depicts putative rainforest refugia identified from forest-dependent termites (Gathorne-Hardy et al., 2002).

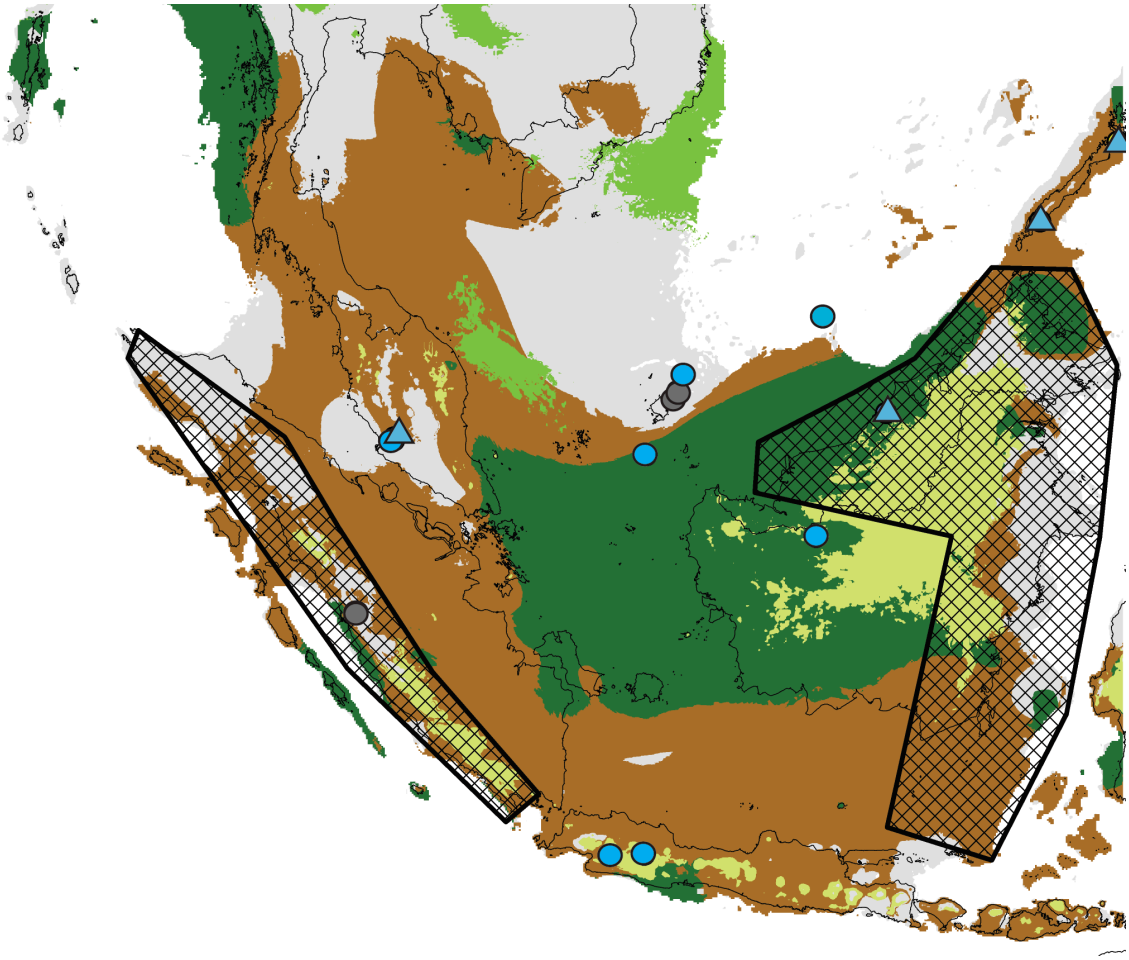


Table 1. Sea-levels for the past 3 million years binned into ten meter increments. The total time sea-levels were within each bin (not contiguous) and the percentage of time over the pat 3 million years.

Max Sea-level	Min Sea-level	Total Time	Percent of Last 3 million years
> 60	60	0	0
59	50	0	0
49	40	0	0
39	30	0	0
29	20	30,000	1
19	10	90,000	3
9	0	135,000	4.5
-1	-9	350,000	11.7
-10	-19	345,000	11.5
-20	-29	390,000	13
-30	-39	350,000	11.7
-40	-49	315,000	10.5
-50	-59	290,000	9.7
-60	-69	190,000	6.3
-70	-79	190,000	6.3
-80	-89	135,000	4.5
-90	-99	95,000	3.2
-100	-109	45,000	1.5
-110	-119	20,000	0.7
-120	< -120	30,000	1

Table 2. Bioclimatic variables used in predictive modeling.

Bioclim 1	Annual mean temperature
Bioclim 2	Mean diurnal range
Bioclim 3	Isothermality
Bioclim 4	Temperature seasonality
Bioclim 5	Max temperature of warmest month
Bioclim 6	Min temperature of coldest month
Bioclim 7	Temperature annual range
Bioclim 8	Mean temperature of wettest quarter
Bioclim 9	Mean temperature of driest quarter
Bioclim 10	Mean temperature of warmest quarter
Bioclim 11	Mean temperature of coldest quarter
Bioclim 12	Annual precipitation
Bioclim 13	Precipitation of wettest month
Bioclim 14	Precipitation of driest month
Bioclim 15	Precipitation seasonality
Bioclim 16	Precipitation of wettest quarter
Bioclim 17	Precipitation of driest quarter
Bioclim 18	Precipitation of warmest quarter
Bioclim 19	Precipitation of coldest quarter

CHAPTER 5

MODEL SELECTION, PARTITIONING SCHEMES, AND PHYLOGENETIC INFORMATIVENESS: A CASE STUDY OF THE DICROGLOSSIDAE (AMPHIBIA: ANURA)

INTRODUCTION

Data partitioning strategies and selection of evolutionary models of nucleotide substitution underlie Bayesian and ML phylogenetic analyses. A suite of recent studies have evaluated how choice of nucleotide substitution models impact phylogeny estimation (Lemmon and Moriarty, 2004; Ripplinger and Sullivan, 2008; Ripplinger and Sullivan, 2010). The choice of nucleotide substitution model effects estimated topologies (Ripplinger and Sullivan, 2010), bipartition estimation (Lemmon and Moriarty, 2004; Ripplinger and Sullivan, 2010), and branch lengths, particularly when rate heterogeneity is ignored (Lemmon and Moriarty, 2004). Ripplinger and Sullivan (2010) demonstrated that while model misspecification does impact topologies, primarily only poorly supported nodes are effected. Under-parameterization can bias branch length and gamma shape parameters (Lemmon and Moriarty, 2004), as well as tree topologies and support values (Ripplinger and Sullivan, 2010). Conversely, over-parameterization decreases the precision of parameter estimates, which can induce error into phylogeny estimation (Lemmon and Moriarty, 2004).

Differing methods for inferring the best-fit model include the Akaikie Information criterion (AIC), the sample size corrected Akaikie Information criterion (AICc), the Bayesian Information Criterion (BIC), and the Decision Theoretic (DT), as well as popular frequentist methods such as hierarchical likelihood ratio tests (hLRTs). While all of these methods have demonstrated utility in model selection, systemic biases among model selection criteria have recently been demonstrated. Ripplinger and Sullivan (2008; 2010) found that the AICc tends to favor more complex models of evolution as compared to the BIC and DT. They also found that tests of model adequacy often failed to reject simple models of nucleotide substitution.

Multiple recent studies have evaluated the impact of data partitioning on phylogeny estimation (Brandley et al., 2005; Brown and Lemmon, 2007; McGuire et al., 2007) and have neatly demonstrated that partitioning schema can greatly impact the estimation of posterior probabilities and phylogeny. Inappropriate over-partitioning can impact the ability of the data to accurately estimate parameters since the amount of data decreases per parameter as the number of partitions increases (Brandley et al., 2005), while under-partitioning utilizes inappropriate models of evolution for particular segments of the data and can provide misleading inference (Brown and Lemmon, 2007) leading to phenomena such as long-branch attraction.

Choices involving both partitioning and nucleotide substitution models can have dramatic impacts on the ability to reconstruct phylogenetic trees, and in course affect all subsequent tree-based analyses of evolutionary phenomona. Here, I investigate the joint impacts of partitioning scheme and nucleotide substitution model on phylogenetic hypotheses. I begin by estimating gene trees for each marker under nucleotide substitution models selected using three different selection criteria. I generate phylogenetic hypotheses using mixed-model Bayesian and Maximum likelihood analyses

under six partitioning schemes, and implementing nucleotide substitution models selected under different criteria. I test the model adequacy for independent and concatenated analyses. Lastly, I evaluate the impact of model selection and partitioning schema on the phylogenetic informativeness (PI) of each marker, tree topology, and likelihoods estimated using Bayesian Inference and Maximum Likelihood methods.

Phylogenetic Informativenss (PI) is a measure of how well any marker is performing within a phylogenetic context. It is the joint probability of a character state change occurring on a short internal branch and the lack of a subsequent character state change on the long tips (Townsend, 2007). It is estimated by first calculating the optimal rate of character state changes along the branches in a phylogenetic tree which maximizes the joint probability of a change occurring on a short internal branch and the lack of change on subsequent terminal tips. The ratio of empirically derived rates of character change and the optimal rates of character change are then calculated to estimate the PI of the characters (Townsend, 2007). Generally, PI has been used as a means of estimating the potential utility of specific markers to resolve particular branching events, and as such temporal partitioning of PI can be used to asses when a particular marker is most useful. Here I use PI to understand how model selection and partitioning scheme can affect phylogenetic inference. Topological concordance among trees was quantified using the *l*cong statistic which is the index of congruency, which estimates the similarity between trees (de Vienne et al., 2007), and a series of additional tree statistics were examined in order to better understand how model selection and partitioning scheme impact tree shape and symmetry.

The study system: The family Dicroglossidae is an Old World family of neobatrachian frogs formerly identified as a subgroup of Ranidae. They are found throughout temperate and tropical regions in Africa, India, and Asia with most of the species diversity located in Asia. The monophyly of the family and the relationships among clades within the family have remained unclear with alternative analyses supporting differing evolutionary hypotheses (Bossuyt et al., 2006; Che et al., 2009; Che et al., 2007a; Che et al., 2007b; Chen et al., 2005; Dubois and Ohler, 2005; Frost et al., 2006; Scott, 2005; Van Der Meijden et al., 2005). There are two generally recognized subfamilies the Occidozyginae consisting of the genera *Ingerana* and *Occidozyga*, and the Dicrglossinae consisting of *Allopaa, Chrysopaa, Euphlyctis, Fejervarya, Hoplobatruchus, Limnonectes, Minervarya, Nannophrys, Nanorana, Ombrana, Quisipaa,* and *Sphaerotheca*. While there have been phylogenetic analyses of both the higher level relationships (e.g. the monophyly of the family) and within family relationships (e.g. the generic and species level relationships), these have been largely generated utilizing single to few markers, or have had limited taxonomic sampling, or have been based on phenetic analyses. Here I expand the taxonomic sampling within the Dicroglossidae and sequence a greater number of independent loci in order to better characterize the phylogenetic relationships of the genera within the family and to assess the membership and monophyly of the family.

METHODS

Taxonomic Sampling

In order to test the monophyly of the family Dicroglossidae, outgroups from both Asian and African ranoids as well as a single bufonoid taxon were selected for inclusion in phylogenetic analyses. Ranoid outgroups included representatives from the families Arthroleptidae, Petropedetidae, Phrynobatracidae, Ptychadinidae, Pyxicephalidae, Ranidae, Hyperoliidae, and Mantellidae. Generic level species assignments and relationships have been in a state of flux with recently proposed molecular based taxonomies (Bossuyt et al., 2006; Che et al., 2009; Che et al., 2007a; Che et al., 2007b; Che et al., 2010; Frost et al., 2006; Van Der Meijden et al., 2005) challenging morphological hypotheses (Dubois, 1992; Dubois and Ohler, 2005). Within the Dicroglossidae, a number of taxonomies have been proposed which result in very different familial, and subfamilial taxonomies (Table 1). The sampling here attempts to include representatives of all genera in all taxonomies. However, samples of three recently erected or described monotypic genera *Minerverya* (Ohler et al., 2009), *Chrysopaa* (Ohler and Dubois, 2006), and *Ombrana* (Dubois, 1992) were not available for inclusion nor was the recently erected genus *Allopaa* (Ohler and Dubois, 2006). To further test the multiple generic level hypotheses, multiple species representing alternative taxonomic schemes were included within these analyses as possible. Table 1 provides a list of all species included in this study as well as the traditional morphological and recent molecular based taxonomies.

Marker Selection and Sequencing

Total genomic DNA was extracted from flash frozen or ethanol preserved tissues using a salt extraction protocol. Amplicon were purified using Exo-sap (USB Corp.), ethanol precipitated, then sequenced using ABI BigDye chemistry on an ABI 3730 capillary sequencer. All fragments were bi-directionally sequenced. Sequences were edited in Sequencher v. 4.8 (GeneCodes).

The taxa of interest encompass tens of millions of years of diversity, so marker selection optimized markers that reliably amplified across divergent taxa but also displayed variation. A total of 13 markers were sequenced for these analyses. A single mitochondrial marker the 16S ribosomal gene (16s) was sequenced. Twelve nuclear markers were sequenced, these include ADNP (activity-dependent neuroprotector), BDNF (brain-derived neurotrophic factor), CYMex3 (proto-oncogene cellular myelocytomatosis), CXCR4 (chemokine receptor type 4), FSTL5:(follistatin-like 5), NCX1 (sodium/calcium ion exchanger), NTS3: (3'-nucleotidase), PTGER4: (prostoglandin E receptor 4 (subtype EP4)), Rhodopsin: (rhodopsin exon 1), SIA: (seven in absentia), TYRO (Tyrosinase), ZEB 2 (zinc finger E-box binding homeobox 2). A thirteenth nuclear marker RAG1 (recombination activating gene 1) compiled solely from genbank data was also incorporated in these analyses for a total of 14 markers. Prior to alignment, this dataset comprises 10,725 bp of sequence data per individual. The initial and aligned lengths of each marker are reported in Table 2.

Sequence Analyses

Alignment and Model Selection: Edited sequences were aligned using MUSCLE (Edgar, 2004) and then further refined. Protein coding genes were translated to amino acids, with codon positions selected to eliminate stop codons, and amino acid alignment was then checked by eye. Geneious v 5.3.6 (Drummond et al., 2010) was utilized to calculate Pairwise Identity, and the number of SNPs within each gene. SNPs were detected by setting minimum coverage to 50% and the minimum frequency to 0.1.

Models of nucleotide evolution were evaluated using jModeltest v.0.1.1 (Posada, 2008). Model analyses included 24 models under three substitution selection criteria. All sets of analyses were evaluated using an optimized maximum likelihood framework in which an optimal maximum likelihood tree is generated in Phyml (Guindon et al., 2010), and then the likelihood of each nucleotide substitution model is calculated using the tree (Posada, 2008). Models of nucleotide evolution were evaluated for each gene, and for each codon position within each gene. Note that in jmodeltest, partitioning of codon positions must be done individually, and these individual files were prepared using PAUP * 4.0 (Swofford, 2002). AICc, BIC, and DT model selection criteria were utilized to infer the best fitting model. Both the AICc and BIC test statistics were calculated using the sequence length as the sample size.

Model Adequacy: To assess the adequacy of nucleotide model specification, posterior probabilities from Bayesian analyses specified under each nucleotide substitution scheme indicated under AICc, BIC, or DT models for independent analyses were used to calculate the multinomial likelihood statistic and then compared with the same statistic calculated from simulation analyses as implemented in PuMA (Brown and ElDabaje, 2009). The multinomial likelihood statistic T(x) gives a measure of how well the model represents the actual data, and determines the predictive probability that the model is inadequate (Bollback, 2002). The T(x) statistic is generated from a posterior distribution of trees and parameter values obtained from a Bayesian analysis (details of Bayesian gene tree analyses provided below). Using seq-gen (Rambaut and Grassly, 1997), PuMA evolves sequence data under the specified nucleotide substitution model onto those trees, and generates a distribution that allows comparison of the multinomial likelihood to that derived from the empirical data and provides a P value. The P value can be thought of as the probability that the model would generate a value as extreme as the value observed from the empirical data (Bollback, 2002). Values close to 0.5 indicate a good fit, and the further the departure from 0.5 the worse the fit of the model to the data. Since the empirical data contained some gaps and missing data, and PuMA does not currently allow for either, both missing and gapped sites were removed from analyses. For coding regions, basepairs were grouped sequentially by codon position before gapped and missing sites were removed to ensure that removal did not affect codon assignment (J. Brown pers. comm.), files were prepared using PAUP* 4 (Swofford, 2002) and Phyutility v.2.2 (Smith and Dunn, 2008).

Next I used Bayes factors to compare the likelihoods of gene trees for partitioned and unpartitioned nucleotide substitution models selected under AICc, BIC, or DT, or as

well as under the GTR + I + Γ model. Bayes factors were calculated as in Suchard et al. (2001), with Bayes Factors >10 considered strong support (Brown and Lemmon, 2007).

Partitioning Scheme: Six partitioning schemes were investigated in these analyses: (p1) a single partition under which all basepairs evolve under the same model (GTR+ I + Γ), (p2) two partitions in which mtDNA and nuclear markers evolve independently under the same model (GTR+ I + Γ), (p3) 14 partitions in which each gene evolves independently under the GTR+ I + Γ model, (p4) 40 partitions in which each codon position within each individual gene evolves independently under the GTR+ I + Γ model, (p5) 14 partitions in which each gene evolves independently under the best fit model selected by AICc, BIC and DT, (p6) 40 partitions in which each codon position within each individual gene evolves independently under the best fit model selected by AICc, BIC, and DT. All six partitioning schemes were examined within a Bayesian Inference framework, but only the first four were examined using Maximum Likelihood. Optimal partitioning schemes for both gene trees and concatenated analyses were evaluated using Bayes factors (Brandley et al., 2005; Brown and Lemmon, 2007; McGuire et al., 2007). As above, Bayes factors were calculated in accordance with the method outlined in Suchard et al. (2001), with Bayes Factors >10 considered strong support (Brown and Lemmon, 2007).

Phylogenetic Analyses-Gene Trees: Maximum likelihood gene trees were generated using RaxML blackbox (Stamatakis, 2004; Stamatakis et al., 2008). RaxML utilizes the GTR+I+Γ model of substitution so unpartitioned and codon partitioned analyses and bootstrapping analyses were carried out under this model with gamma rate heterogenity. In all instances, 100 bootstraps were generated under the GTRCAT model, while the tree search was executed using the GTRGAMMA. Bayesian gene trees were estimated utilizing MrBayes v. 3.1.2 (Huelsenbeck and Ronquist, 2002). Bayesian analyses were run with three heated and one cold chain for 15-30 million generations sampled every 1000 generations, two independent runs were performed for each analysis. Both AWTY (Wilgenbusch et al., 2004) and the PSRF statistic reported in MrBayes were utilized to assess stationarity. In AWTY, two plots were utilized to assess stationarity; the cumulative plot of the posterior probability of splits were generated for the 20 most variable splits divided into 20 increments, and plots comparing the posterior probabilities of all splits both with 25% burn in. The averaged harmonic mean and the averaged arithmetic mean were obtained using the sump command in MrBayes. Gene trees were estimated under each of the three best fit nucleotide substitution models selected by AICc, BIC, and DT as well as under the GTR+ I + Γ model.

Phylogenetic Analyses-Concatenation: Partitioned ML and Bayesian analyses of the concatenated data were run using RaxML blackbox (Stamatakis, 2004; Stamatakis et al., 2008) and MrBayes 3.2.1 (Huelsenbeck and Ronquist, 2002) respectively. Bayesian analyses were run with three heated and one cold chain for 40 million generations sampled every 1000 generations. Stationarity was assessed as detailed above using AWTY (Wilgenbusch et al., 2004).

Tree Statistics: Topological concordance between pairs of trees generated via

concatenation analyses was assessed using the index of congruency test statistic (*l*cong) implemented on the *l*cong webserver (de Vienne et al., 2007). The *l*cong statistic estimates the MAST (maximum agreement subtrees) shared between two trees, with the number of shared leaves is indicative of the shared concordance (de Vienne et al., 2007). In these analyses, maximum concordance would mean that 185 of 185 taxa were shared in the MAST. TreeStat v1. 2 (Rambaut and Drummond, 2011) was used to calculate tree length, tree height, the $\bar{N}$ statistic (Kirkpatrick and Slatkin, 1993), treeness (Phillips et al., 2001), and the ratio of external to internal branch lengths. The $\bar{N}$ statistic is a measure of tree symmetry. $\bar{N}$ is the number of internal branches between the root and the tips. The analytical value is compared to a distribution of random draws from random topologies with the same number of taxa to assess if the tree is symmetric (Kirkpatrick and Slatkin, 1993). The treeness statistic measures the stemminess of the tree by measuring the proportion of tree distance on internal branches. Treeness values (T) can be thought of as a signal to noise measure if signal is considered to be characters that land on internal branches and support taxonomic groupings and noise to be other state changes. Treeness values (T) range from 0-1 with values closer to 1 indicative of a high signal to noise ratio (Phillips et al., 2001; Phillips and Penny, 2003). Lastly, the ratio of external to internal branch lengths is a measure of the starlike-quality of the tree. All trees were rooted with *Bufo crocus* (Wogan et al., 2003).

Phylogenetic Informativeness: I use PI as a means to assess how model selection and partitioning scheme affect phylogenetic inference. I do this by holding the data matrix constant and evaluating the PI of individual markers on phylogenies estimated using different models and partitioning schemes. Phydesign is an online web-interface (http://phydesign.townsend.yale.edu/ (Lopez-Giraldez and Townsend, 2011) designed to measure PI (Townsend, 2007). Since ultrametric trees are required for these analyses, and the analyses generated here are not time-calibrated, ultrametric trees that simply reflect the relative time were utilized (Fong and Fujita, 2011). TreeEdit v.1.0a10 (Rambaut and Charleston, 2001) was used to convert either the best ML tree or the consensus tree from Bayesian analyses to ultrametric. Time is simply reported as time 0 at tips and scaled to time 1 at the roots. HyPhy (Pond et al., 2005) was used to calculate the rates under empirical base frequencies and a time-reversible model with transitions=2 and transversions=1. Each analysis was performed using the same dataset with each gene identified as a character set, but with the tree estimated from each of the different ML or Bayesian analyses. This approach allows assessment of the PI of each gene under each of the differing inferential parameter combinations.

RESULTS

The genes sequenced here ranged from 315-1398 bp and the number of SNPs ranged from 0-375. Genes with 0 SNPs were not invariable, but SNPs were not detected under the criteria utilized here. Relaxing the minimum coverage requirement revealed additional SNPs. Pairwise identity and GC content ranged from 84.8-95.9% and 38.3-51.2%, respectively (Table 2). The combined analyses consisted of 9382 bp. A total of 324 individuals were in the dataset, redundant haplotypes were collapsed for a final dataset of 185 taxa.

Nucleotide Substitution Models: A total of 160 model selection analyses were performed (Tables 3, 4). AICc and BIC selected models were in agreement more than AICc and DT or BIC and DT selected models. AIC selected models tended to be less parameter rich than BIC models and more parameter rich than DT selected models. (Table 5).

Tests of model adequacy revealed that in very few cases were the nucleotide substitution models able to model the empirical data adequately (Table 7). The models that were best able to model the data were the most complex parameter rich GTR+ I + Γ codon models. For some genes (e.g. PTGER4), none of the models of evolution selected by AICc, BIC, DT, or the GTR+ I + Γ models were able to capture the complexity of the underlying nucleotide substitution pattern (Table 7).

Not surprisingly, analyses of likelihood scores estimated from individual gene trees, suggested that partitioning by codon greatly improved the likelihood scores of the data, and that in every case, the GTR + I + Γ codon model had the best likelihood score. Note that the likelihood scores calculated from RaxML are not directly comparable to those calculated from Bayesian analyses, but are included in the graphs to demonstrate that codon partitioning also greatly influences the likelihood score (Figure 1, Table 6).

Bayes factors comparing harmonic mean log likelihoods of unpartitioned gene trees to codon partitioned gene trees overwhelmingly supported the use of codon partitioned models, although there were a few exceptions in which the unpartitioned model had higher supprt (Table 8a). In most instances the most complex GTR+ I + Γ codon partitioned models were favored (Table 8b). The exceptions are the AICc selected models were favored for CMYex, and Sia, and the DT selected models were favored for CXCR4 and Rag1. In only six of the comparisons were the Bayes Factors less than 10.

Partitioning Models: For the concatenated Bayesian analyses, the GTR model was evaluated under four partitioning schemes, and the AICc, BIC, and DT models were evaluated under 2 partitioning schemes. Bayes factor analyses support the most parameter rich and most partitioned GTR40 model over all other models (Table 9). The BIC40 model is also supported. Based on the fact that the BIC selected more parameter rich models than the AICc and DT, this is likely to be the second most parameter rich and highly partitioned model. Since the BIC40 is less parameter rich than the GTR40, and the goal is to select the model that best represents the data while minimizing unnecessary parameters (McGuire et al., 2007), these analyses support the use of this tree as the primary framework for further discussion of Dicroglossid phylogenetics.

Tree Statistics: Measures of tree length varied greatly among the concatenation trees, from 4.51-1296.87, demonstrating the effect of model selection and partitioning scheme on tree topology (Table 10). The symmetry of the tree remained stable across all of the trees evaluated here with $\bar{N}$ values hovering between 13.27 – 14.87 (Table 10). Treeness (T) values ranged from .3749 - .454 across all of the tree topologies. The relatively low <0.5 values obtained here suggest that there is quite a bit of noise in the signal, e.g. few characters are falling onto internal branches and lending support to

groupings. In particular, the T value was particularly poor for the ML tree with two partitions (mtDNA versus all nuclear autosomal markers), while the highest T value was obtained in the parameter rich GTR + I + Γ Bayesian model partitioned by gene and codon (Table 10). The external/internal branch length ratio was relatively stable across all of the Bayesian trees, but more dramatic values emerged from the ML analyses, with the ratio dramatically changing across partitioning schemes (Table 10). Tree topology was found to be statistically congruent in all pairwise comparisons, although, the number of taxa encompassed by the MAST dropped drastically in some of the comparisons, meaning that there was topological variation among the trees (Table 11). The number of taxa held within MAST among all pair-wise comparisons ranged from a minimum of 131 (between AIC14 and MLP4) to a maximum of 176 (between BIC14 and BIC40) out of a possible 185. It was not possible to calculate the *l*cong statistic for some pair-wise comparisons, in particular the MLP4 tree generated from ML analyses with 40 partitions was not comparable to most of the remaining trees. The *l*cong could only be calculated for the MLP40 and the AIC14 and the MLP14, and as mentioned above the MLP40 and AIC14 MAST had the lowest number of shared taxa (Table 11).

Phylogenetic Informativeness: The analyses of PI revealed that both model selection and partitioning scheme affected the informativeness of the characters (Table 12). For most of the markers examined, PI was highest in the Bayesian analyses generated from models selected by the DT selection criterion and partitioned by gene (Table 12). As mentioned above, the DT -selected models were simpler than those selected by AIC and BIC selection criteria in most instances. In the ML analyses, PI was highest for the analyses where fourteen partitions (e.g. partitioned by gene) were used as opposed to the analyses that were more or less partitioned (Figures 2, 3). The PI was lowest for the unpartitioned analyses (Figures 2,3). Bayesian analyses performed under the GTR model partitioning by gene gave higher PI values than the other partitioning schemes (Figures 4, 5), as with the ML analyses, the PI was lowest under the unpartitioned analyses (Figures 4, 5). Conversely, both the AICc and BIC models had higher PI under the codon partitioned than gene partitioned analyses (Figures 6,7).

Since the Bayes Factors analyses support the concatenation tree generated under the BIC40 model, and the PI is highest for most of the markers under the DT14 concatenated tree, both of these trees could potentially serve as the optimal framework within which to evaluate the evolutionary relationships of the Dicroglossidae. While these two trees are largely congruent, the MAST indicates complete concordance of 163 taxa out of 185 possible taxa, which means that there are some topological variations between these two trees (Table 11). Furthermore, the likelihoods of these trees differ significantly as do the tree heights, although the remaining tree statistics are similar (Table 10). We evaluate both of these trees.

Dicroglossid Frogs: The DT14 tree (Figure 9) was well resolved with strong posterior probabilities at most major nodes. Among the African ranoids, multiple non-monophyletic families were recovered, in particular the Pyxicephalidae was scattered among other groups. The Dicroglossidae were recovered with strong support values at the base of the clade. Sister to the Dicrogossidae is a poorly supported polytomy containing

Conraua (Petropedetidae), *Pyxicephalus* (Pyxicephalidae) and the *Ranidae*. Within the Dicroglossidae both the Occidozyginae and Dicroglossinae subfamilies were recovered. The subfamily Occidozyginae contained *Ingerana* and *Occidozyga*. Within the Dicroglossinae, *Hoplobatrachus*, *Euphlyctis, Limnonectes, and Ombrana* were all monophyletic, however three clades of *Fejervarya* were recovered, as well as a non-monophyletic *Nanorana*. One of the *Fejervarya* clades is sister to a clade containing *Euphylctis, Hoplobatachus, Sphaerotheca*, and all other *Fejervarya*. It consists of a single sample from western Myanmar, additional analyses and sampling is required to determine the status of this lineage. Of the remaining two *Fejervarya* clades, one clade is primarily composed of Indian species and is sister to the genus *Sphaerotheca*. This South Asian clade contains *Fejervarya andamanensis, F. mudduraja*, *F. rufescens, F. kudremukhensis, F. caperata, F. syhadrensis, F. greenei, F. kirtsinghhei, F. pierrei*, and *F. granosa*. *F. syhadrensis* is recovered as paraphyletic. In addition to the named species, cryptic lineages from Myanmar (sister to the Andaman Island species) and India and Myanmar were found. The second clade, the Asian clade includes species from east Asia (Japan, Taiwan, and China), insular Asia and Sundaland, and mainland Southeast Asia (Indochina and Myanmar). Species within this clade include the crab-eating frog *F. cancrivora*, *F. triora* from Thailand, *F. iskandari* from Java, *F. multistriata* from China, *F. sakishimensis* from Japan, and two undescribed lineages from Myanmar.

The BIC40 tree (Figure 10) recovered *Conraua* as sister to all other Dicroglossidae. The Occidozyginae and Dicroglossinae were both recovered. *Occidozyga, Limnonectes, Euphlyctis, Hoplobatachus, Sphaerotheca*, and *Ombrana* were all monophyletic whereas *Fejervarya*, and *Nanorana* were not. Three clades of *Fejervarya* were recovered, matching the same clade memberships as above. The clade composed of a single specimen from western Myanmar was grouped with the Occidozyginae in this tree, suggesting that additional sampling is needed to resolve the placement of this taxon.

DISCUSSION

This study has demonstrated 1) that model selection criteria have systemic biases that affect phylogenetic inference, 2) that selected models are often not able to adequately capture the complexity of nucleotide substitution underlying empirical data, 3) that partitioning scheme affects phylogenetic inference and 4) that choice of nucleotide substitution model and partitioning scheme affects the phylogenetic informativeness of the data.

The selection criteria examined in this study found that the BIC tended to select the most parameter-rich models as compared to the AICc or DT. These results differ from those of Ripplinger and Sullivan (2010), which found that the AIC tended to select more complex models than the BIC, DT, or hLRT. More worrisome than the selection bias however is the inadequacy of most of our models to capture the evolutionary dynamics of model substitution in empirical data. The tests of model adequacy here showed that, of the 14 markers, the model was only able to characterize the underlying substitution dynamics for 9 of them, while for the remaining markers, non of the selected models were working well. One aspect of the model adequacy tests used here, the multinomial

likelihood statistic, is that missing and ambiguous sites have to be removed prior to analyses, so for many empirical datasets, particularly nuclear datasets, this will impact the calculation of the test statistic. Calculation of the T(x) statistic with the inclusion of missing and or ambiguous data is currently too computationally intensive (Brown and Eldabaje, 2009).

Relatively congruent topologies were recovered from all of our concatenation analyses, suggesting that topology is relatively robust to both nucleotide substitution model and partitioning scheme. However, the degree of congruency varied, with the number of shared taxa in MAST varying greatly. So while the topology was relatively stable, there were significant topological differences among the trees obtained from differing model selection and partitioning strategies. Enumeration of these differences using more sensitive topology tests might be important to understand under which sets of conditions topologies become unstable. The properties of the trees were less stable, in particular the estimated tree lengths and tree heights varied by orders of magnitude. The symmetry of the trees ($\overline{N}$), the external/internal branch length ratio, and the treeness (signal to noise ratios; T) were also stable across model selection, inference method and partitioning schema. This suggests that these statistics are also relatively robust to many phylogenetic analytical decisions.

Fong and Fujita (2011) examined Phylogenetic Informativeness (PI) across the vertebrate tree of life to evaluate the optimal data type at differing temporal scales. They compared PI for models utilizing nucleotides, amino acids, first and second codon positions only, and degenerative nucleotide data. They demonstrated that the use of nucleotides at all time scales was appropriate since the PI of these data were much higher than for any of the other data forms. Here I demonstrated that the selection of nucleotide substitution model and partitioning scheme also affects the PI. The PI was highest for the simpler models selected by DT and for analyses employing partitioning by gene rather than by gene and codons. PI is based upon (1) the joint probability of character state changes occurring on an internal branch, (2) the lack of character state changes along the tips, and (3) the ratio of the empirical character state change rates to the optimal character state change rates, it suggests that changes in PI could reflect change in the relative lengths of external and internal branch lengths, or differing branching orders, or overall tree length. Since the topological tests of congruence suggest that all of the trees evaluated here were highly similar, and the treeness statistic suggests that the signal to noise ratio was very little altered among analyses, and the external to internal branch length ratio was also relatively stable, that leaves the tree length as the primary explanation for differences in PI. It might be beneficial to further evaluate the topological differences on a node-by-node basis to provide more insight into what features of the different trees are changing the PI.

The Dicroglossidae are recovered in both analyses presented here. In the DT14 analysis, the family is sister to a polytomy of frogs from three families (Ranidae, Pyxicephalidae, and Petropedatidae (*Conraua*), whereas in the BIC40 analysis they are sister to *Conraua* alone. The placement of *Conraua* is interesting as it has been a taxonomic puzzle with morphological analyses placing it within the Dicrogossidae

(Dubois, 1992), and molecular analyses placing it sister to South African ranid families (Van Der Meijden et al., 2005), within the Petropedetidae (Frost et al., 2006), or sister to *Limnonectes* within Dicroglossidae (Kosuch et al., 2001).

Our analyses found support for the widely recognized subfamilies, Occidozyginae and the Dicroglossinae. While many of the genera within the Dicroglossinae were monophyletic such as *Euphlyctis, Hoplobatachus*, and *Limnonectes* other recognized taxa such as *Fejervarya, Nanorana*, and *Quasipaa* were not recovered as monophyletic, indicating that they still require additional taxonomic attention. Che et al (2009; 2010) have been examining the frogs formerly of the genus *Paa* and have been working to resolve the generic placement of the *Quasipaa*, *Allopaa, Chrysopaa* and *Nanorana*. Our findings provide further evidence that the taxonomies of these groups require further investigation. Finally, our study suggests that *Fejervarya* should either be split into two genera, or that *Sphaerotheca* should be synonomyzed with *Fejervarya*. Furthermore, the taxonomic status of the third Fejervarya linage (from western Myanmar) requires attention.

Literature Cited

Bollback, J. P. 2002. Bayesian model adequacy and choice in phylogenetics. Molecular Biology and Evolution 19:1171-1180.

Bossuyt, F., R. M. Brown, D. M. Hillis, D. Cannatella, and M. Milinkovitch. 2006. Phylogeny and Biogeography of a Cosmopolitan Frog Radiation: Late Cretaceous Diversification Resulted in Continent-Scale Endemism in the Family Ranidae. Syst. Biol. 55:579-594.

Brandley, M., A. Schmitz, and T. Reeder. 2005. Partitioned Bayesian analyses, partition choice, and the phylogenetic relationships of scincid lizards. Syst. Biol. 54:373-90.

Brown, J., and A. Lemmon. 2007. The importance of data partitioning and the utility of bayes factors in Bayesian phylogenetics. Syst. Biol. 56:643-55.

Brown, J. M., and R. ElDabaje. 2009. Applications Note PuMA: Bayesian analysis of partitioned (and unpartitioned) model adequacy. Bioinformatics 25:537-538.

Brown, J. M., and R. Eldabaje. 2009. PuMA: Bayesian analysis of partitioned (and unpartitioned) model adequacy. Bioinformatics 25:537-538.

Che, J., J. Hu, W. Zhou, R. Murphy, T. Papenfuss, M. Chen, D. Rao, P. Li, and Y. Zhang. 2009. Phylogeny of the Asian spiny frog tribe Paini (Family Dicroglossidae) sensu Dubois. Molecular Phylogenetics and Evolution 50:59-73.

Che, J., J. Pang, H. Zhao, G. Wu, E. Zhao, and Y. Zhang. 2007a. Molecular phylogeny of the Chinese ranids inferred from nuclear and mitochondrial DNA sequences. Biochemical Systematics and Ecology 35:29-39.

Che, J., J. Pang, H. Zhao, G. Wu, E. Zhao, and Y. Zhang. 2007b. Phylogeny of Raninae (Anura: Ranidae) inferred from mitochondrial and nuclear sequences. Mol. Phylogenet. Evol. 43:1-13.

Che, J., W. Zhou, J. Hu, F. Yan, T. J. Papenfuss, D. Wake, and Y. Zhang. 2010. Spiny frogs (Paini) illuminate the history of the Himalayan region and Southeast Asia. Proceedings of the National Academy of Sciences 107:13765.

Chen, L., R. Murphy, A. Lathrop, A. Ngo, and N. Orlov. 2005. Taxonomic chaos in Asian ranid frogs: an initial phylogenetic resolution. Herpetological Journal.

de Vienne, D. M., T. Giraud, and O. C. Martin. 2007. A congruence index for testing topological similarity between trees. Bioinformatics 23:3119-3124.

Drummond, A., B. Ashton, S. Buxton, M. Cheung, A. Cooper, J. Heled, M. Kearse, R. Moir, S. Stones-Havas, S. Sturrock, T. Thierer, and A. Wilson. 2010. Geneious v5.1, version 5.3.6. Available from http://www.geneious.com.

Dubois, A. 1992. Notes sur la classification des Ranidae (amphibiens anoures). Bulletin Mensuel de la Societe Linneenne de Lyon 61:305-352.

Dubois, A., and A. Ohler. 2005. Taxonomic notes on the Asian frogs of the tribe Paini (Ranidae, Dicroglossinae): 1. Morphology and synonymy of Chaparana aenea (Smith, 1922), with proposal of a new statistical method for testing homogeneity of small samples. Journal of Natural History 39:1759-1778.

Edgar, R. C. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Research 32:1792-1797.

Fong, J., and M. Fujita. 2011. Evaluating phylogenetic informativenss and data-type usage for new protein-coding genes across Vertebrata. Molecular Phylogenetics and Evolution 61:300-307.

Frost, D. R., T. Grant, J. Faivovich, R. Bain, A. Haas, C. Haddad, R. De Sa, A. Channing, M. Wilkinson, S. C. Donnellan, C. Raxworthy, J. A. Campbell, B. L. Blotto, P. Moler, R. C. Drewes, R. A. Nussbaum, J. D. Lynch, D. M. Green, and W. C. Wheeler. 2006. The Amphibian Tree of Life. Bulletin of the American museum of Natural History 297:1-370.

Guindon, S., J. F. Dufayard, V. Lefort, M. Anisimova, W. Hordijk, and O. Gascuel. 2010. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. Systematic Biology 59:307-321.

Huelsenbeck, J. P., and Ronquist. 2002. MrBayes: Bayesian inference of phylogeny.

Kirkpatrick, M., and M. Slatkin. 1993. Searching for evolutionary patterns in the shape of a phylogenetic tree. Evolution 47:1171-1181.

Kosuch, J., M. Vences, A. Dubois, A. Ohler, and W. Bohme. 2001. Out of Asia: mitochondrial DNA evidence for an Oriental origin of tiger frogs, genus *Hoplobatrachus*.

Lemmon, A., and E. C. Moriarty. 2004. The importance of proper model assumption in Bayesian phylogenetics. Systematic Biology 53:265-277.

Lopez-Giraldez, F., and J. P. Townsend. 2011. PhyDesign: a webapp for profiling phylogenetic informativeness BMC Evolutionary Biology 11:152.

McGuire, J., C. Witt, D. Altshuler, and J. Remsen. 2007. Phylogenetic systematics and biogeography of hummingbirds: bayesian and maximum likelihood analyses of partitioned data and selection of an appropriate partitioning strategy. Syst. Biol. 56:837-56.

Ohler, A., K. Deuti, S. Grosjean, S. Paul, A. K. Ayyaswamy, M. F. Ahmed, and S. K. Dutta. 2009. Small-sized dicroglossids from India, with the description of a new species from West Bengal, India. Zootaxa:43-56.

Ohler, A., and A. Dubois. 2006. Phylogenetic relationships and generic taxonomy of the tribe Paini (Amphibia, Anura, Ranidae, Dicroglossinae) with diagnoses of two new genera. Zoosystema 28:769-784.

Phillips, M. J., Y.-H. Lin, G. L. Harrison, and D. Penny. 2001. Mitochondrial genomes of a bandicoot and a brushtail possum confirm the monophyly of Australidelphian marsupials. Proceedings of the Royal Society Biological Sciences 268:1533-1538.

Phillips, M. J., and D. Penny. 2003. The root of the mammalian tree inferred from whole mitochondrial genomes. Molecular Phylogenetics and Evolution 28:171-185.

Pond, S. L., S. D. Frost, and S. V. Muse. 2005. HyPhy: hypothesis testing using phylogenies. Bioinformatics 21:676-679.

Posada, D. 2008. jModelTest: Phylogenetic Model Averaging. Molecular Biology and Evolution 25:1253-1256.

Rambaut, A., and M. Charleston. 2001. TreeEdit http://tree.bio.ed.ac.uk/software/treeedit/.

Rambaut, A., and A. Drummond. 2011. TreeStat v.1.6.2. Tree Statistical Calculation Tool.

Rambaut, A., and N. C. Grassly. 1997. Seq-Gen: an application for the Monte carlo simulation of DNA sewuence evolution along phylogenetic trees. Computational Applied Biosciences 13:235-238.

Ripplinger, J., and J. Sullivan. 2008. Does choice in model selection affect maximum likelihood analysis? Systematic Biology 57:76-85.

Ripplinger, J., and J. Sullivan. 2010. Assessment of substitution model adequacy using frequentist and Bayesian methods. Molecular Biology and Evolution 27:2790-2803.

Scott, E. 2005. A phylogeny of ranid frogs (Anura: Ranoidea: Ranidae), based on a simultaneous analysis of morphological and molecular data. Cladistics 21:507-574.

Smith, S. A., and C. W. Dunn. 2008. Phyutility: a phyloinformatics tool for trees, alignments, and molecular data. Bioinformatics 24:715-716.

Stamatakis, A. 2004. RAxML-III: a fast program for maximum likelihood-based inference of large phylogenetic trees. Bioinformatics 21:456-463.

Stamatakis, A., P. Hoover, and J. Rougemont. 2008. A rapid bootstrap algorithm for the RAxML Web-Servers. Systematic Biology 75:758-771.

Suchard, M. A., R. E. Weiss, and J. S. Sinsheimer. 2001. Bayesian selection of continuous-time Markov chain evolutionary models. Molecular Biology and Evolution 18:1001-1013.

Swofford, D. L. 2002. PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods). version Version 4. Sinauer Associates.

Townsend, J. P. 2007. Profiling phylogenetic informativeness. Systematic Biology 56:222-231.

Van Der Meijden, A., M. Vences, S. Hoegg, and A. Meyer. 2005. A previously unrecognized radiation of ranid frogs in Southern Africa revealed by nuclear and mitochondrial DNA sequences. Molecular Phylogenetics and Evolution 37:674-685.

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

Wogan, G. O. U., H. Win, T. Thin, K. S. Lwin, A. K. Shein, S. W. Kyi, and H. Tun. 2003. A new species of *Bufo* (Anura: Bufonidae) from Myanmar (Burma), and redescription of the little known species *Bufo stuarti* Smith 1929. Proceedings of The California Academy of Sciences 54:141-153.

Figure 1. Likelihood scores from from Bayesian and ML analyses of individual genes generated under the evolutionary model of nucleotide substitution selected by AICc, BIC, DT selection criteria, and also GTR + I + Γ for both unpartitioned and codon partitioned datasets. Note that the likelihood scores from RaxML are not directly comparable to those calculated from Bayesian analyses.

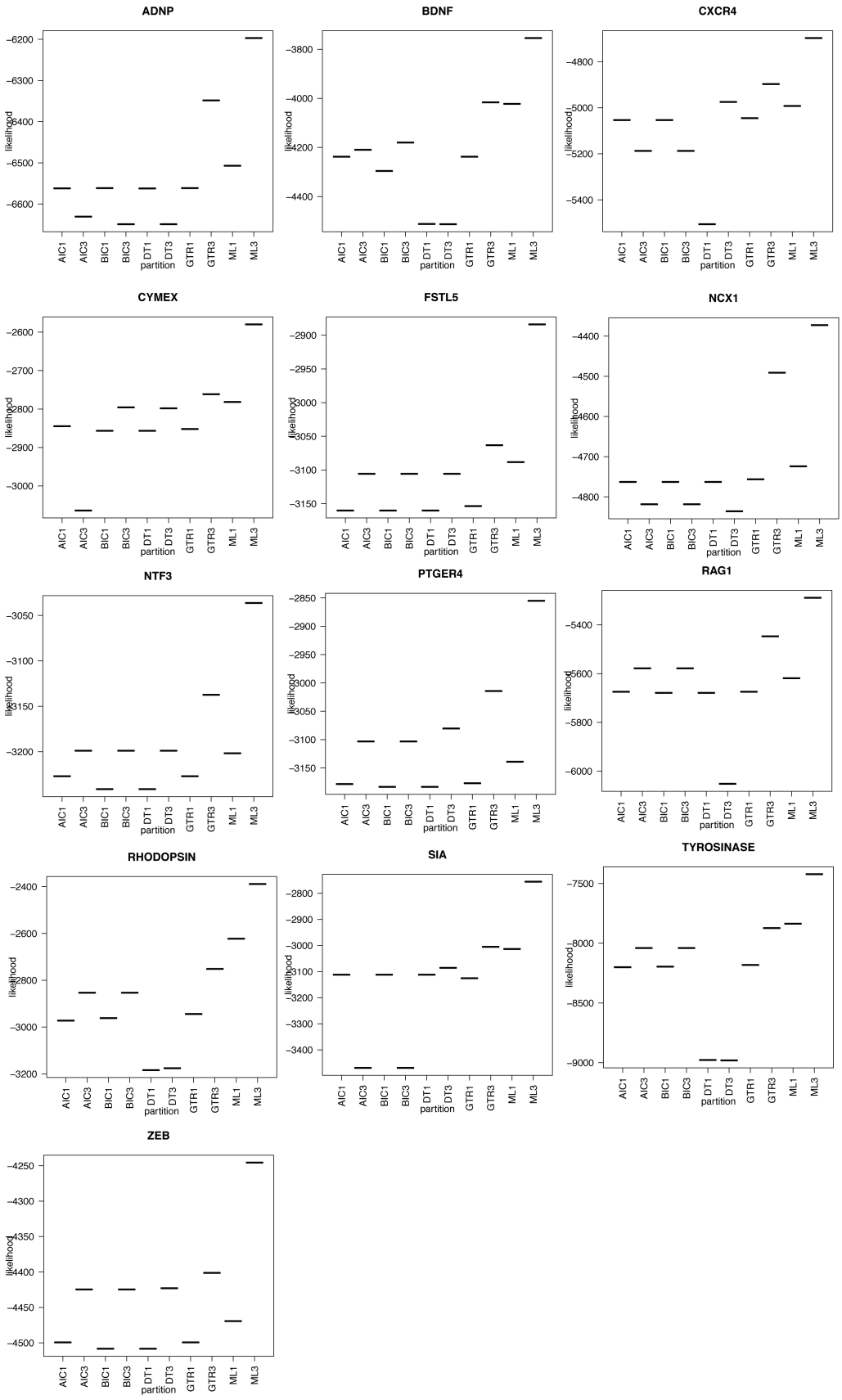
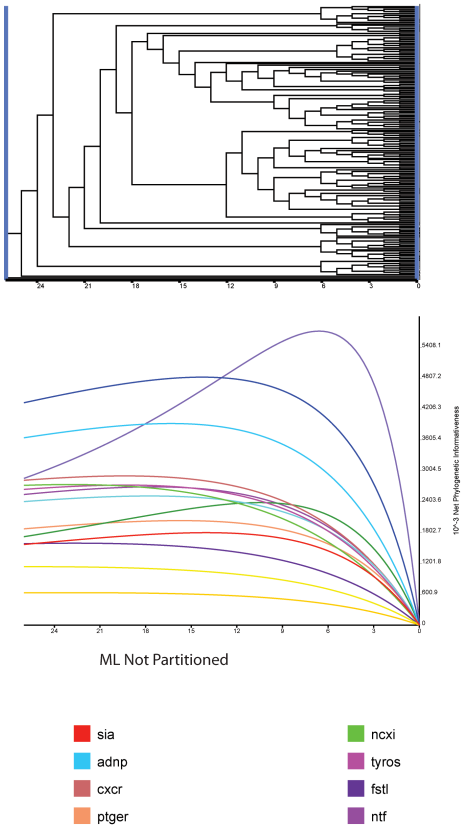
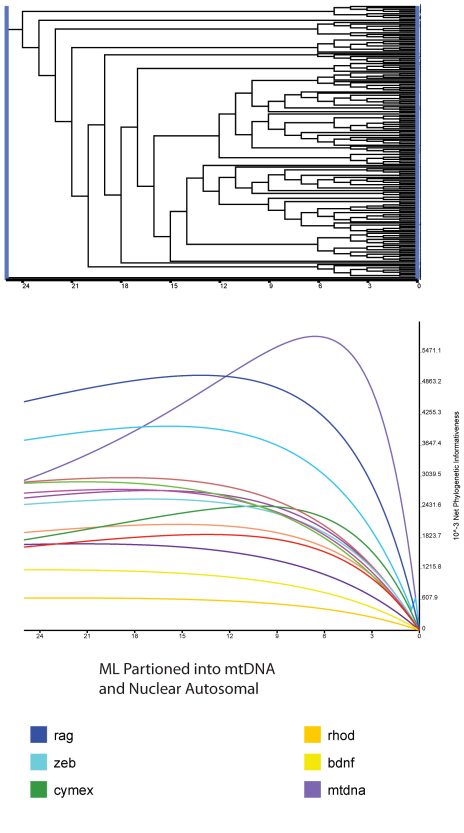

ADNP
BDNF
CXCR4
CYMEX
FSTL5
NCX1
NTF3
PTGER4
RAG1
RHODOPSIN
SIA
TYROSINASE
ZEB
likelihood
partition
AIC1
AIC3
BIC1
BIC3
DT1
DT3
GTR1
GTR3
ML1
ML3

Figure 2. Phylogenetic Informativeness (PI) of 14 markers estimated based on Maximum Likelihood trees. Two of the partitioning schemes are depicted; not partitioned (left), two partitions (right).



ML Not Partitioned



ML Partioned into mtDNA
and Nuclear Autosomal

sia
adnp
cxcr
ptger
ncxi
tyros
fstl
ntf

rag
zeb
cymex
rhod
bdnf
mtdna

Figure 3. Phylogenetic Informativeness (PI) of 14 markers estimated based on Maximum Likelihood trees. Two of the partitioning schemes are depicted; fourteen partitions (left), forty partitions (right).

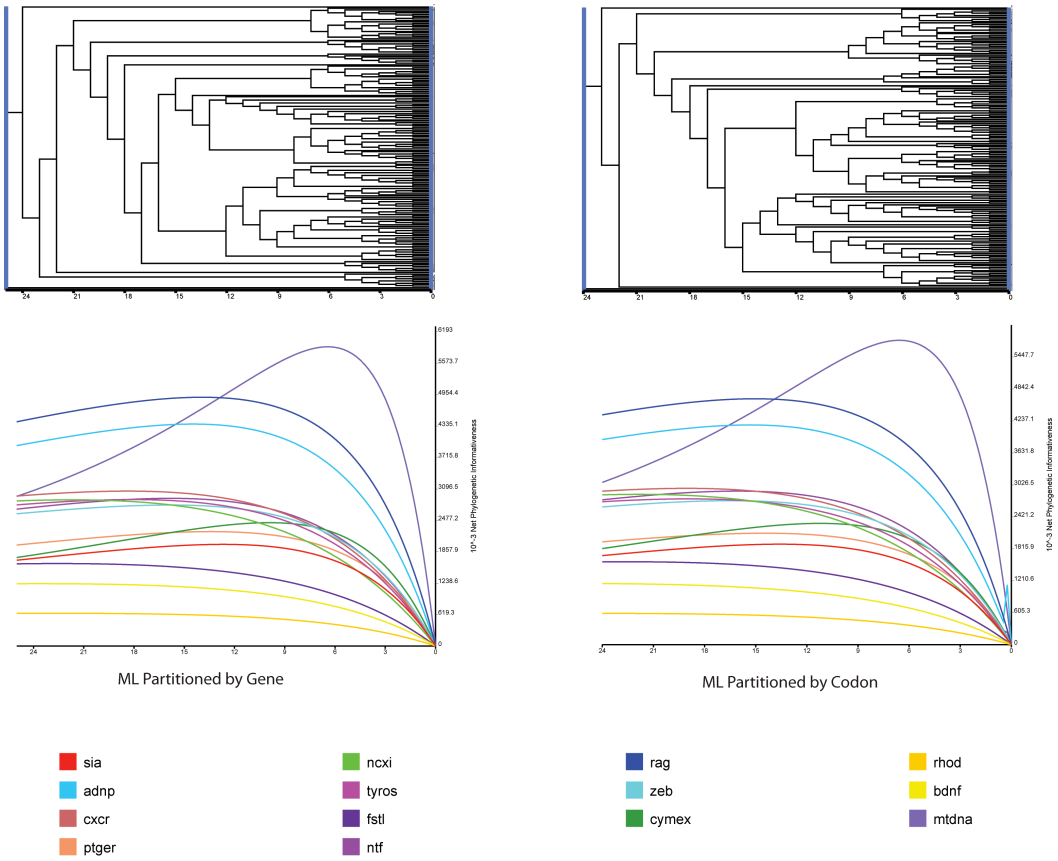
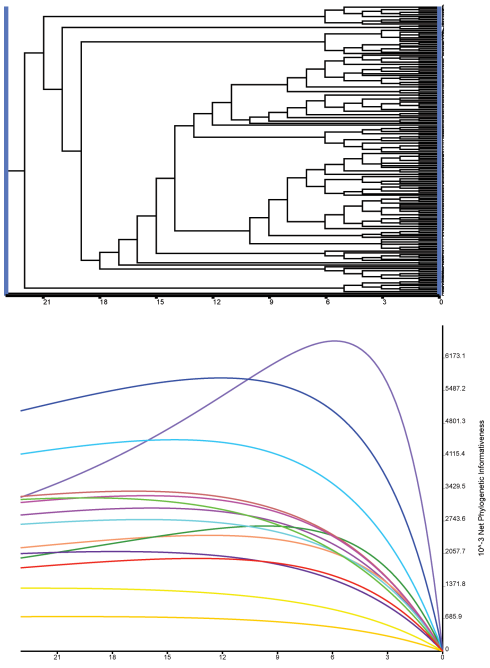
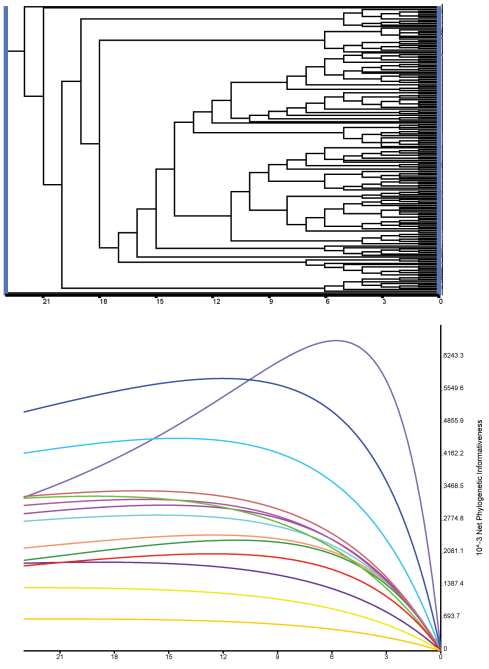

24 21 18 15 12 9 6 3 0
6193
5573.7
4954.4
4335.1
3715.8
3096.5
2477.2
1857.9
1238.8
619.3
0
10^-3 Net Phylogenetic Informativeness
5447.7
4842.4
4237.1
3631.8
3026.5
2421.2
1815.9
1210.6
605.3
ML Partitioned by Gene
ML Partitioned by Codon
sia
adnp
cxcr
ptger
ncxi
tyros
fstl
ntf
rag
zeb
cymex
rhod
bdnf
mtdna

Figure 4. Phylogenetic Informativenss (PI) of 14 markers estimated based on Bayesian Inference trees under the GTR nucleotide substitution model. Two of the partitioning schemes are depicted; not partitioned (left), two partitions (right).



GTR Not Partitioned



GTR Partitioned into mtDNA
and Nuclear Autosomal

sia
adnp
cxcr
ptger
ncxi
tyros
fstl
ntf
rag
zeb
cymex
rhod
bdnf
mtdna

Figure 5. Phylogenetic Informativenss (PI) of 14 markers estimated based on Bayesian Inference trees under the GTR nucleotide substitution model. Two of the partitioning schemes are depicted; fourteen partitions (left), forty partitions (right).

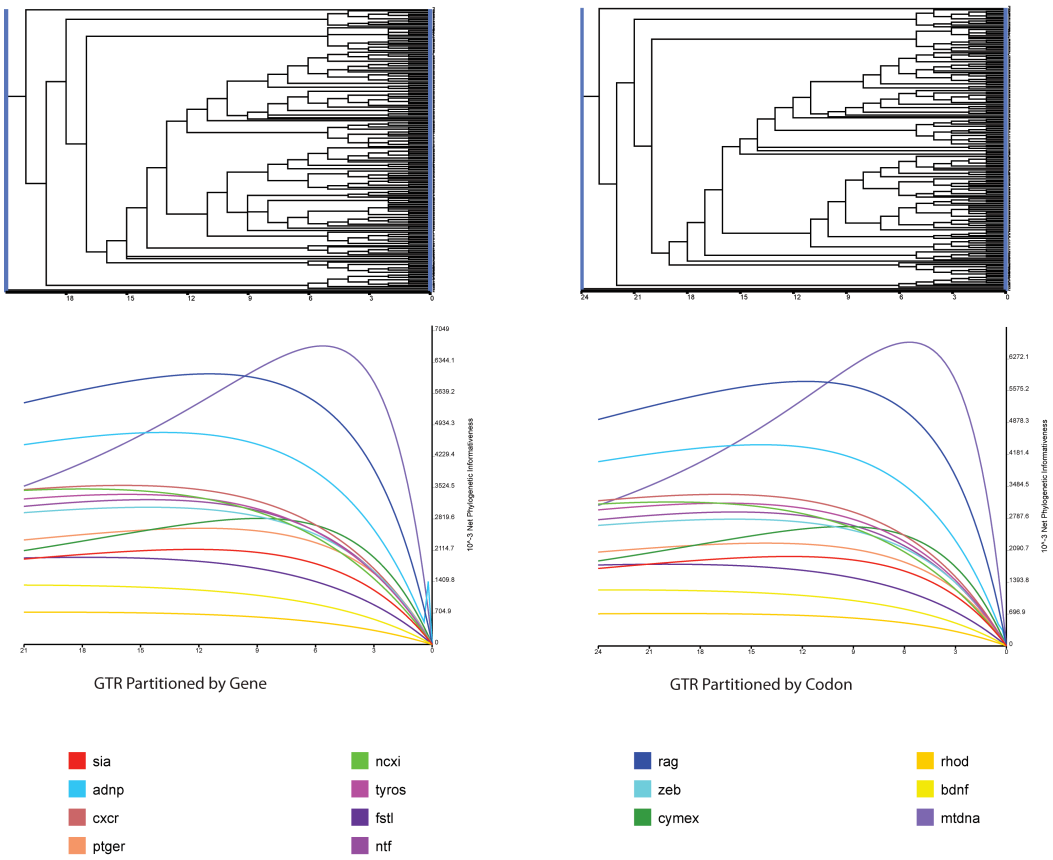
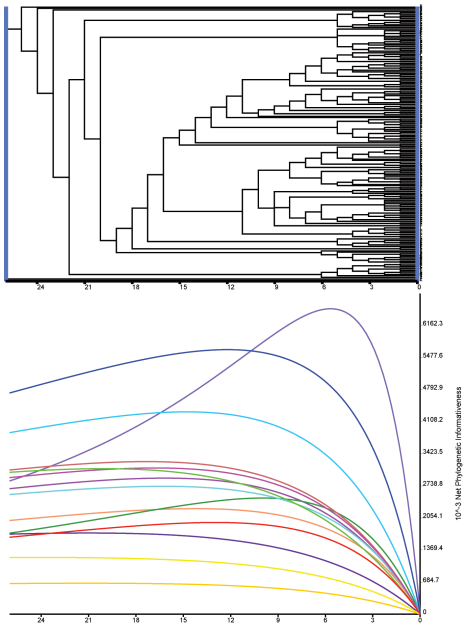
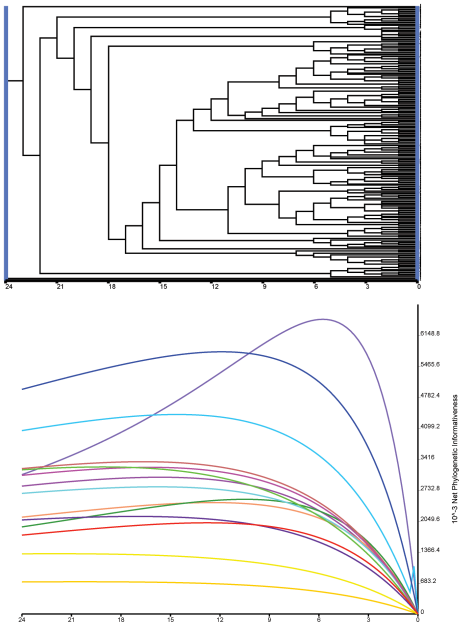

GTR Partitioned by Gene
GTR Partitioned by Codon
sia
adnp
cxcr
ptger
ncxi
tyros
fstl
ntf
rag
zeb
cymex
rhod
bdnf
mtdna

Figure 6. Phylogenetic Informativenss (PI) of 14 markers estimated based on Bayesian Inference trees under the nucleotide substitution models chosen by AICc. Two of the partitioning schemes are depicted; fourteen partitions (left), forty partitions (right).



AICc Partitioned by Gene



AICc Partitioned by Codon

sia
adnp
cxcr
ptger
ncxi
tyros
fstl
ntf
rag
zeb
cymex
rhod
bdnf
mtdna

Figure 7. Phylogenetic Informativenss (PI) of 14 markers estimated based on Bayesian Inference trees under the nucleotide substitution models chosen by BIC. Two of the partitioning schemes are depicted; fourteen partitions (left), forty partitions (right).

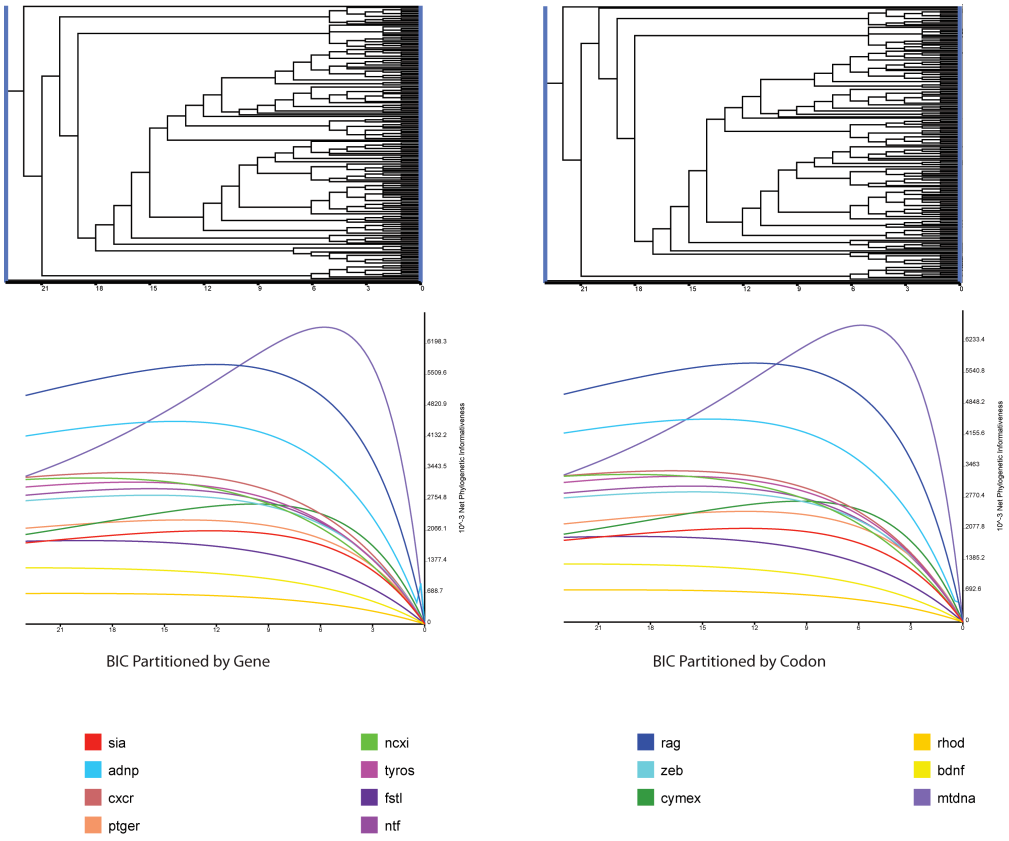
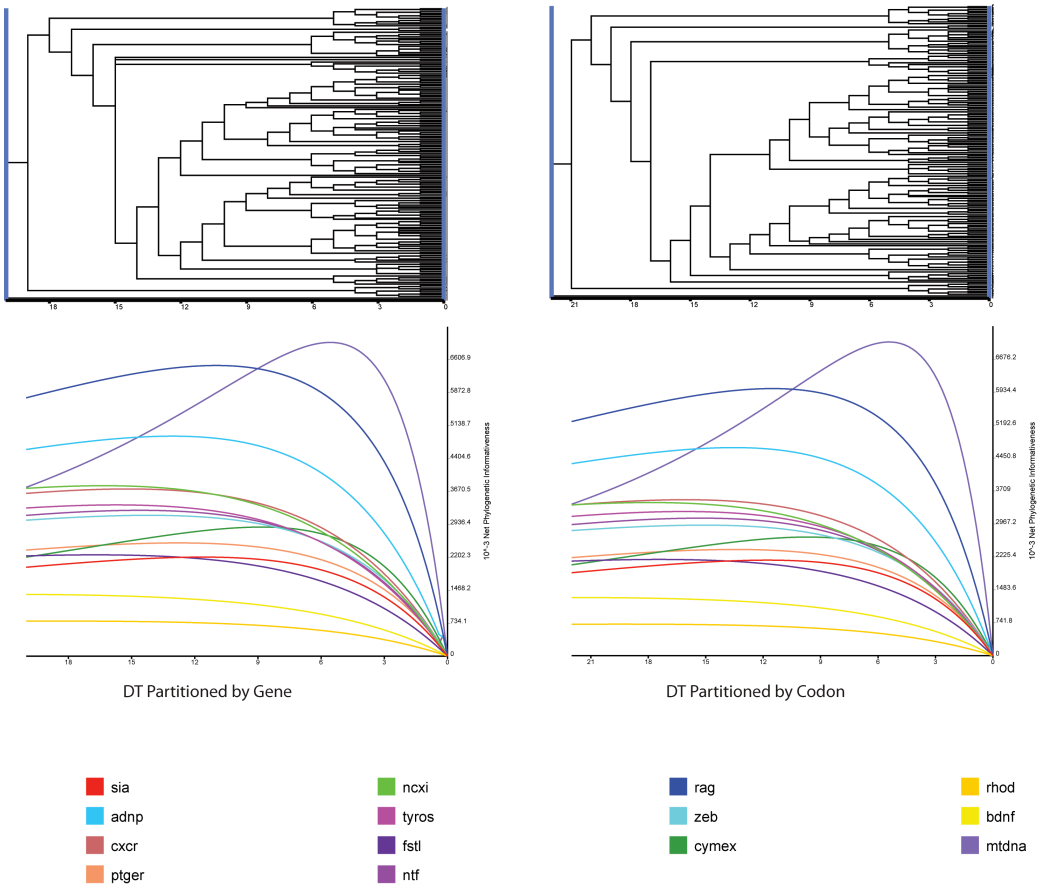

BIC Partitioned by Gene
BIC Partitioned by Codon
sia
adnp
cxcr
ptger
ncxi
tyros
fstl
ntf
rag
zeb
cymex
rhod
bdnf
mtdna

Figure 8. Phylogenetic Informativenss (PI) of 14 markers estimated based on Bayesian Inference trees under the nucleotide substitution models chosen by DT. Two of the partitioning schemes are depicted; fourteen partitions (left), forty partitions (right).

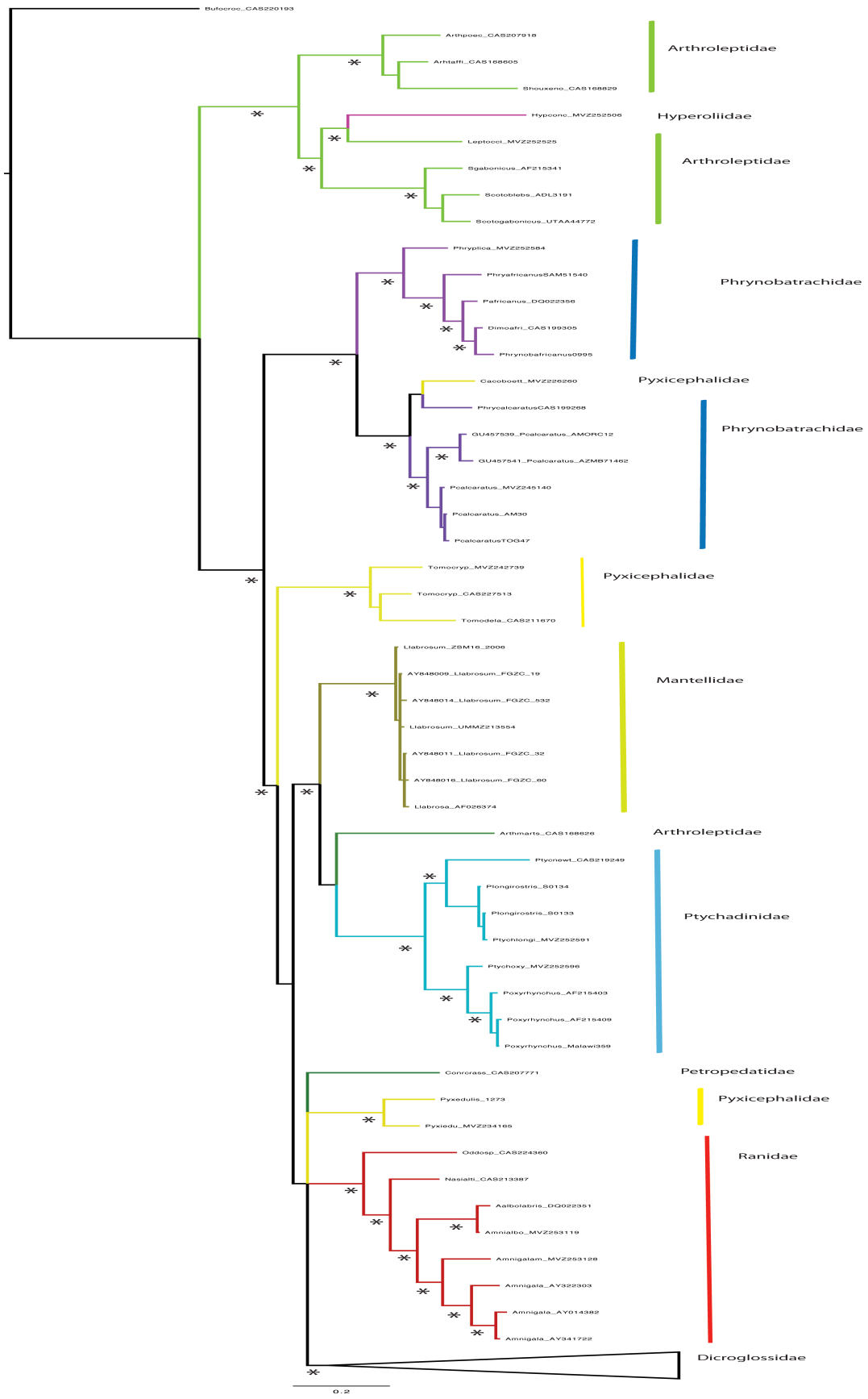


DT Partitioned by Gene

DT Partitioned by Codon

sia
adnp
cxcr
ptger
ncxi
tyros
fstl
ntf
rag
zeb
cymex
rhod
bdnf
mtdna

Figure 9. Bayesian Inference phylogenetic tree from concatenated analyses generated under the nucleotide substitution models selected by the DT with 14 partitions. Posterior probabilities >.95 are demarcated with an *. First page depicting familial relationships, the second page focused in upon Dicroglossidae (indicated as a collapsed clade in the familial tree).


Bufocroc_CAS220193
Arthpoec_CAS207918
Arhtaffi_CAS168605
Shouxeno_CAS168829
Arthroleptidae
Hypconc_MVZ252506
Hyperoliidae
Leptocci_MVZ252525
Sgabonicus_AF215341
Scotoblebs_ADL3191
Scotogabonicus_UTAA44772
Arthroleptidae
Phryplica_MVZ252584
PhryafricanusSAM51540
Pafricanus_DQ022356
Dimoafri_CAS199305
Phrynobafricanus0995
Phrynobatrachidae
Cacoboett_MVZ226260
PhrycalcaratusCAS199288
GU457539_Pcalcaratus_AMORC12
GU457541_Pcalcaratus_AZMB71462
Pcalcaratus_MVZ245140
Pcalcaratus_AM30
PcalcaratusTOG47
Pyxicephalidae
Phrynobatrachidae
Tomocryp_MVZ242739
Tomocryp_CAS227513
Tomodela_CAS211670
Pyxicephalidae
Liabrosum_ZSM16_2006
AY848009_Liabrosum_FGZC_19
AY848014_Liabrosum_FGZC_532
Liabrosum_UMMZ213554
AY848011_Liabrosum_FGZC_32
AY848016_Liabrosum_FGZC_60
Liabrosa_AF026374
Mantellidae
Arthmarts_CAS168626
Arthroleptidae
Ptyonewt_CAS219249
Plongirostris_80134
Plongirostris_80133
Ptychlongi_MVZ252591
Ptychoxy_MVZ252596
Poxyrhynchus_AF215403
Poxyrhynchus_AF215409
Poxyrhynchus_Malawi359
Ptychadinidae
Conrcrass_CAS207771
Petropedatidae
Pyxedulis_1273
Pyxiedu_MVZ234165
Pyxicephalidae
Oddosp_CAS224360
Nasialti_CAS213387
Aalbolabris_DQ022351
Amnialbo_MVZ253119
Amnigalam_MVZ253128
Amnigala_AY322303
Amnigala_AY014382
Amnigala_AY341722
Ranidae
Dicroglossidae
0.2

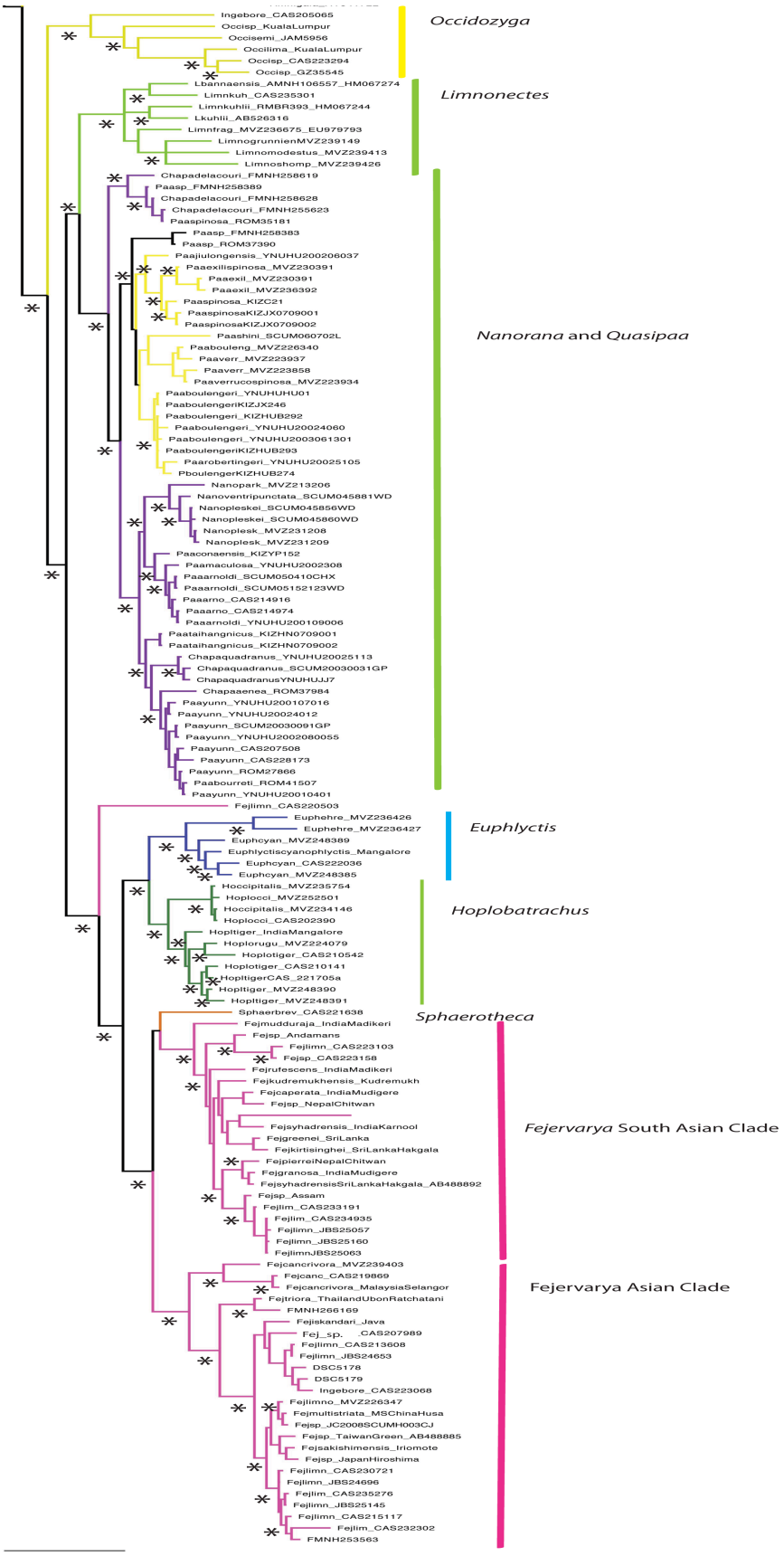
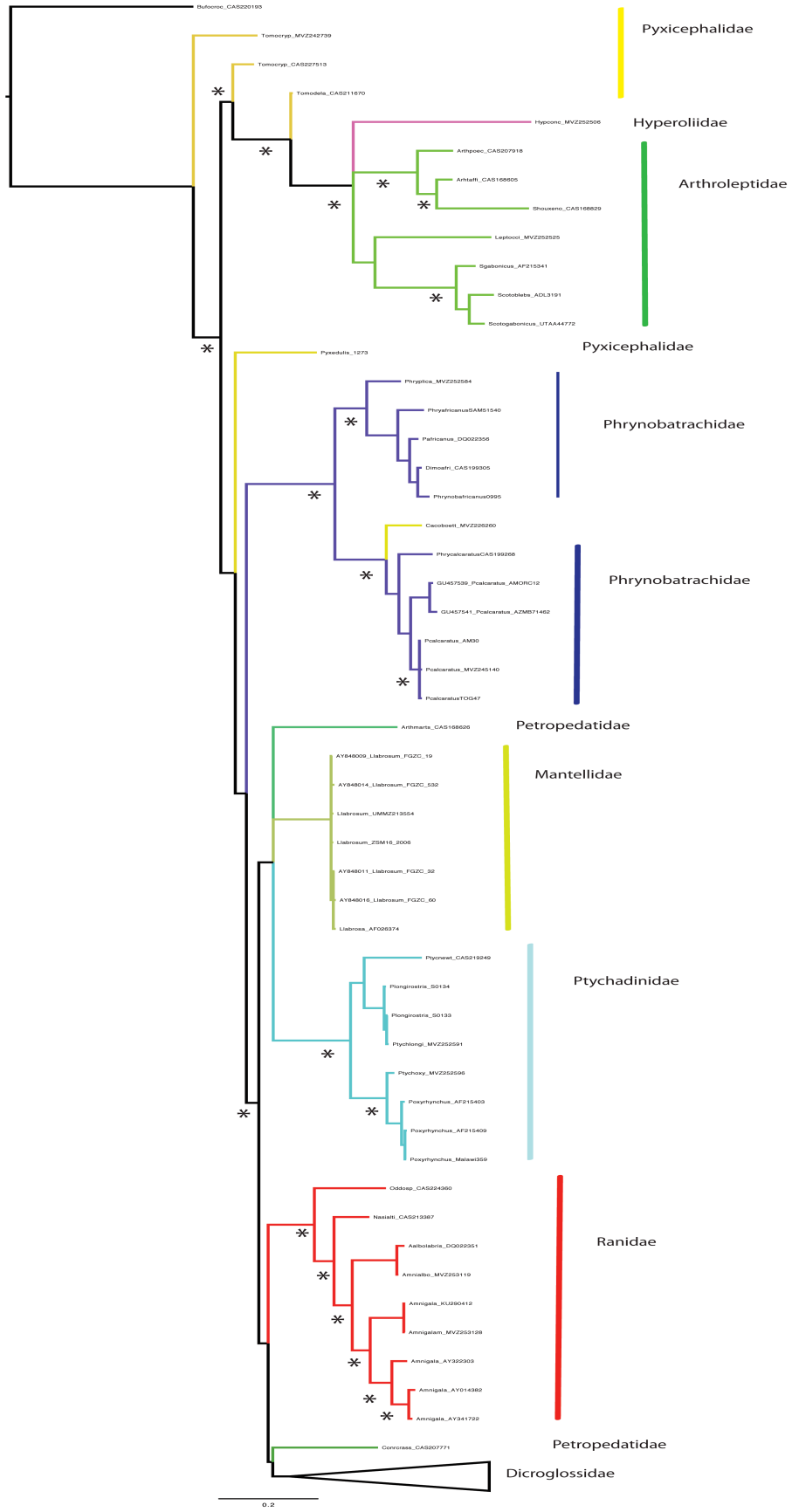

Occidozyga
Limnonectes
Nanorana and Quasipaa
Euphlyctis
Hoplobatrachus
Sphaerotheca
Fejervarya South Asian Clade
Fejervarya Asian Clade
0.2

Figure 10. Bayesian Inference phylogenetic tree from concatenated analyses generated under the nucleotide substitution models selected by the BIC with 40 partitions. Posterior probabilities >.95 are demarcated with an *. First page depicting familial relationships, the second page focused in upon Dicroglossidae (indicated as a collapsed clade in the familial tree).


Bufocroc_CAS220193
Tomocryp_MVZ242739
Tomocryp_CAS227513
Tomodela_CAS211670
Hypoconc_MVZ252506
Arthpoec_CAS207918
Arthtafl_CAS168605
Shouxeno_CAS168629
Leptocci_MVZ252525
Sgabonicus_AF215341
Scotoblabs_ADL3191
Scotogabonicus_UTAA44772
Pyxicephalidae
Hyperoliidae
Arthroleptidae
Pyxicephalidae
Pyxedulis_1273
Phryplica_MVZ252584
PhryafricanusSAM51540
Pafricanus_DQ022356
Dimoatri_CAS199305
PhrynobafricanusQ995
Cacobeett_MVZ226260
PhrycalcaratusCAS199266
GU457539_Pcalcaratus_AMORC12
GU457541_Pcalcaratus_AZM571462
Pcalcaratus_AM30
Pcalcaratus_MVZ245140
PcalcaratusTOG47
Phrynobatrachidae
Phrynobatrachidae
Arthmarts_CAS168626
AY848009_Llabrosum_FGZC_19
AY848014_Llabrosum_FGZC_532
Llabrosum_UMMZ213554
Llabrosum_ZSM16_2006
AY848011_Llabrosum_FGZC_32
AY848016_Llabrosum_FGZC_60
Llabrosa_AF026374
Petropedatidae
Mantellidae
Ptycnewt_CAS219249
Plongirostris_S0134
Plongirostris_S0133
Ptychlongi_MVZ252591
Ptychoxy_MVZ252596
Poxyrhynchus_AF215403
Poxyrhynchus_AF215409
Poxyrhynchus_Malawi359
Ptychadinidae
Oddosp_CAS224360
Nasialli_CAS213387
Aalbolabris_DQ022351
Amnialbo_MVZ253119
Amnigala_KU290412
Amnigalam_MVZ253128
Amnigala_AY322303
Amnigala_AY014382
Amnigala_AY341722
Ranidae
Conrcrass_CAS207771
Petropedatidae
Dicroglossidae
0.2

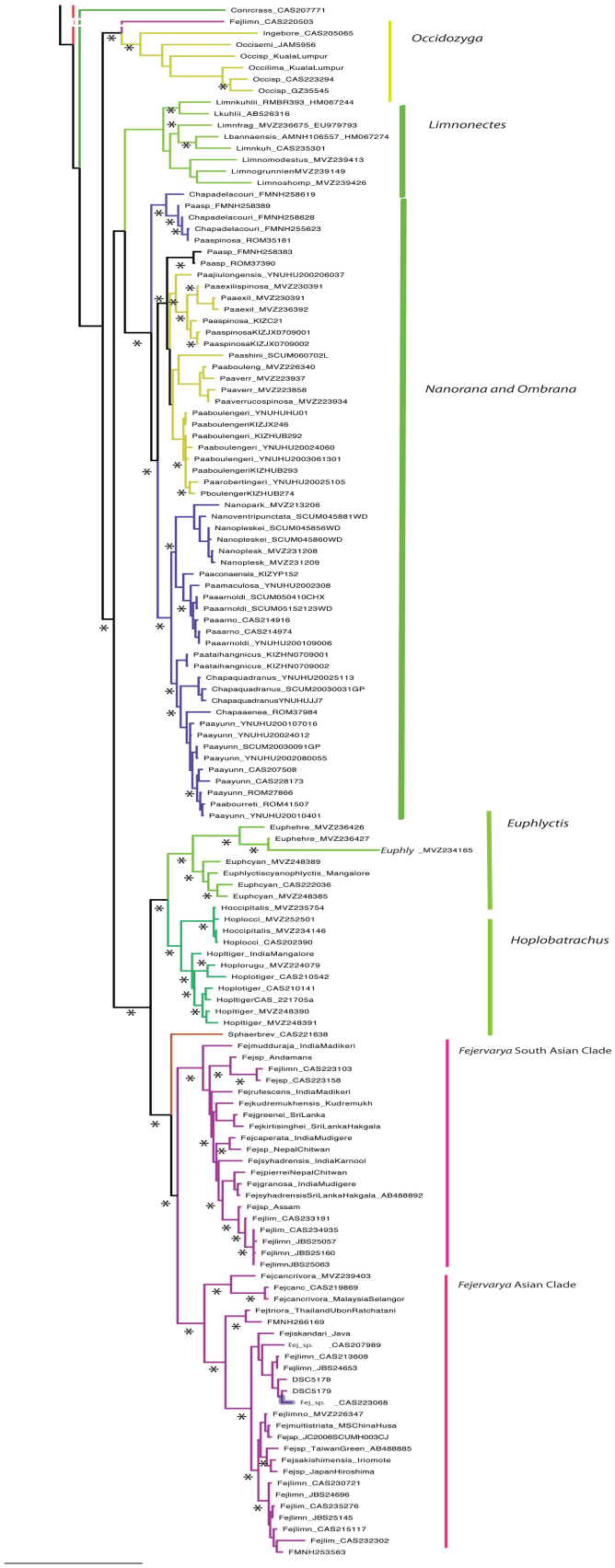

Conrcrass_CAS207771
Fejlimn_CAS220503
Ingebore_CAS205065
Occisemi_JAM5956
Occisp_KualaLumpur
Occilima_KualaLumpur
Occisp_CAS223294
Occisp_GZ35545
Occidozyga
Limnkuhlii_RMBR393_HM067244
Lkuhlii_AB526316
Limnfrag_MVZ236675_EU979793
Lbannaensis_AMNH106557_HM067274
Limnkuh_CAS235301
Limnomodestus_MVZ239413
LimnogrunnienMVZ239149
Limnoshomp_MVZ239426
Limnonectes
Chapadelacouri_FMNH258619
Paasp_FMNH258389
Chapadelacouri_FMNH258628
Chapadelacouri_FMNH255623
Paaspinosa_ROM35181
Paasp_FMNH258383
Paasp_ROM37390
Paajiulongensis_YNUHU200206037
Paaexilispinosa_MVZ230391
Paaexil_MVZ230391
Paaexil_MVZ236392
Paaspinosa_KIZC21
PaaspinosaKIZJX0709001
PaaspinosaKIZJX0709002
Paashini_SCUM060702L
Paabouleng_MVZ226340
Paaverr_MVZ223937
Paaverr_MVZ223858
Paaverrucospinosa_MVZ223934
Paaboulengeri_YNUHUHU01
PaaboulengeriKIZJX246
Paaboulengeri_KIZHUB292
Paaboulengeri_YNUHU20024060
Paaboulengeri_YNUHU2003061301
PaaboulengeriKIZHUB293
Paarobertingeri_YNUHU20025105
PboulengerKIZHUB274
Nanopark_MVZ213206
Nanoventripunctata_SCUM045881WD
Nanopleskei_SCUM045856WD
Nanopleskei_SCUM045860WD
Nanoplesk_MVZ231208
Nanoplesk_MVZ231209
Paaconaensis_KIZYP152
Paamaculosa_YNUHU2002308
Paaarnoldi_SCUM050410CHX
Paaarnoldi_SCUM05152123WD
Paaarno_CAS214916
Paaarno_CAS214974
Paaarnoldi_YNUHU200109006
Paataihangnicus_KIZHN0709001
Paataihangnicus_KIZHN0709002
Chapaquadranus_YNUHU20025113
Chapaquadranus_SCUM20030031GP
ChapaquadranusYNUHUJ7
Chapaaenea_ROM37984
Paayunn_YNUHU200107016
Paayunn_YNUHU20024012
Paayunn_SCUM20030091GP
Paayunn_YNUHU2002080055
Paayunn_CAS207508
Paayunn_CAS228173
Paayunn_ROM27866
Paabourreti_ROM41507
Paayunn_YNUHU20010401
Nanoparks and Ombrana
Euphehre_MVZ236426
Euphehre_MVZ236427
Euphly _MVZ234165
Euphcyan_MVZ248389
Euphlyctiscyanophlyctis_Mangalore
Euphcyan_CAS222036
Euphcyan_MVZ248385
Euphlyctis
Hoccipitalis_MVZ235754
Hoplocci_MVZ252501
Hoccipitalis_MVZ234146
Hoplocci_CAS202390
Hoptiger_IndiaMangalore
Hoplorugu_MVZ224079
Hoplotiger_CAS210542
Hoplotiger_CAS210141
HoptigerCAS_221705a
Hoptiger_MVZ248390
Hoptiger_MVZ248391
Hoplobatrachus
Sphaerbrev_CAS221638
Fejmudduraja_IndiaMadikeri
Fejsp_Andamans
Fejlimn_CAS223103
Fejsp_CAS223158
Fejrufescens_IndiaMadikeri
Fejkudremukhensis_Kudremukh
Fejgreenei_SriLanka
Fejkirtisinghei_SriLankaHakgala
Fejcaperata_IndiaMudigere
Fejsp_NepalChitwan
Fejsyhadrensis_IndiaKarnool
FejpierreiNepalChitwan
Fejgranosa_IndiaMudigere
FejsyhadrensisSriLankaHakgala_AB488892
Fejsp_Assam
Fejlim_CAS233191
Fejlim_CAS234935
Fejlimn_JBS25057
Fejlimn_JBS25160
FejlimnJBS25063
Fejervarya South Asian Clade
Fejcancrivora_MVZ239403
Fejcanc_CAS219869
Fejcancrivora_MalaysiaSelangor
Fejtriora_ThailandUbonRatchatani
FMNH266169
Fejiskandari_Java
Fej_sp._CAS207989
Fejlimn_CAS213608
Fejlimn_JBS24653
DSC5178
DSC5179
Fej_sp._CAS223068
Fejlimno_MVZ226347
Fejmultistriata_MSChinaHusa
Fejsp_JC2008SCUMH003CJ
Fejsp_TaiwanGreen_AB488885
Fejsakishimensis_Iriomote
Fejsp_JapanHiroshima
Fejlimn_CAS230721
Fejlimn_JBS24696
Fejlim_CAS235276
Fejlimn_JBS25145
Fejlimn_CAS215117
Fejlim_CAS232302
FMNH253563
Fejervarya Asian Clade
0.2

Table 1. Two differing taxonomic arrangements for frogs of the family Dicroglossidae, and number of individuals sampled in this study.

Species	Num. Indivs.	AmphibiaWeb	Frost Taxonomy
		Family: Subfamily	Family: Subfamily
Amnirana albolabris	2	Ranidae: Raninae	Ranidae
Amnirana galamensis	9	Ranidae: Raninae	Ranidae
Arthroleptides martiensseni	2	Ranidae: Petropedetinae	Petropedetidae
Arthroleptis affinis	1	Arthroleptidae: Arthroleptinae	Arthroleptidae: Arthroleptinae
Arthroleptis poecilonotus	1	Arthroleptidae: Arthroleptinae	Arthroleptidae: Arthroleptinae
Bufo crocus	2	Bufonidae	Bufonidae
Cacosternum boettgeri	5	Ranidae: Pyxicephalidae	Pyxicephalidae: Cacosterninae
Chaparana aenea	1	Ranidae: Dicroglossinae	Dicroglossidae: Dicroglossinae
Chaparana delacouri	3	Ranidae: Dicroglossinae	Dicroglossidae: Dicroglossinae
Chaparana quadranus	5	Ranidae: Dicroglossinae	Dicroglossidae: Dicroglossinae
Chaparana unculuanus	1	Ranidae: Dicroglossinae	Dicroglossidae: Dicroglossinae
Conraua crassipes	3	Ranidae: Conrauinae	Petropedetidae
Euphlyctis cyanophlyctis	4	Ranidae: Dicroglossinae	Dicroglossidae: Dicroglossinae
Euphlyctis ehrenbergii	3	Ranidae: Dicroglossinae	Dicroglossidae: Dicroglossinae
Fejervarya cancrivora	5	Ranidae: Dicroglossinae	Dicroglossidae: Dicroglossinae
Fejervarya caperata	1	Ranidae: Dicroglossinae	Dicroglossidae: Dicroglossinae
Fejervarya granosa	1	Ranidae: Dicroglossinae	Dicroglossidae: Dicroglossinae
Fejervarya greenei	1	Ranidae: Dicroglossinae	Dicroglossidae: Dicroglossinae
Fejervarya iskandari	2	Ranidae: Dicroglossinae	Dicroglossidae: Dicroglossinae
Fejervarya kirtsinghei	1	Ranidae: Dicroglossinae	Dicroglossidae: Dicroglossinae
Fejervarya kudremukhensis	1	Ranidae: Dicroglossinae	Dicroglossidae: Dicroglossinae
Fejervarya limnocharis	27	Ranidae: Dicroglossinae	Dicroglossidae: Dicroglossinae
Fejervarya mudduraja	1	Ranidae: Dicroglossinae	Dicroglossidae: Dicroglossinae
Fejervarya multistriata	3	Ranidae: Dicroglossinae	Dicroglossidae: Dicroglossinae
Fejervarya orissaensis	1	Ranidae: Dicroglossinae	Dicroglossidae: Dicroglossinae
Fejervarya pierrei	1	Ranidae: Dicroglossinae	Dicroglossidae: Dicroglossinae
Fejervarya rufescens	1	Ranidae: Dicroglossinae	Dicroglossidae: Dicroglossinae
Fejervarya sakishimensis	1	Ranidae: Dicroglossinae	Dicroglossidae: Dicroglossinae
Fejervarya sp.	20	Ranidae: Dicroglossinae	Dicroglossidae: Dicroglossinae
Fejervarya syhadrensis	5	Ranidae: Dicroglossinae	Dicroglossidae: Dicroglossinae
Fejervarya triora	2	Ranidae: Dicroglossinae	Dicroglossidae: Dicroglossinae
Hoplobatrachus rugulosa	1	Ranidae: Dicroglossinae	Dicroglossidae: Dicroglossinae
Hoplobatrachus tigerinus	7	Ranidae: Dicroglossinae	Dicroglossidae: Dicroglossinae
Hoplobratrachus occipitalis	4	Ranidae: Dicroglossinae	Dicroglossidae: Dicroglossinae
Hyperolius concolor	1	Hyperoliidae	Hyperoliidae
Ingerana tenasserimensis	1	Ranidae: Dicroglossinae	Dicroglossidae: Occidozyginae
Laliostoma labrosum	10	Ranidae: Mantellidae	Mantellidae: Laliostominae
Laliostoma labrosum	7	Ranidae: Mantellidae	Mantellidae: Laliostominae
Laliostoma labrosum	8	Ranidae: Mantellidae	Mantellidae: Laliostominae
Leptopelis occidentalis	1	Arthroleptidae: Leptopelinae	Arthroleptidae: Leptopelinae
Limnonectes bannaensis	2	Ranidae: Dicroglossinae	Dicroglossidae: Dicroglossinae
Limnonectes fragilis	3	Ranidae: Dicroglossinae	Dicroglossidae: Dicroglossinae
Limnonectes fujianensis	1	Ranidae: Dicroglossinae	Dicroglossidae: Dicroglossinae
Limnonectes grunniens	1	Ranidae: Dicroglossinae	Dicroglossidae: Dicroglossinae
Limnonectes jarujini	4	Ranidae: Dicroglossinae	Dicroglossidae: Dicroglossinae
Limnonectes kuhlii	5	Ranidae: Dicroglossinae	Dicroglossidae: Dicroglossinae
Limnonectes kuhlii	4	Ranidae: Dicroglossinae	Dicroglossidae: Dicroglossinae
Limnonectes laticeps	2	Ranidae: Dicroglossinae	Dicroglossidae: Dicroglossinae
Limnonectes megastomias	3	Ranidae: Dicroglossinae	Dicroglossidae: Dicroglossinae
Limnonectes modestus	1	Ranidae: Dicroglossinae	Dicroglossidae: Dicroglossinae
Limnonectes namiyei	1	Ranidae: Dicroglossinae	Dicroglossidae: Dicroglossinae
Limnonectes shompenorum	1	Ranidae: Dicroglossinae	Dicroglossidae: Dicroglossinae
Limnonectes taylori	4	Ranidae: Dicroglossinae	Dicroglossidae: Dicroglossinae
Nanorana parkeri	4	Ranidae: Dicroglossinae	Dicroglossidae: Dicroglossinae
Nanorana pleskei	5	Ranidae: Dicroglossinae	Dicroglossidae: Dicroglossinae
Nanorana ventripunctata	3	Ranidae: Dicroglossinae	Dicroglossidae: Dicroglossinae
Nasirana alticola	2	Ranidae: Raninae	Ranidae
Occidozyga borealis	2	Ranidae: Dicroglossinae	Dicroglossidae: Occidozyginae
Occidozyga lima	1	Ranidae: Dicroglossinae	Dicroglossidae: Occidozyginae
Occidozyga semipalmata	1	Ranidae: Dicroglossinae	Dicroglossidae: Occidozyginae
Occidozyga sp.	4	Ranidae: Dicroglossinae	Dicroglossidae: Occidozyginae
Odorrana sp.	1	Ranidae: Raninae	Ranidae
Paa arnoldi	6	Ranidae: Dicroglossinae	Dicroglossidae: Dicroglossinae
Paa boulengeri	9	Ranidae: Dicroglossinae	Dicroglossidae: Dicroglossinae
Paa bourreti	1	Ranidae: Dicroglossinae	Dicroglossidae: Dicroglossinae
Paa chayuensis	3	Ranidae: Dicroglossinae	Dicroglossidae: Dicroglossinae
Paa conaensis	1	Ranidae: Dicroglossinae	Dicroglossidae: Dicroglossinae
Paa exilspinosa	5	Ranidae: Dicroglossinae	Dicroglossidae: Dicroglossinae
Paa jiulongensis	2	Ranidae: Dicroglossinae	Dicroglossidae: Dicroglossinae
Paa liebigii	2	Ranidae: Dicroglossinae	Dicroglossidae: Dicroglossinae
Paa maculosa	2	Ranidae: Dicroglossinae	Dicroglossidae: Dicroglossinae
Paa medogensis	2	Ranidae: Dicroglossinae	Dicroglossidae: Dicroglossinae
Paa robertingeri	2	Ranidae: Dicroglossinae	Dicroglossidae: Dicroglossinae
Paa shini	3	Ranidae: Dicroglossinae	Dicroglossidae: Dicroglossinae
Paa sp.	3	Ranidae: Dicroglossinae	Dicroglossidae: Dicroglossinae
Paa sp.inosa	9	Ranidae: Dicroglossinae	Dicroglossidae: Dicroglossinae
Paa taihangnicus	2	Ranidae: Dicroglossinae	Dicroglossidae: Dicroglossinae
Paa verrucospinosa	6	Ranidae: Dicroglossinae	Dicroglossidae: Dicroglossinae
Paa yei	1	Ranidae: Dicroglossinae	Dicroglossidae: Dicroglossinae
Paa yunnanensis	12	Ranidae: Dicroglossinae	Dicroglossidae: Dicroglossinae
Phrynobatrachus africanus	6	Ranidae: Phrynobatrachinae	Phrynobatrachidae
Phrynobatrachus calcaratus	10	Ranidae: Phrynobatrachinae	Phrynobatrachidae
Phrynobatrachus plicatus	1	Ranidae: Phrynobatrachinae	Phrynobatrachidae
Ptychadena longirostris	3	Ranidae: Ptychadeninae	Ptychadenidae
Ptychadena newtoni	1	Ranidae: Ptychadeninae	Ptychadenidae
Ptychadena oxyrhynchus	4	Ranidae: Ptychadeninae	Ptychadenidae
Pyxicephalus edulis	7	Ranidae: Pyxicephalidae	Pyxicephalidae: Pyxicephalinae
Scotobleps gabonicus	4	Arthroleptidae	Arthroleptidae: Arthroleptinae
Shoutedenella xenodactylus	1	Arthroleptidae: Arthroleptinae	Arthroleptidae: Arthroleptinae
Sphaerotheca breviceps	1	Ranidae: Dicroglossinae	Dicroglossidae: Dicroglossinae
Sphaerotheca dobsonii	2	Ranidae: Dicroglossinae	Dicroglossidae: Dicroglossinae
Tomopterna cryptosis	2	Ranidae: Pyxicephalidae	Pyxicephalidae: Cacosterninae
Tomopterna delalandii	4	Ranidae: Pyxicephalidae	Pyxicephalidae: Cacosterninae

Table 2. Characterization of the gene regions and sampling used within these analyses. NAPC= nuclear autosomal proein coding MtNC=mitochondrial non-protein coding

gene region	gene name	gene type	initial fragment length	aligned fragment length	number of tips in dataset	%pairwise identity	SNPs	GC content(%)
ADNP	activity-dependent neuroprotector	NAPC	885	776	53	88.5	372	47
BDNF	brain-derived neurotrophic factor	NAPC	759	705	96	95.9	35	49.2
CYMex3	proto-oncogene cellular myelocytomatosis	NAPC	420	376	54	90.3	54	49.8
CXCR4	chemokine receptor type 4	NAPC	764	687	52	89.2	59	46.8
FSTL5	follistatin-like 5	NAPC	685	654	54	94.5	56	38.3
NCX1	sodium/calcium ion exhanger	NAPC	1323	975	35	89.4	0	41.3
NTF3	3'-nucleotidase	NAPC	655	525	28	86.3	0	44.3
PTGER4	prostoglandin E receptor 4 subtype EP4	NAPC	550	522	39	90.4	0	57
RAG1	recombination activiting gene 1	NAPC	1445	1398	51	93.7	49	44
RHOD	rhodopsin exon X	NAPC	368	315	193	94.8	28	47.3
SIA	seven in absentia	NAPC	484	405	62	92.1	47	51.2
TYR	tyrosinase	NAPC	758	634	152	90.4	109	49.9
ZEB2	zinc finger E-box homeobox 2	NAPC	1075	900	37	92.3	0	42.5
16s mtDNA	ribosomal RNA	MtNC	554	510	185	84.8	135	43.3

Table 3. Selection of Evolutionary Model of Nucleotide Substitution. Each gene treated under a single unpartitioned model. Selection based on three differing model selection criteria. AICc is the sample size corrected Akaike Information Criterion, BIC is the Bayesian Information Criterion, and DT is the Decision Theoretic. Gray highlighting indicates that the same model was selected by differing criteria.

		Nucleotide Substitution Model Analyses		
Gene	**Model Selection**	**AICc**	**BIC**	**DT**
ADNP	all basepairs	HKY + G / 6513.442 / 109	HKY + G / 6513.442 / 109	HKY + G / 6513.442 / 109
BDNF	all basepairs	GTR + I + G / 4019.3304 / 200	K80 + I + G / 4040.6758 / 193	JC / 4432.1782 / 190
CXCR4	all basepairs	HKY + G / 5002.418 / 107	HKY + G / 5002.418 / 107	JC / 5467.0773 / 102
CYMEX	all basepairs	GTR + G / 2782.4268 / 115	HKY + G / 2793.4481 / 111	HKY + G / 2793.4481 / 111
FSTL5	all basepairs	HKY + G / 3094.975 / 111	HKY + G / 3094.975 / 111	HKY + G / 3094.975 / 111
NCX1	all basepairs	HKY + G / 4730.0265 / 73	HKY + G / 4730.0265 / 73	HKY + G / 4730.0265 / 73
NTF3	all basepairs	GTR + G / 3203.5521/ 63	K80 + G / 3220.413 / 56	K80 + G / 3220.413 / 56
PTGER4	all basepairs	GTR + G / 3142.5038 / 85	HKY + G / 3149.8664 / 81	HKY + G / 3149.8664 / 81
RAG1	all basepairs	GTR + I + G / 5628.0669 / 110	GTR + G / 5631.1542 / 109	GTR + G / 5631.1542 / 109
RHOD	all basepairs	K80 + G / 2653.4525 / 386	HKY + G / 2642.2213 / 389	JC / 2926.5832 / 384
SIA	all basepairs	K80 + G / 3065.2911 / 124	K80 + G / 3065.2911 / 124	K80 + G / 3065.2911 / 124
TYRO	all basepairs	K80 + I + G / 7870.002 / 345	SYM + I + G / 7851.3093 / 349	JC / 8837.2828 / 342
ZEB	all basepairs	GTR + I + G / 4464.8113 / 82	HKY + I + G / 4474.0047 / 78	HKY + I + G / 4474.0047 / 78
16s mtDNA	all basepairs	HKY + I + G / 12061.0 / 374	GTR + I + G / 12032.8 / 378	JC / 15113.392 / 368

Table 4. Selection of Evolutionary Model of Nucleotide Substitution. Each codon within each gene treated as a unit. Selection based on three differing model selection criteria under a single partition. AICc is the sample size corrected Akaike Information Criterion, BIC is the Bayesian Information Criterion, and DT is the Decision Theoretic. Gray highlighting indicates that the same model was selected by differing criteria.

		Nucleotide Substitution Model Analyses		
Gene	**Model Selection**	**AICc**	**BIC**	**DT**
ADNP	first codon positions	K80 + G / 1504.6723 / 106	K80 + G / 1504.6723 / 106	K80 + G / 1504.6723 / 106
ADNP	second codon positions	F81 + G / 971.0715 / 108	F81 + G / 971.0715 / 108	F81 + G / 971.0715 / 108
ADNP	third codon positions	K80 + G / 3712.7537 / 106	HKY + G / 3704.1975 / 109	HKY + G / 3704.1975 / 109
BDNF	first codon positions	K80 + G / 3224.5016 / 192	K80 + I + G / 3164.9742 / 193	JC / 3509.7883 / 190
BDNF	second codon positions	JC / 885.9097 / 190	JC + I / 860.5577 / 191	JC / 885.9097 / 190
BDNF	third codon positions	K80 + G / 3224.5016 / 192	K80 + I + G / 3164.9742 / 193	JC / 3509.7883 / 190
CXCR4	first codon positions	JC / 1129.4321 / 102	K80 + G / 1120.7861 / 104	JC + G / 1124.868 / 103
CXCR4	second codon positions	JC / 909.5247 / 102	F81 / 894.0805 / 105	HKY + G / 894.1448 / 107
CXCR4	third codon positions	HKY + G / 2489.8508 / 107	HKY + G / 2489.8508 / 107	HKY + G / 2489.8508 / 107
CYMEX	first codon positions	JC / 595.5205 / 106	SYM + G / 546.7333 / 112	SYM + G / 546.7333 / 112
CYMEX	second codon positions	JC/ 434.9875 / 106	HKY + G / 409.8748 / 111	HKY + G / 409.8748 / 111
CYMEX	third codon positions	K80 / 1638.242 / 107	HKY + G / 1566.9772 / 111	HKY + G / 1566.9772 / 111
FSTL5	first codon positions	K80 + G / 684.9815 / 108	K80 + G / 684.9815 / 108	K80 + G / 684.9815 / 108
FSTL5	second codon positions	GTR + G / 507.2008 / 115	GTR + G / 507.2008 / 115	GTR + G / 507.2008 / 115
FSTL5	third codon positions	HKY + G / 1690.1151 / 111	HKY + G / 1690.1151 / 111	HKY + G / 1690.1151 / 111
NCX1	first codon positions	HKY + I + G / 903.9653 / 74	HKY + I / 906.0756 / 73	HKY + I / 906.0756 / 73
NCX1	second codon positions	F81 / 592.8941 / 71	F81 / 592.8941 / 71	HKY / 591.5922 / 72
NCX1	third codon positions	SYM + G / 2921.9681 / 74	K80 + G / 2933.008 / 70	K80 + G / 2933.008 / 70

		Nucleotide Substitution Model Analyses		
Gene	**Model Selection**	**AICc**	**BIC**	**DT**
NTF3	first codon positions	SYM + G / 802.6351 / 60	SYM + G / 802.6351 / 60	SYM + G / 802.6351 / 60
NTF3	second codon positions	HKY + G / 590.5211 / 59	HKY + G / 590.5211 / 59	HKY + G / 590.5211 / 59
NTF3	third codon positions	K80 + G / 1631.4706 / 56	K80 + G / 1631.4706 / 56	K80 + G / 1631.4706 / 56
PTGER4	first codon positions	K80 + G / 723.3878 / 78	K80 + G / 723.3878 / 78	K80 + G / 723.3878 / 78
PTGER4	second codon positions	JC + I / 446.6561 / 77	JC + I / 446.6561 / 77	JC + G / 446.8784 / 77
PTGER4	third codon positions	HKY + G / 1692.6118 / 81	GTR + G / 1679.3069 / 85	GTR + G / 1679.3069 / 85
RAG1	first codon positions	GTR + G / 1929.8925 / 109	GTR + G / 1929.8925 / 109	JC / 2055.7537 / 100
RAG1	second codon positions	HKY + G / 1713.0293 / 105	HKY + G / 1713.0293 / 105	K80 / 1829.2466 / 101
RAG1	third codon positions	HKY + G / 4005.0303 / 105	HKY + G / 4005.0303 / 105	JC / 4228.6741 / 100
RHOD	first codon positions	GTR + I + G / 598.8899 / 394	K80 + I + G / 609.1277 / 387	JC / 673.9055 / 384
RHOD	second codon positions	F81 + G / 271.548 / 388	F81 + G / 271.548 / 388	JC / 285.3424 / 384
RHOD	third codon positions	GTR + G / 1428.9358 / 393	HKY + G / 1431.348 / 389	JC / 1549.6738 / 384
SIA	first codon positions	JC / 509.996 / 122	K80 + I / 472.4462 / 124	K80 + I / 472.4462 / 124
SIA	second codon positions	JC / 399.9264 / 122	JC + I / 385.3583 / 123	JC + G / 385.4925 / 123
SIA	third codon positions	JC / 2063.5444 / 122	HKY + I + G / 1861.1054 / 128	HKY + I + G / 1861.1054 / 128
TYRO	first codon positions	GTR + G / 2262.9563 / 351	GTR + G / 2262.9563 / 351	JC / 2573.0702 / 342
TYRO	second codon positions	HKY + I + G / 1736.4116 / 348	HKY + I + G / 1736.4116 / 348	JC / 1848.7454 / 342
TYRO	third codon positions	GTR + I + G / 5037.6383 / 352	GTR + I + G / 5037.6383 / 352	JC / 6076.9926 / 342
ZEB	first codon positions	GTR + G / 1377.4561 / 81	GTR + G / 1377.4561 / 81	GTR + G / 1377.4561 / 81
ZEB	second codon positions	F81 + I + G / 1162.4823 / 77	F81 + I + G / 1162.4823 / 77	F81 + I + G / 1162.4823 / 77
ZEB	third codon positions	HKY + G / 2768.7837 / 77	HKY + G / 2768.7837 / 77	HKY + G / 2768.7837 / 77

Table 5. Summary of model selection in regard to model complexity.

		Number of Parameters		
Method A	**Method B**	**A > B**	**A = B**	**A < B**
AIC	BIC	10	28	19
AIC	DT	25	18	10
BIC	DT	19	23	1

Table 6. Likelihoods and Harmonic Likelihoods of Gene Trees. Likelihoods are reported on the left and harmonic likelihoods are given on the left.

	Likelihoods															
	Single Partition Nucleotide Substitution Models								Codon Partitioned Nucleotide Substitution Models							
	GTR + I+ G		AICc		BIC		DT		GTR + I+ G		AICc		BIC		DT	
ADNP	-6561.38	-6615.52	-6561.90	-6619.07	-6561.45	-6622.58	-6562.26	-6617.35	-6348.39	-6469.74	-6630.66	-6691.83	-6649.05	-6700.84	-6649.05	-6700.84
BDNF	-4273.66	-4362.21	-4273.66	-4362.21	-4296.06	-4380.15	-4512.09	-4601.49	-4016.05	-4085.65	-4209.84	-4310.85	-4180.33	-4285.28	-4512.87	-4604.62
CMYEX	-2851.96	-2937.10	-2844.87	-2938.94	-2856.59	-2942.85	-2856.59	-2942.85	-2761.35	-2818.07	-3064.00	-3117.80	-2795.57	-2869.30	-2798.36	-2860.39
CXCR4	-5045.11	-5119.40	-5053.31	-5112.53	-5053.31	-5112.53	-5505.72	-5554.17	-4897.37	-4961.22	-5187.37	-5242.46	-5106.16	-5165.98	-4975.08	-5033.03
FSTL5	-3153.62	-3242.00	-3160.19	-3225.85	-3160.19	-3225.85	-3160.19	-3225.85	-3063.27	-3117.01	-3105.74	-3178.32	-3105.74	-3178.32	-3105.74	-3178.32
NCX1	-4756.27	-4806.16	-4762.85	-4806.11	-4762.85	-4806.11	-4762.85	-4806.11	-4491.10	-4537.89	-4818.64	-4870.79	-4838.69	-4882.71	-4835.99	-4878.30
NTF3	-3227.02	-3262.48	-3227.06	-3260.58	-3241.29	-3275.77	-3241.29	-3275.77	-3137.43	-3181.69	-3198.95	-3231.93	-3198.95	-3231.93	-3198.95	-3231.93
PTGER4	-3176.88	-3232.13	-3178.40	-3229.18	-3183.46	-3233.64	-3183.46	-3233.64	-3014.37	-3058.02	-3103.27	-3156.82	-3092.08	-3139.76	-3080.45	-3132.94
RAG1	-5674.73	-5732.35	-5674.73	-5732.35	-5679.18	-5733.65	-5679.18	-5733.65	-5447.13	-5502.37	-5578.34	-5636.28	-5577.99	-5637.90	-6052.41	-6101.23
RHOD	-2944.45	-3031.85	-2972.72	-3068.42	-2961.84	-3070.01	-3184.25	-3328.31	-2751.78	-2862.34	-2853.50	-2955.03	-2873.24	_2969.62	-3176.01	-3307.98
SIA	-3125.30	-3208.33	-3112.01	-3208.04	-3112.01	-3208.04	-3112.01	-3208.04	-3005.18	-3064.29	-3469.08	-3524.03	-3090.82	-3160.71	-3085.49	-3169.06
TYRO	-8181.58	-8300.45	-8200.52	-8301.00	-8195.64	-8288.86	-8977.38	-9079.65	-7873.70	-7967.58	-8040.56	-8147.33	-8040.56	-8147.33	-8980.56	-9073.96
ZEB	-4499.40	-4549.70	-4499.40	-4549.70	-45-8.39	-4553.45	-45-8.39	-4553.45	-4401.24	-4454.29	-4424.60	-4472.15	-4423.14	-4470.02	-4422.95	-4484.18
16s mtDNA	-12204.23	-12342.21	-12229.89	-12343.18	-12204.23	-12342.21	-15285.49	-15402.39	na		na		na		na	

Table 7. Tests of Model Adequacy. Values in first column are the multinomial likelihood T(x) test statistic calculated in PuMA. P values should be centered in the predictive distribution around 0.5. Values on the tails are extremes and suggest that the model is inadequate. The best model for each gene is in bold.

Gene	Bayesian Analyses								
	T(x)	GTR 1 partition	AICc 1 partition	BIC 1 partition	DT 1 partition	GTR codon partitions	AICc codon partitions	BIC codon partitions	DT codon partitions
ADNP	-1094.6816	0.0001	0.0000	0.0000	0.0000	**0.0028**	0.0000	0.0000	0.0000
BDNF	-801.0168	0.0601	0.0601	0.0000	0.0000	**0.5388**	0.0000	0.0000	0.0000
CMYEX	-905.8302	0.0000	0.0000	0.0000	0.0000	**0.0181**	0.0000	0.0000	0.0000
CXCR4	-902.6631	0.0009	0.0000	0.0000	0.0000	**0.1097**	0.0000	0.0001	0.0000
FSTL5	-920.2201	0.0000	0.0000	0.0000	0.0000	**0.0853**	0.0000	0.0000	0.0000
NCX1	-538.7126	0.0188	0.0182	0.0182	0.0182	**0.1160**	0.0005	0.0003	0.0002
NTF3	-1272.2202	**0.0031**	0.0021	0.0012	0.0012	1.0000	0.0002	0.0000	0.0000
PTGER4	-472.3399	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
RAG1	-1047.5491	0.0389	0.0389	0.0161	0.0161	**0.7887**	0.0000	0.0001	0.0000
RHOD	-537.2077	0.0000	0.0000	0.0000	0.0000	**0.0333**	0.0024	0.0024	0.0000
SIA	-321.7051	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
TYRO	-753.8400	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
ZEB	-1503.9734	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

Table 8. (a) Bayes factors comparing unpartitioned and codon partitioned Bayesian gene trees generated under either GTR substitution models, or the nucleotide substitution models selected by AICc, BIC, and DT. Positive values support the codon partitioned model, negative values support the unpartitioned model. Numbers in bold show strong support for a model. (b) Bayes factors comparing among AICc, BIC, DT and GTR nucleotide models. A positive Bayes factor supports the first model listed, a negative Bayes factor supports the second model listed. Bold numbers show strong support for a model.

a.

	GTR (3) versus GTR (1)	AIC (3) versus AIC (1)	BIC (3) versus BIC (1)	DT (3) versus DT (1)
ADNP	**291.56**	**-145.52**	**-156.52**	**-166.98**
BDNF	**553.12**	**102.72**	**189.74**	-6.26
CMYEX	**238.06**	**-357.72**	**147.1**	**164.92**
CXCR4	**316.36**	**-259.86**	**-106.9**	**1042.28**
FSTL5	**249.98**	**95.06**	**95.06**	**95.06**
NCX1	**536.54**	**-129.36**	**-153.2**	**-144.38**
NTF3	**161.58**	**57.3**	**87.68**	**87.68**
PTGER4	**348.22**	**144.72**	**187.76**	**201.4**
RAG1	**459.96**	**192.14**	**191.5**	**-735.16**
RHOD	**339.02**	**226.78**	**200.78**	**40.66**
SIA	**288.08**	**-631.98**	**94.66**	**77.96**
TYRO	**665.74**	**307.34**	**283.06**	**11.38**
ZEB	**190.82**	**155.1**	**166.86**	**138.54**

b.

	GTR versus AIC	GTR versus BIC	GTR versus DT	AIC versus BIC	AIC versus DT	BIC versus DT
ADNP	**146.04**	**135.04**	**124.58**	**-11**	**-21.46**	**-10.46**
BDNF	**450.4**	**363.38**	**546.86**	**-87.02**	**96.46**	**183.48**
CMYEX	**-119.66**	**90.96**	**73.14**	**210.62**	**192.8**	**-17.82**
CXCR4	**56.5**	**209.46**	**-725.92**	**152.96**	**-782.42**	**-935.38**
FSTL5	**154.92**	**154.92**	**154.92**	0	0	0
NCX1	**407.18**	**383.34**	**392.16**	**-23.84**	**-15.02**	8.82
NTF3	**104.28**	**73.9**	**73.9**	**-30.38**	**-30.38**	0
PTGER4	**203.5**	**160.46**	**146.82**	**-43.04**	**-56.68**	**-13.64**
RAG1	**267.82**	**268.46**	**-275.2**	0.64	**-543.02**	**-543.66**
RHOD	**112.24**	**138.24**	**298.36**	**26**	**186.12**	**160.12**
SIA	**-343.9**	**193.42**	**210.12**	**537.32**	**554.02**	**16.7**
TYRO	**358.4**	**382.68**	**654.36**	**24.28**	**295.96**	**271.68**
ZEB	**35.72**	**23.96**	**52.28**	**-11.76**	**16.56**	**28.32**

Table 9. Bayes factors calculated for concatenated analyses. A positive value supports the model along the horizontal axis, a negative value supports the model along the vertical axis. Bold numbers show strong support for a model.

	AIC14	AIC40	BIC14	BIC40	DT14	DT40	GTR1	GTR2	GTR14	GTR40
AIC14										
AIC40	**1007.6**									
BIC14	**-539.52**	**-1547.12**								
BIC40	**-2389.1**	**-3396.7**	**-1849.58**							
DT14	**10263.5**	**9255.9**	**10803.02**	**12652.6**						
DT40	**8991.28**	**7983.68**	**9530.8**	**11380.38**	**-1272.22**					
GTR1	**649.98**	**-357.62**	**1189.5**	**3039.08**	**-9613.52**	**-8341.3**				
GTR2	**-244.24**	**-1251.84**	**295.28**	**2144.86**	**-10507.74**	**-9235.52**	**-894.22**			
GTR14	**-656.14**	**-1663.74**	**-116.62**	**1732.96**	**-10919.64**	**-9647.42**	**-1306.12**	**-411.9**		
GTR40	**-2454.96**	**-3462.56**	**-1915.44**	**-65.86**	**-12718.46**	**-11446.24**	**-3104.94**	**-2210.72**	**-1798.82**	

Table 10. Tree statistics for concatenated analyses.

tree	Partitions	Likelihood Score	Tree Length	Tree Height	N	Treeness (T)	External/Internal ratio
AICc	14	-66679.74	28.025091	3.086467	14.0810811	0.454160666	1.201863953
AICc	40	-67191.01	16.619656	1.853685	14.6702703	0.437373313	1.286376354
BIC	14	-66403.07	26.110275	2.885258	14.8702703	0.451703745	1.213840401
BIC	40	-65449.13	10.118862	1.15123	13.8378378	0.42630021	1.345764735
DT	14	-71813.36	15.16475	1.368702	13.2702703	0.427381065	1.339832254
DT	40	-71202.09	16.398847	1.459858	14.9351351	0.440149542	1.271955106
GTR	1	-66995.57	4.510612	0.423349	14.4918919	0.434913932	1.29930551
GTR	2	-66557.32	7.15001	0.810315	14.5189189	0.441824277	1.26334326
GTR	14	-66336.48	29.046756	3.114063	13.4810811	0.435518307	1.29611473
GTR	40	-65420.27	19.403642	2.125695	14.3459459	0.458597566	1.18056107
ML	1	-66825.608810	6.58394967	0.68026857	14.7243243	0.446479393	1.239745026
ML	2	-65417.353652	4.986348898	0.35060491	14.4378378	0.374964598	1.666918439
ML	14	-63356.455323	1296.875598	91.1886148	14.6756757	0.407365783	1.454796256
ML	40	-60347.941626	3559.650772	185.936256	14.8918919	0.432113061	1.31420915

Table 11. Results of topological congruency tests. The numbers below the diagonal are the pairwise *l*cong statistical scores, the numbers above the diagonal are the number of taxa held in common in the MAST between each tree pair. Higher numbers of taxa are indicative of greater topological congruence. The maximum number of taxa in these analyses is 185. For some tree pairs, the *l*cong statistic could not be calculated.

lcong	AIC14	AIC40	BIC14	BIC40	DT14	DT40	GTR1	GTR2	GTR14	GTR40	MLP1	MLP2	MLP3	MLP4
AIC14	x	156	163	161	145	149	150	163	172	157	147	185	138	131
AIC40	7.79	x	161	161	150	153	159	168	159	154	150	146	142	na
BIC14	8.14	8.04	x	176	160	165	163	169	166	171	160	139	na	na
BIC40	8.04	8.04	8.79	x	163	161	168	168	167	169	162	143	142	na
DT14	7.24	7.49	7.99	8.14	x	168	163	150	149	156	155	138	na	na
DT40	7.44	7.64	8.24	8.04	8.39	x	168	158	150	159	157	139	134	na
GTR1	7.49	7.94	8.14	8.39	8.14	8.39	x	160	157	162	160	141	136	na
GTR2	8.14	8.39	8.44	8.39	7.49	7.89	7.99	x	165	164	154	146	145	na
GTR14	8.59	7.94	8.29	8.34	7.44	7.49	7.84	8.24	x	161	149	142	139	na
GTR40	7.84	7.69	8.54	8.44	7.79	7.94	8.09	8.19	8.04	x	159	137	na	na
MLP1	7.34	7.49	7.99	8.09	7.74	7.84	7.99	7.69	7.44	7.94	x	147	144	na
MLP2	6.99	7.29	6.94	7.14	6.89	6.94	7.04	7.29	7.09	6.84	7.34	x	na	na
MLP3	6.89	7.09	na	7.09	na	6.69	6.79	7.24	6.94	na	7.19	na	x	145
MLP4	6.54	na	na	na	na	na	na	na	na	na	na	na	7.24	x

Table 12. Phylogenetic Informativeness. The net PI of each marker under each of the concatenated analyses. The highest PI value for each marker is indicated in bold.

	AIC14	AIC40	BIC14	BIC40	DT14	DT40	GTR1	GTR2	GTR14	GTR40	ML1	ML2	ML14	ML40
adnp	86.62	88.37	89.98	90.59	**98.93**	94.51	89.34	90.88	96.14	88.25	77.68	79.78	87.80	82.97
bdnf	22.61	25.19	23.30	24.29	**26.22**	24.77	25.11	25.44	25.47	22.94	20.74	21.84	22.72	20.81
cxcr	63.59	65.98	65.64	65.95	**74.01**	69.28	66.11	66.95	70.80	64.87	55.42	57.70	58.83	56.61
cymex	49.01	50.31	52.19	52.61	**55.89**	52.67	51.90	47.06	55.01	51.24	47.34	48.49	48.09	45.78
fstl	32.76	42.40	35.38	36.67	**44.47**	42.40	41.04	36.59	38.40	34.02	29.54	31.88	29.92	28.97
mtdna	113.56	113.90	114.79	115.28	120.27	**120.60**	114.21	115.36	116.43	115.24	104.39	105.37	106.46	104.97
ncxi	58.43	61.63	61.14	62.38	**74.38**	66.44	61.21	62.67	67.42	59.26	49.46	54.00	52.86	52.16
ntf	57.28	59.71	59.10	59.82	64.77	61.45	59.58	61.57	**64.96**	57.71	52.38	53.84	57.44	57.28
ptger	45.63	50.09	46.59	50.03	51.32	48.48	49.62	50.16	**53.21**	45.41	40.76	41.87	45.78	42.67
rag	116.00	118.75	117.80	118.01	**132.41**	123.23	118.01	118.73	124.74	118.83	98.51	102.36	99.92	94.27
rhod	12.64	13.70	12.94	13.78	**15.35**	13.80	14.10	12.95	14.06	13.74	11.83	12.00	12.11	10.83
sia	39.40	40.52	41.53	41.87	**44.40**	43.36	39.21	41.53	42.99	39.33	36.10	37.93	40.29	38.10
tyros	61.23	63.67	61.43	63.68	66.61	63.93	64.49	63.39	**66.78**	61.18	52.01	53.00	55.69	52.78
zeb	54.43	56.17	56.98	58.01	**63.47**	59.10	55.26	57.89	62.28	55.26	49.40	51.06	55.47	54.10