Mountains and Migrations: Using Biogeography, Population and Landscape Genetics to Elucidate Patterns and Processes of Diversification in a Biodiversity Hotspot

By

Chan Kin Onn

Submitted to the graduate degree program in Ecology and Evolutionary Biology and the Graduate Faculty of the University of Kansas in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

Date Defended: 10 Nov 2017

The dissertation committee for Chan Kin Onn certifies that this is the approved version of the following dissertation:

Mountains and Migrations: Using Biogeography, Population and Landscape Genetics to Elucidate Patterns and Processes of Diversification in a Biodiversity Hotspot

_____

Chairperson Rafe M. Brown

Date Approved: 6 December 2017

ABSTRACT

Elucidating the tempo and mode of diversification is a major goal of evolutionary biology and represents a fundamental step towards understanding how biodiversity is generated and maintained. Achieving this goal is challenging due to the multidimensional complexity of macro- and micro-evolutionary forces that act across time, space and different phylogenetic levels. Furthermore, these evolutionary forces can involve both adaptive and neutral processes that form confounding interactions with landscape/environmental characteristics. Understanding how genes, geography, and ecology interact to generate, maintain and distribute biodiversity, therefore requires numerous datatypes and analytical resources that span different disciplines such as biogeography, population, and landscape genetics. The primary goal of this study is to provide a better understanding of the patterns and processes involved in the diversification of frogs from the family Ranidae through the use of genomic data and recent advances in analytical methods.

Malaysia is one of the most biodiverse and environmentally threatened countries in the world. It is part of Sundaland, a biodiversity hotspot that has undergone dramatic climatic fluctuations in the past few million years. The dynamic geological history, coupled with the highly heterogeneous landscape of Malaysia, provides an ideal system to study the patterns and processes of diversification within a spatial and temporal context. Such studies are of particular importance in Malaysia as it is the country with the highest rate of deforestation in the world. Moreover, most biodiversity research in Malaysia revolves around systematic revisions and species descriptions with relatively few studies focused on understanding the evolutionary underpinnings that generate diversity. This study will fill a significant gap in the region's

biodiversity research, especially as it represents the first genomic study on Malaysian amphibians.

Chapter 1 examines the broad biogeographic patterns of Ranid diversification at the family and generic level. Using the most comprehensive and robust time-calibrated phylogeny to date, we estimated the timing and patterns of major dispersal events to test the hypothesis that colonization of new geographic areas triggers a concomitant acceleration in diversification rates. Additionally, we determined whether the Eocene-Oligocene extinction event (EOEE) had a significant impact on the diversification of Ranids. Our results showed that the EOEE had no effect on diversification rates; most major dispersal events occurred over a relatively short period of time during the end of the Eocene, and the colonization of new geographic areas was not followed by increased net-diversification. On the contrary, diversification rate declined or did not shift following geographic expansion. Thus, the diversification history of Ranid frogs contradicts the prevailing expectation that amphibian net-diversification accelerated towards the present or increased following range expansion. Rather, our results demonstrate that despite their dynamic biogeographic history, the family Ranidae diversified at a relatively constant rate despite their present high diversity and circumglobal distribution.

Chapter 2 compares the efficacy of commonly-used species delimitation methods (SDMs) and a population genomics approach based on genome-wide single nucleotide polymorphisms (SNPs) to assess lineage separation in the Malaysian Torrent Frog Complex currently recognized as a single species (*Amolops larutensis*). First, we used morphological, mitochondrial DNA and genome-wide SNPs to identify putative species boundaries by implementing non-coalescent and coalescent-based SDMs. We then tested the validity of putative boundaries by estimating spatiotemporal gene flow to assess the extent of genetic

separation/cohesion among putative species. Our results show that SDMs were effective at delimiting divergent lineages in the absence of gene flow but overestimated species in the presence of marked population structure and gene flow. However, using a population genomics approach and the concept of species as separately evolving metapopulation lineages as the only necessary property of a species, we were able to objectively elucidate cryptic species boundaries in the presence of past and present gene flow.

Chapter 3 builds on the findings of the previous chapter to determine the spatiotemporal factors that generated the genetic differentiation in *Amolops*. We tested the significance and relative contributions of (1) geographic distance (isolation-by-distance, IBD); (2) landscape/environmental variables including mountain ranges, river basins, forest cover, and habitat suitability (isolation-by-environment, IBE); and (3) historical events (isolation-by-colonization, IBC) towards genetic differentiation. Results showed that interspecies diversification was primarily driven by historical events (IBC); IBD was responsible for intraspecies population structure, and IBE did not play a significant role in the diversification of *Amolops*. Additionally, demographic analyses detected significant population bottlenecks, indicating that speciation was most likely due to founder effect speciation.

ACKNOWLEDGEMENTS

I thank the Biodiversity Institute (KUBI), Department of Ecology and Evolutionary Biology (EEB), Graduate Student Organization, National Science Foundation Doctoral Dissertation Improvement Grant (DEB-170203) and National Geographic Society (Grant number 9722-15) for providing support and funding for this dissertation research. I would like to specifically thank the Director of the KUBI Dr. Leonard Krishtalka, the Chair of EEB Dr. Christopher Haufler, the EEB Graduate Coordinator Aagje Ashe, Lori Schlenker and Jamie Keeler at KUBI for their time, patience and technical assistance throughout my time at KU.

My committee members Town Peterson, Rob Moyle, Mark Mort, Li Xingong, and especially my major advisor Rafe Brown have been invaluable sources of guidance and support. My graduate career would not have been possible without the support and mentorship of my friend, collaborator, and mentor, Lee Grismer who played a major role in my scientific development. I would like to thank past and present graduate students, and especially members of the herpetology division who have assisted and supported me in various ways: Bill Duellman, Linda Trueb, Rich Glor, Jeet Sukumaran, Jamie Oaks, Charles Linkem, Cameron Siler, Carl Oliveros, Patrick Monnahan, Anthony Barley, Luke Welton, Alana Alexander, Jesse Grismer, Perry Wood Jr., Scott Travers, Karen Olson, Tashitso Anamza, Kerry Cobb, David Blackburn, Luke Mahler, Su Yong Chao, Matt Buehler, Robin Abraham, Carl Hutter, Kevin Chovanec, Pietro de Mello, Katie Allen, Walter Tapondjou, Jeff Wienell, Matt Jones, Javier Torres and Mark Herr.

I thank my family who nurtured my love and interest in nature and never once pressured me to pursue anything except that which I am passionate about. Last but not least, I thank my partner in life, Lynn for providing unwavering love and moral support. She, together with our

dogs Sophie, Tori and Suki have brought joy, balance and sanity to my life when times got difficult.

CONTENTS

Chapter 2. Species delimitation with gene flow: a methodological comparison and population genomics approach to elucidate cryptic species boundaries in Malaysian Torrent Frogs

CHAPTER 1

Spatiotemporal diversification of True Frogs (family: Ranidae): testing the "Dispersification" and Mass-Extinction Hypotheses

Chan, KO and Brown, RM. 2017. Did true frogs "dispersify"? *Biology Letters*, 13(8): 20170299

1.1 Abstract

The interplay between range expansion and concomitant diversification is of fundamental interest to evolutionary biologists, particularly when linked to intercontinental dispersal and/or large scale extinctions. The evolutionary history of True Frogs has been characterized by circumglobal range expansion. As a lineage that survived the Eocene-Oligocene extinction event (EOEE), the group provides an ideal system to test the prediction that range expansion triggers increased net-diversification. We constructed the most densely sampled, time-calibrated phylogeny to date in order to: (1) characterize tempo and patterns of diversification; (2) assess the impact of the EOEE; and (3) test the hypothesis that range expansion was followed by increased net-diversification. We show that late Eocene colonization of novel biogeographic regions was not affected by the EOEE and surprisingly, global expansion was not followed by increased net-diversification. On the contrary, diversification rate declined or did not shift

following geographic expansion. Thus, the diversification history of True Frogs contradicts the prevailing expectation that amphibian net-diversification accelerated towards the present or increased following range expansion. Rather, our results demonstrate that despite their dynamic biogeographic history, True Frogs diversified at a relatively constantly rate, even as they colonized the major landmasses of Earth.

1.2 Introduction

Geographic distribution of species richness is a function of lineage diversification through time (rate of diversification) and space (movement of species). Therefore, changes in diversification rates can be an important factor underlying geographic patterns of biodiversity. Rate of diversification is defined as speciation minus extinction over time. In order for a diversification rate shift to occur, the probability of speciation and/or extinction has to be altered, for example through increased speciation (Rolland *et al.* 2014), elevated extinction (Brocklehurst *et al.* 2015), or lower extinction rates (Cracraft 1985). Increased speciation is often linked to intrinsic factors such as the evolution of key innovations (Drummond *et al.* 2012; Hodges 1997; Ostrom 1979; Woodburne *et al.* 2003) or movement of lineages into new areas (Moore & Donoghue 2007; Uribe-Convers & Tank 2015). Although key innovations play an important role, Moore and Donoghue (2007) suggested that dispersal followed by subsequent diversification ("dispersification") may play a larger role in shaping global biodiversity patterns than previously recognized. Diversification rate shifts have also been associated with extrinsic events such as mass extinctions that can cause increased extinction (Brocklehurst *et al.* 2015) or a burst of speciation if surviving lineages rapidly diversify into vacant niches following an extinction event (Feng *et al.* 2017; Ksepka *et al.* 2017; Ruta *et al.* 2007). Understanding the deterministic factors

that underlie diversification rate shifts therefore requires reliable calculations of speciation and extinction rates as well as robust estimates of phylogenetic relationships, divergence times and inference of geographic range evolution.

Estimating diversification rates across phylogenetic trees is a central component of macroevolutionary studies. Although this field has seen much advancement in recent years (Laurent *et al.* 2015; May *et al.* 2016; Rabosky 2014; Stadler & Smrckova 2016), the reliable estimation of this parameter remains complicated. This is exemplified by the "Great BAMM Controversy" that revolves around reputed theoretical flaws in the popular program BAMM that has been widely used to detect the timing and location of diversification rate shifts in phylogenies (Moore *et al.* 2016; Rabosky 2014; Rabosky *et al.* 2017). Nevertheless, this program is still being used, even in a number of recent high-profile publications involving the estimation of diversification rates across large amphibian phylogenies (Feng *et al.* 2017; Yuan *et al.* 2016), thereby casting doubts on the validity of those results.

Studies have shown that the diversification rates of amphibians have accelerated toward the present (Roelants *et al.* 2007), were positively associated with global range expansion (Pramuk *et al.* 2008) and increased following mass extinction events (Feng *et al.* 2017). Yuan *et al.* (2016) used the controversial program BAMM to infer spatiotemporal diversification patterns of the genus *Rana* (family Ranidae) but no study has been conducted on the entire family or other subclades within the family. This represents a significant gap in our understanding of the evolutionary history of this diverse (~380 species) and cosmopolitan amphibian family (Frost 2015). Furthermore, the evolutionary history of Ranidae spans the Eocene-Oligocene mass extinction event (EOEE), which triggered worldwide extinctions in marine invertebrates (Pearson *et al.* 2008) and a marked synchronous turnover of flora and fauna in Europe and Asia

(Hooker *et al.* 2004; Sun *et al.* 2014). However, the impact of the EOEE on amphibians has never been explicitly studied.

To obtain a well resolved phylogeny for systematic treatments and downstream macroevolutionary analyses, I synthesized the most comprehensive multilocus molecular dataset to date from a large proportion (77%) of species and genera from the family Ranidae. Having a fully resolved, time-calibrated phylogeny will facilitate biogeographic reconstructions of the family Ranidae to understand the mode and tempo of long-distance intercontinental dispersal in amphibians. I can then estimate the significance and timing of shifts in diversification rates using a program that hasn't been shown to be flawed (Hohna *et al.* 2015) to test the hypothesis that movement into new geographic areas and mass extinction events are followed by a concomitant increase in diversification rate (the "Dispersification" and Mass Extinction Hypotheses).

1.3 Methods

1.3.1 Sampling and time-calibrated phylogenetic estimation

A total of 402 genetic samples representing 292 of the known 380 Ranidae species were obtained from Genbank representing two mitochondrial (16S, Cytochrome-b) and two nuclear genes (RAG-1, Tyrosinase; Appendix Table 1). These samples were aligned using the MUSCLE algorithm implemented in the program Geneious (Kearse *et al.* 2012). A total of 4,328 base-pairs were concatenated and partitioned by gene prior to phylogenetic analysis using the Bayesian program BEAST (Drummond & Bouckaert 2015). We used BEAST's bModelTest to explore substitution model space while simultaneously estimating model parameters and the phylogeny (Bouckaert & Drummond 2017). To establish a temporal framework for Ranidae diversification,

we used four fossil calibration points: three within the genus *Rana* (Yuan *et al.* 2016) and one to calibrate the most recent common ancestor of the genus *Pelophylax* (Roček & Rage 2003). The fossilized birth-death process was used to model speciation times using the following prior settings: originFBD ~ Uniform(30, 100); samplingProportionFBD and turnoverFBD ~ Uniform(0, 1.0); ucldMean ~ Exponential(10, 0); and ucldStdev ~ Gamma(0.5396, 0.3819). Fossil calibration times for the genus *Rana* follows Yuan *et al.* (2016) while the age of the most recent common ancestor of the genus *Pelophylax* was constrained to 30 million years (Roček & Rage 2003). Sampling of the posterior distribution was performed using two independent MCMC chains at 500,000,000 million generations each. The posterior distribution was considered adequately sampled when effective sample size (ESS) values for parameters were greater than 200. The BEAST plugin LogCombiner was used to combine sampled trees from separate MCMC runs and TreeAnnotator was used to generate a maximum clade credibility tree for downstream analyses.

1.3.2 Ancestral range reconstruction

The resulting time-calibrated phylogeny was used to reconstruct the spatio-temporal evolution of geographic ranges. We defined seven biogeographic regions that are known to be separate landmasses during the Cenozoic: America, Europe, Africa, Asia, India, Philippines and Australasia (Fig. 1.1). We distinguished the Indian subcontinent based on its unique geological history as a separate landmass prior to its collision with mainland Asia during the Eocene (Morley 2012). Model parameters and ancestral range probabilities were estimated using the DEC, DIVALIKE and BAYAREALIKE models. These models were implemented with and without the parameter *j*, which describes the relative weight of founder event speciation during

cladogenesis. Under this model, which resembles founder event speciation, an ancestor in area A instantaneously "jumps" to area B, leaving one descendent in A and one in B (Matzke 2013). Likelihood rations and the Akaike Information Criterion was used to evaluate the fit of our data to the biogeographic models. Ancestral range reconstruction analysis was performed using the BioGeoBEARS R package (Matzke 2014).

1.3.3 Diversification Rate Analyses

Due to the theoretical flaws in the widely used program BAMM (Moore *et al.* 2016; Rabosky 2014), I estimated diversification rates using the program TESS that provides a flexible framework for specifying diversification models, where diversification rates are constant, vary continuously, or change episodically through time. One of the major features of TESS is the ability to include methods of incomplete taxon sampling while providing robust Bayesian methods for assessing the relative fit of lineage diversification models to a given tree. Reversible-jump MCMC (rjMCMC) simulation is performed over all possible episodically varying birth-death processes with explicitly modeled mass extinction events. Briefly, this method treats the number of speciation-rate shifts, extinction-rate shifts and mass extinction events as well as the parameters associated with these events as random variables and estimates their joint posterior distribution (Höhna *et al.* 2015; May *et al.* 2016). Because it can be difficult to specify the prior distributions for the speciation- and extinction-rate parameters, I first implemented an automatic empirical hyperprior procedure which performs an initial Bayesian MCMC analysis under a constant-rate birth-death process model to determine reasonable values for the hyperparameters of the diversification priors. Using the estimated parameter values, I performed the TESS analysis, allowing the MCMC chain to run until all parameters reached a

minimum effective sampling size of 500. Because finer-scale or lineage-specific changes may be obscured when larger groups are analysed as a whole (Yuan *et al.* 2016), we performed analyses at three hierarchical levels to capture diversification shifts at different phylogenetic scales: (1) entire phylogeny; (2) three major subclades: Torrent Frogs (*Amolops*), Afro-Asian Stream Frogs [*Hylarana* Complex *sensu* Oliver *et al.* (2015)], Core True Frogs (*Babina, Odorrana, Rana*); and (3) individual subclades that colonized a new region. This analysis was performed using the R package TESS (Höhna *et al.* 2015).

1.4 Results

1.4.1 Time calibration and phylogenetic relatiosnhips

The phylogenetic analysis produced high support (posterior probability ≥0.9) for most major nodes and divergence time estimates (Appendix Fig. 1) were generally consistent (± 5 million years) with two previous, more sparsely-sampled phylogenetic studies of Ranidae (Wiens *et al.* 2009; Yuan *et al.* 2016) but strongly conflicted with another study, which recovered substantially younger ages (Oliver *et al.* 2015). The family Ranidae diverged from its sister lineage at ~87 mya (95%HPD: 73–99 mya), followed by a cladogenetic event at ~70.4 mya (95%HPD: 57–86) that split the most basal genus *Staurois* with the other lineages. Most genera were recovered as monophyletic clades except for the genera *Huia, Glandirana, Sylvirana* and *Amnirana*. *Huia cavitympanum* was reciprocally monophyletic with respect to *Meristogenys* (pp = 1.0) and polyphyletic with regard to *H. masonii* and *H. sumatrana* that formed a sister relationship with *Clinotarsus*. *Glandirana* was paraphyletic with respect to *Sanguirana* with high support (pp = 0.9). The true *Sylvirana* clade was sister to *Papurana* + *Hydrophylax* with high support (pp = 1.0). However, *S. guentheri* samples were sister to the genus *Humerana* while *S. attigua,* which

was only represented by a single specimen was sister to *Amnirana nicobariensis*. African *Amnirana* was recovered as a distinct monophyletic clade sister to *Indosylvirana* while *A. nicobariensis*, which is widely distributed across Asia was sister to the clade that includes African *Amnirana* + *Indosylvirana* + *Hydrophylax* + *Papurana* + *Sylvirana* albeit with low support (pp = 0.3).The sister relationship between *Babina* and *Odorrana* was also weakly supported (*pp* = 0.2) (Appendix Fig. 1).

1.4.2 Ancestral range reconstruction

Ancestral range reconstructions unambiguously support Asia as the origin of True Frogs (Fig. 1.1). Of the six biogeographic models assessed, the BAYAREALIKE+J model was favored (Appendix Fig. 2; Table 1.1). A total of 11 major dispersal events were detected, the clear majority of which (n=9) occurred during the Eocene and Oligocene including three colonizations of the Indian subcontinent by the genera *Clinotarsus* (~25–39 mya), *Indosylvirana* (~ 30–40 mya) and *Hydrophylax* (~ 21–35 mya); one colonization of Europe by *Pelophylax* (~ 23–35 mya); one dispersal into the Philippines by S*anguirana* (~ 22–38 mya); two separate invasions into the Americas by *Rana* (~ 33–39 mya and 17–26 mya); one to Africa by *Amnirana* (~ 37–40 mya); and one to Australasia by *Papurana* (~ 24–35 mya). Two major dispersals occurred during the Miocene including a second invasion of the Philippines by *Pulchrana* (~ 20–25 mya); a second colonization of Europe by *Rana* (~ 14–20 mya).

1.4.3 Diversification rate-shifts

There was a strong decrease in speciation rates in Afro-Asian Stream Frogs at ~35 mya and in Core True Frogs at ~7 mya ($2 < 2\ln BF > 6$), but Torrent Frogs showed no significant shifts (Fig.

1.2). No shifts were detected in individual subclades except for the genus *Rana* which had a similar pattern to the entire Core True Frog clade (Fig. 1.2, 1.8). When the analysis was performed on the entire phylogeny, a strong decrease in speciation rate ($2\ln BF \approx 5$) was detected at ~6 mya and no shifts in extinction rates, nor signatures of mass-extinction were detected (Fig. 1.9).

1.5 Discussion

1.5.1 Diversification rates

Two centuries of global biogeographical and paleontological treatments have reinforced the expectation that major biotic range expansion into new geographic areas is often followed by an increase in net-diversification rates, due to a broad range of phenomena that fall under the general concepts of ecological opportunity and evolutionary innovation (Wiens & Donoghue 2004). Interestingly, and in striking contrast to expectations, the rapid global range expansion of True Frogs was not associated with increased net-diversification. On the contrary, diversification rates either decreased or remained unchanged, even as remarkable, circumglobal range expansion via dispersal and colonization of novel regions occurred. The Afro-Asian Stream Frogs, which underwent the most extensive and rapid range expansion, actually exhibited a decrease in net-diversification following range expansions at the Eocene-Oligocene boundary. A similar but stronger shift has been reported in Core True Frogs during the Miocene (Yuan *et al.* 2016). Although studies have shown that an overall decrease in diversification rates in larger groups (e.g. at the family or generic level) may obscure increased diversification rates in smaller

subgroups (Yuan *et al.* 2016), our results showed that diversification rate patterns were consistent across different phylogenetic scales.

1.5.2 Timing and patterns of diversification

Our divergence time estimates are largely congruent with the timing of several well characterized tectonic events. All three independent colonizations of the Indian subcontinent occurred between 35–40 mya, after the Indian-Eurasian collision at 40 mya (Bouilhol *et al.* 2013). The separation between eastern and western Paleartic lineages of *Pelophylax* was estimated at ~35 mya, coinciding with the closure of the Turgai Straits, an event which resulted in a land connection between southern Europe and southwestern Asia, and facilitation of faunal exchange between these two regions (Sun *et al.* 2014). The colonization of the Philippines from east Asia by members of the genus *Glandirana* precisely matches the timing, polarity of inferred dispersal, and phylogenetic relationships postulated by the "Palawan Ark" (Blackburn *et al.* 2010; Yumul *et al.* 2009), reinforcing the interpretation of isolation and paleotransport of True Frog lineages on the Palawan Microcontinent Block. This minor continental fragment rifted from the Asian mainland ~40 mya, opening up the South China Sea, eventually lodged north of Borneo (~10–6 mya) as the oceanic portions of the Philippines drifted north along the Mobile Philippine Belt. The existence of an extensive, archipelago-wide *Glandirana* radiation in the Philippines, most closely related to northern coastal Asian and Eurasian lineages (but conspicuously absent from the remainder of the Sunda Shelf landmasses to the south), and with the Palawan endemic lineage first branching and sister to the remainder of the Philippine species, precisely resembles patterns inferred previously in several other radiations of Philippine land vertebrates (Brown *et al.* 2013) , lending strong statistical support in yet another independent

lineage, for the Palawan Ark biogeographic mechanism which has initiated several spectacular and entirely endemic Philippine radiations.

True Frogs began to disperse out of Asia at the end of the Eocene and by the beginning of the Miocene, colonized every continent except Antarctica. Yuan and colleagues demonstrated that the New World was colonized via the Beringian land bridge (Yuan *et al.* 2016). In contrast, our results showed that the dispersal of *Amnirana* from India/Asia into Africa at ~37–40 mya could not have occurred over land as Africa and Eurasia were separated by the neo-Tethys ocean until ~27 mya (McQuarrie & Van Hinsbergen 2013). This dispersal event coincides with the middle Eocene climatic optimum (MECO), a period of pronounced warming in the middle to late Eocene (Bohaty & Zachos 2003). Numerous other intercontinental faunal exchanges have been documented during this period (Beard *et al.* 1994; Chaimanee *et al.* 2012), indicating that the MECO could have been an important facilitator of intercontinental faunal exchange in many unrelated vertebrate groups.

1.5.3 Amphibian diversification during the EOEE

The EOEE was followed by accelerated extinction rates in marine-life, mammals and vegetation (Sun *et al.* 2014). However, we found no evidence for a EOEE associated reduction of diversity in True Frogs. Extinction rates remained relatively constant through time, indicating that shifts in net-diversification were caused by decreased speciation rates, likely due to global cooling during the Oligocene or the lack of key innovations that prevented ecological generalists from competing with incumbent species (Duellman & Trueb 1994). The diversification history of True Frogs goes against the current thesis that amphibian net-diversification accelerated towards the

present (Roelants *et al.* 2007) or increased following range expansion (Pramuk *et al.* 2008). Our study demonstrates that, despite their dynamic biogeographic history of pan-global range expansion, True Frog diversity did not increase with range expansion. Rather, a relatively constant, linear accumulation of taxonomic diversity through time, coupled with instances of decreased speciation, characterized the evolutionary history of the planet's cosmopolitan True Frogs.

1.5.4 Systematics

Although our phylogenetic analysis resolved a number of clades that were poorly supported in previous related studies (Oliver *et al.* 2015; Pyron & Wiens 2011; Wiens *et al.* 2009), numerous clades (e.g. *Huia, Glandirana, Sylvirana, Amnirana, Babina*) remain poorly supported with unstable topologies. The polyphyly of *Huia* has been shown in previous studies (Che *et al.* 2007; Stuart 2008) and likely reflects incorrect taxonomy as opposed to evolutionary convergence or parallel evolution. This study also provides stronger support for the paraphyly of *Glandirana* (Brown *et al.* 2013) which will most likely result in the synonymization of *Sanguirana* with the more senior nomen *Glandirana* (Frost 2015). The taxonomic placement and monophyly of *Amnirana* remains a challenging problem to resolve. Almost every study that has included this genus has produced conflicting phylogenetic relationships albeit with low support: sister to *Chalcorana* + *Pulchrana* (Pyron & Wiens 2011; Wiens *et al.* 2009); sister to *Amolops ricketti* (Huang *et al.* 2014); sister to *Humerana* + *Hylarana* + *Indosylvirana* + *Sylvirana* + *Hydrophylax* + *Papurana* (Oliver *et al.* 2015). Additionally Oliver *et al.* (2015) and Pyron & Wiens (2011) found African and Asian *Amnirana* to be monophyletic while Wiens, Sukumaran, Pyron & Brown (2009) and our study found African *Amnirana* to be polyphyletic with regard to

Asian *Amnirana*. This conflict has significant biogeographic implications regarding how *Amnirana* colonized Africa. Unfortunately our analysis also failed to resolve the phylogenetic placement of *Amnirana* and more data is probably needed to fully resolve the poorly supported nodes within the Ranidae phylogeny.

1.6 Tables

Table 1.1 Model comparison results from the BioGeoBEARS analysis.

	LnL	numparams	d	e	j
DEC	-147.5982351	2	0.000644772	1.00E-12	0
DEC+J	-141.8861955	3	0.000437178	1.00E-12	0.001632332
DIVALIKE	-157.8676818	2	0.000789643	1.00E-12	0
DIVALIKE+J	-152.3083492	3	0.000454217	2.85E-11	0.002507451
BAYAREALIKE	-157.5969216	2	0.000417869	0.003936621	0
BAYAREALIKE+J	-109.5805316	3	8.03E-05	0.001191031	0.004599812

1.7 Figures

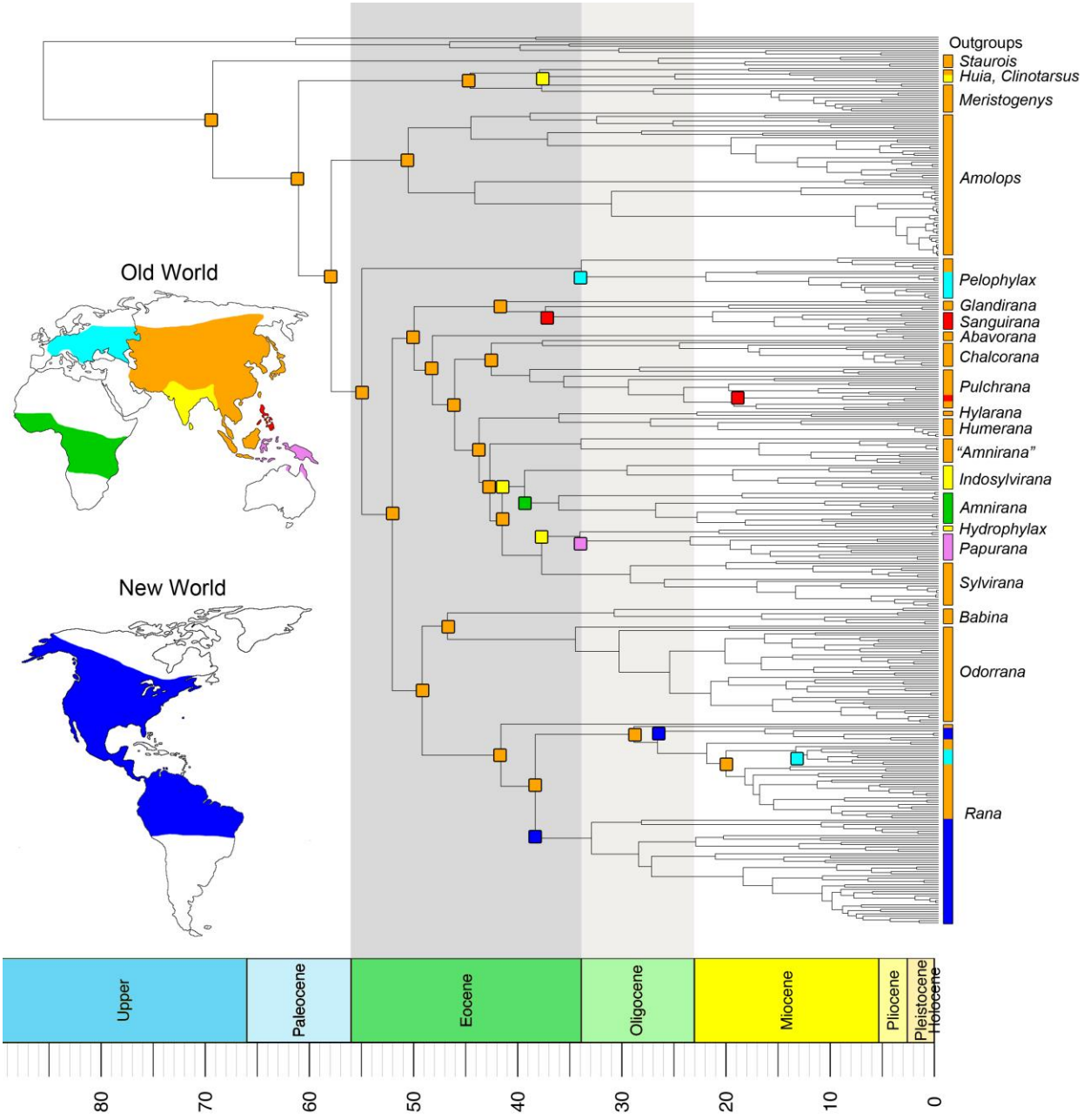


Figure 1.1 Time-calibrated phylogeny and ancestral range reconstructions of the family Ranidae

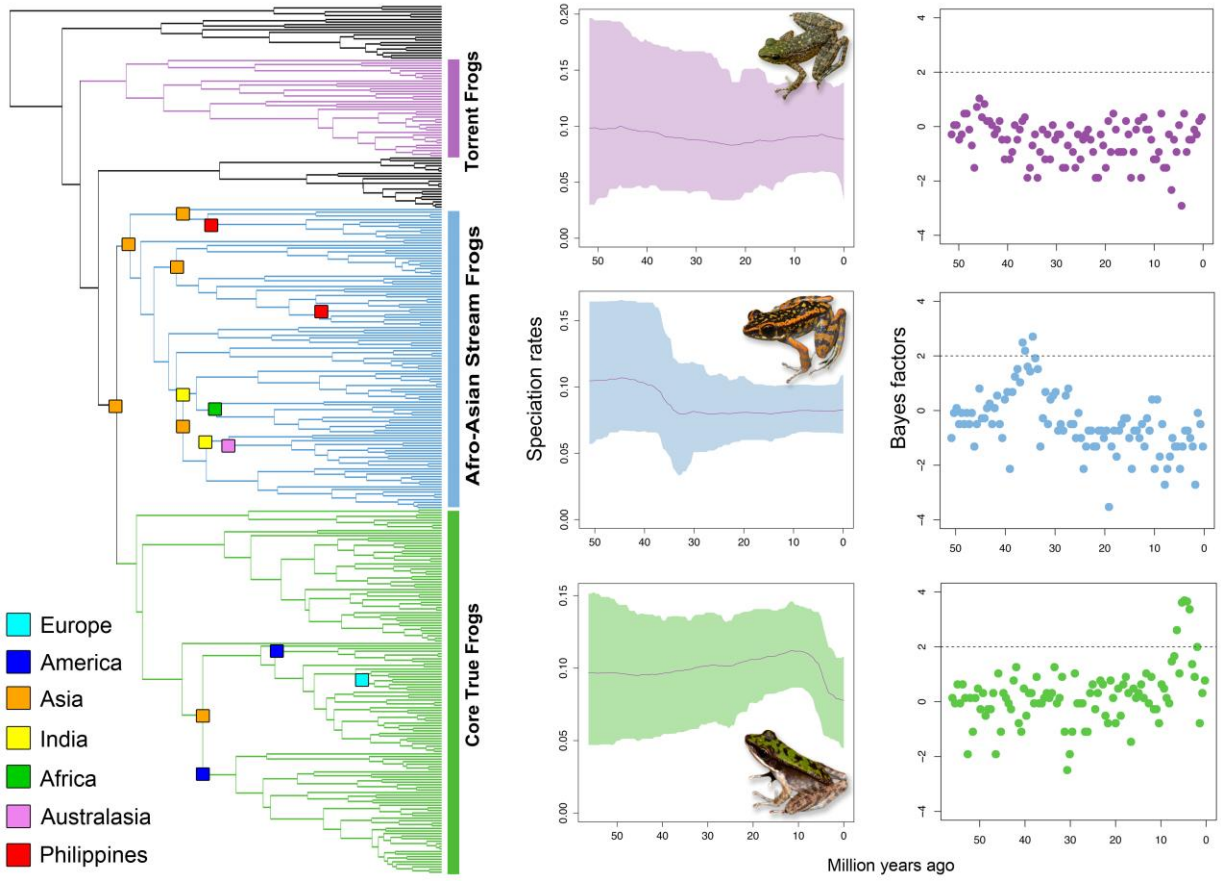


Figure 1.2 Speciation rates through time in three major clades with significance assessed using Bayes factors.

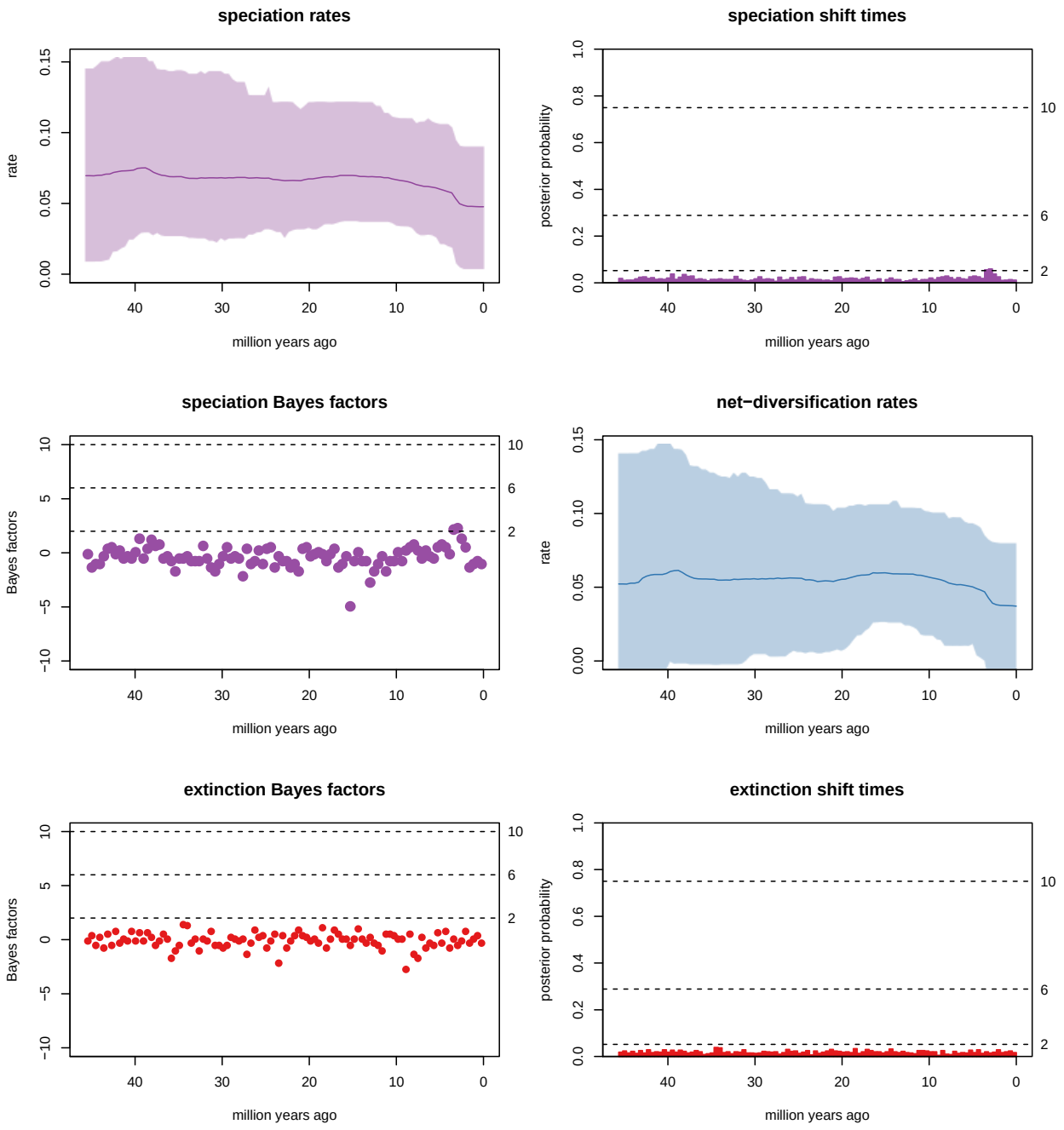


Figure 1.3 Results of the TESS analysis for the combined clade of *Clinotarsus* + *Huia* + *Meristogenys*

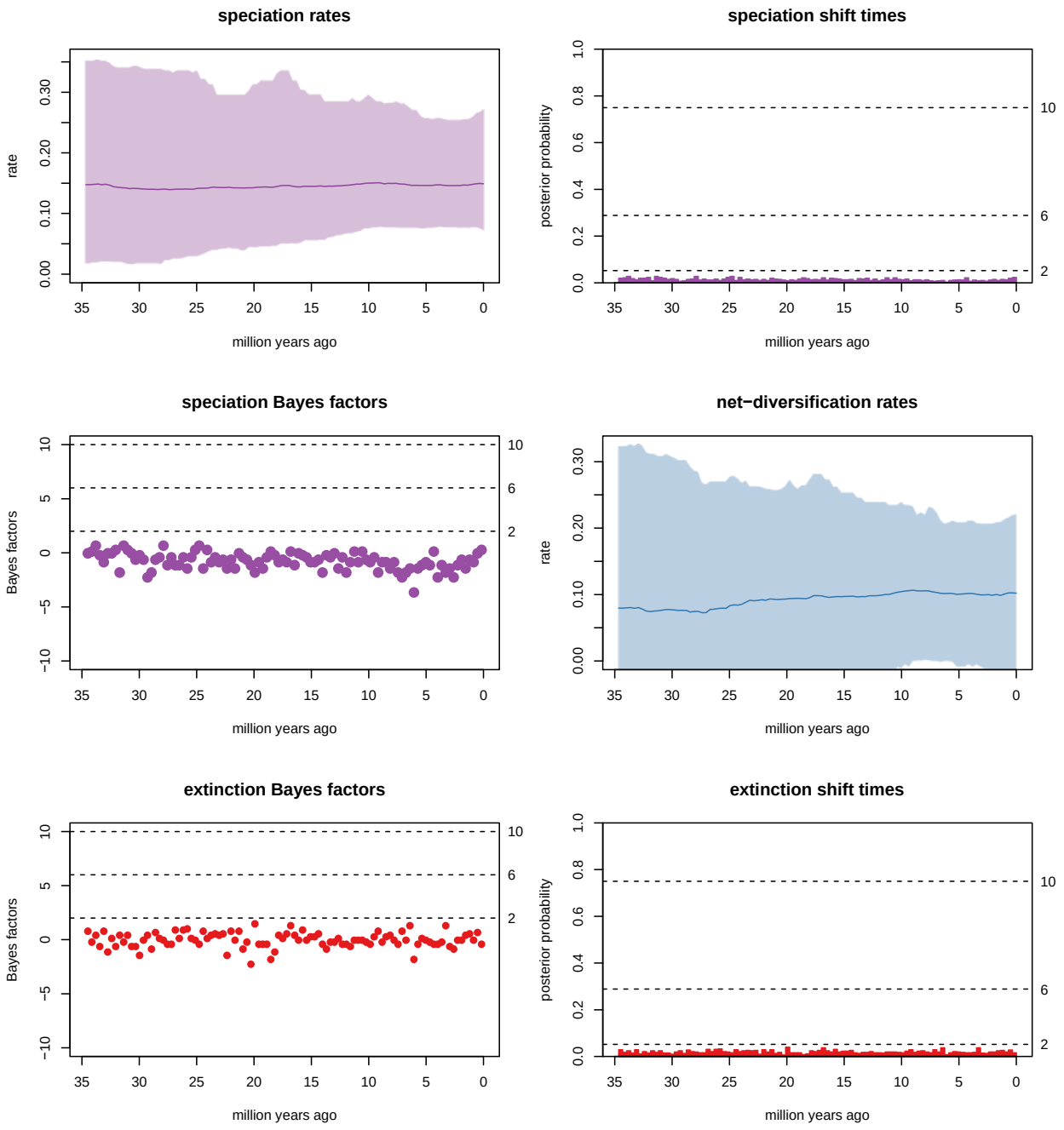


Fig. 1.4 Results of the TESS analysis for the genus *Pelophylax*

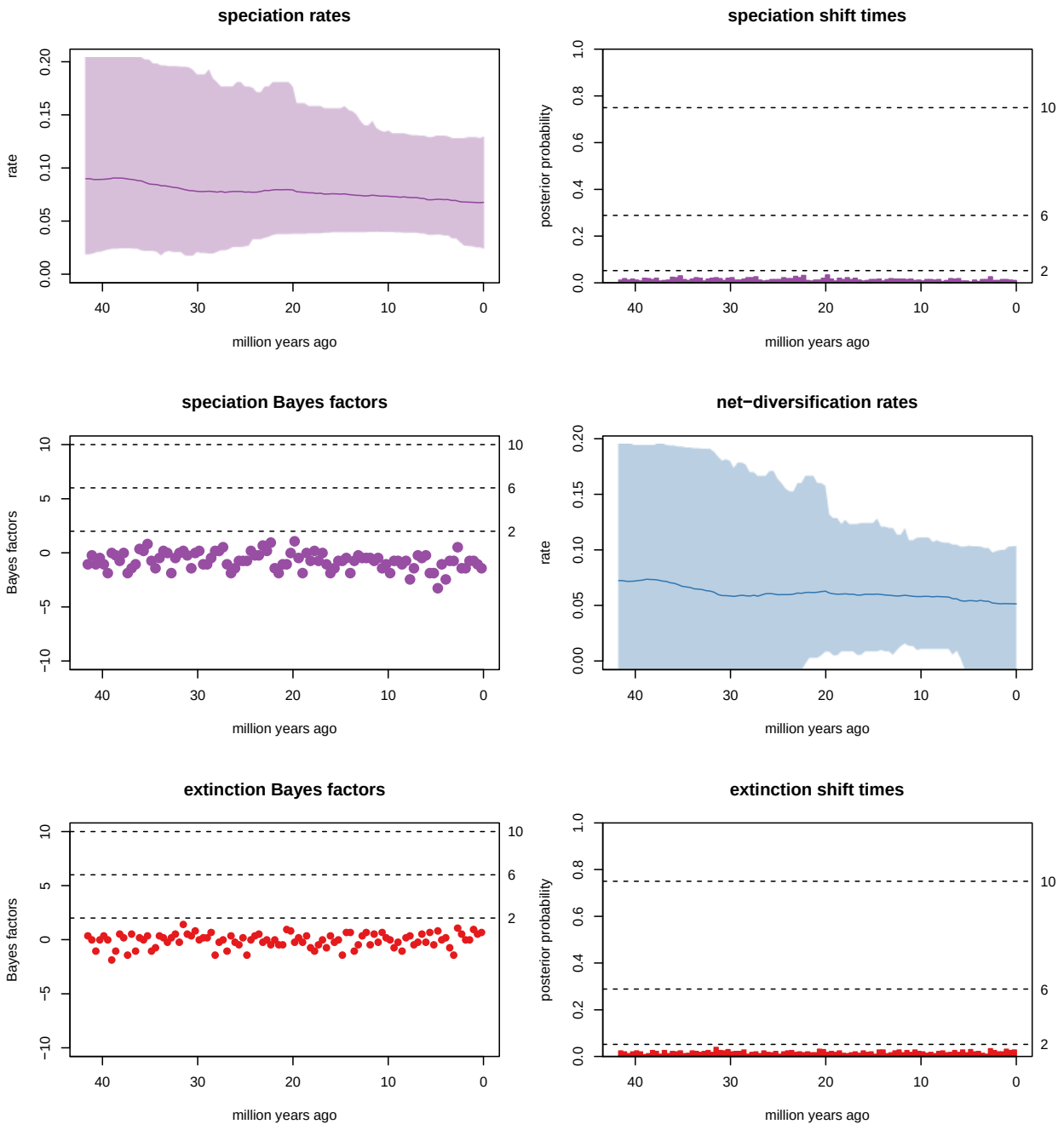


Fig. 1.5 Results of the TESS analysis for the combined clade of *Glandirana* + *Sanguirana*

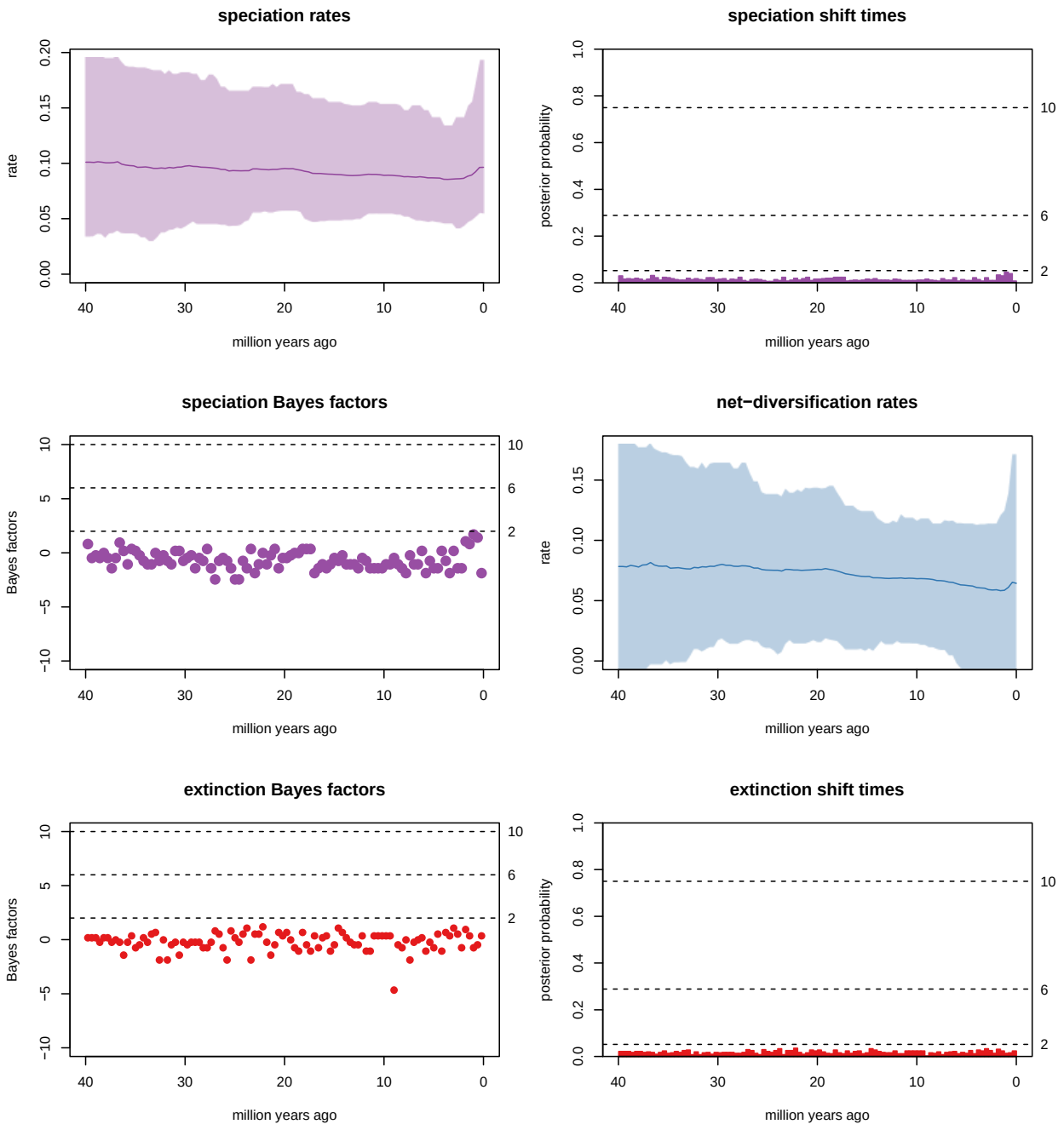


Fig. 1.6 Results of the TESS analysis for the combined clade of *Amnirana* + *Indosylvirana*

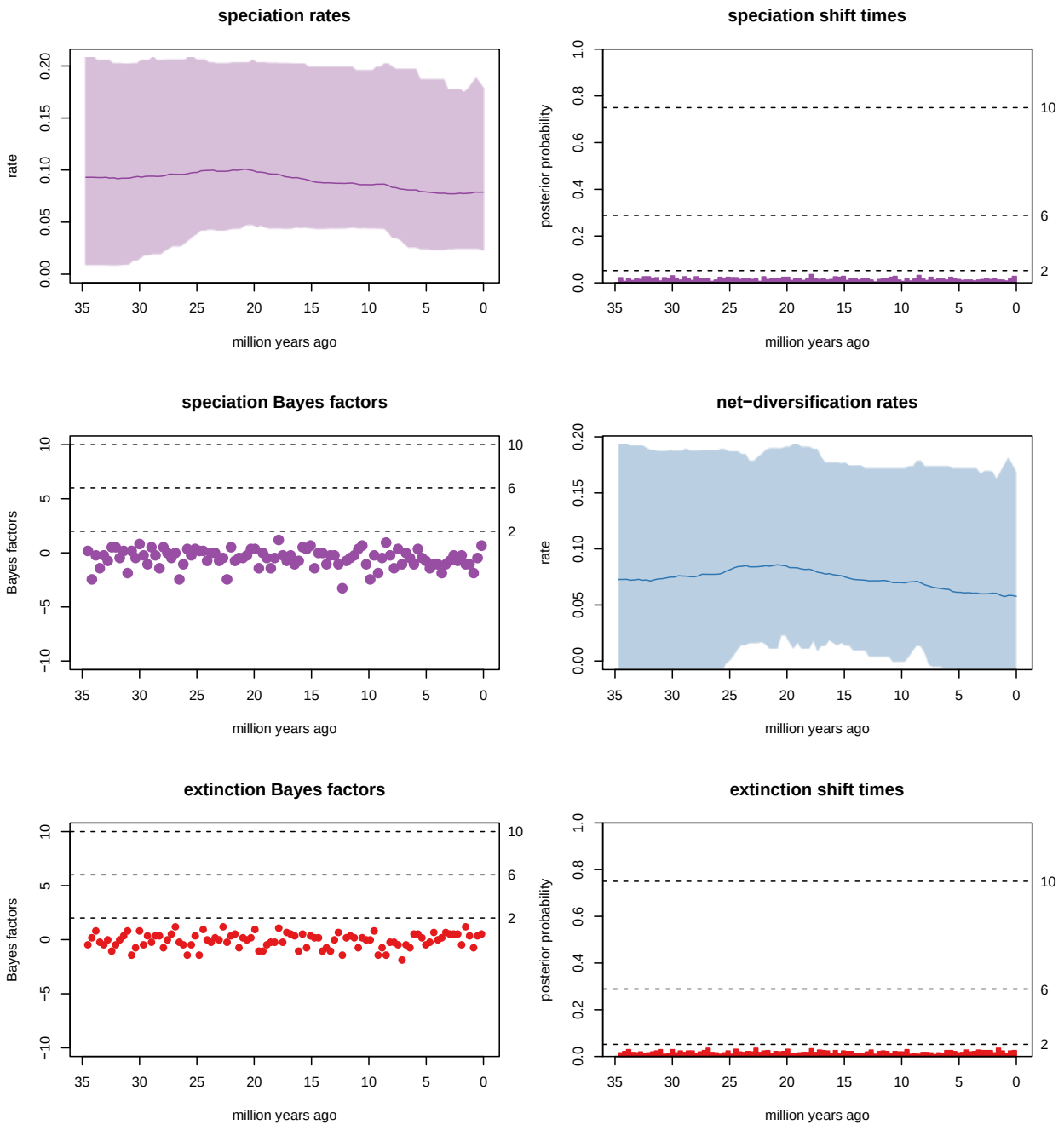


Fig. 1.7 Results of the TESS analysis for the combined clade of *Papurana + Hydrophylax*

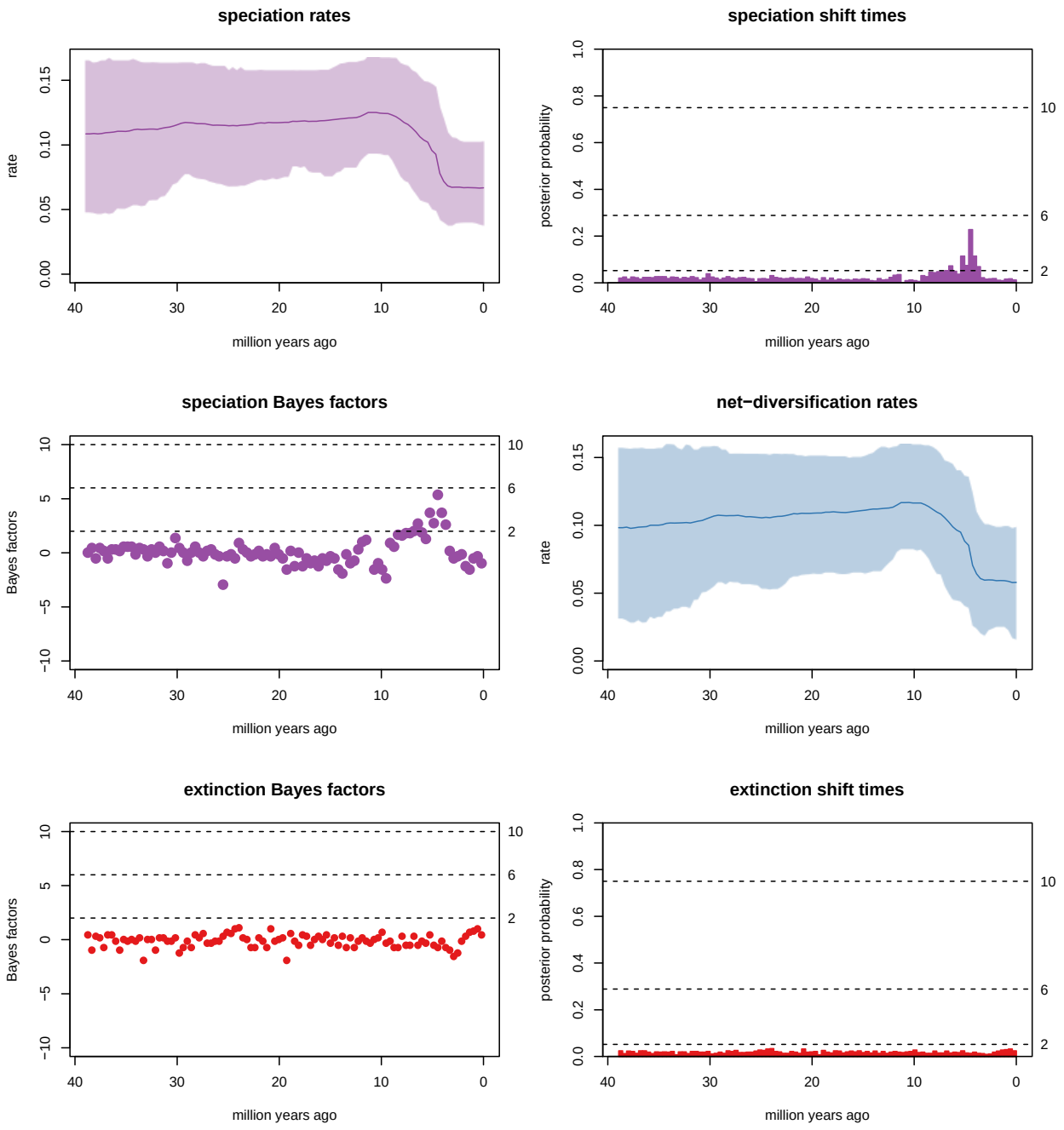


Fig. 1.8 Results of the TESS analysis for the genus *Rana*

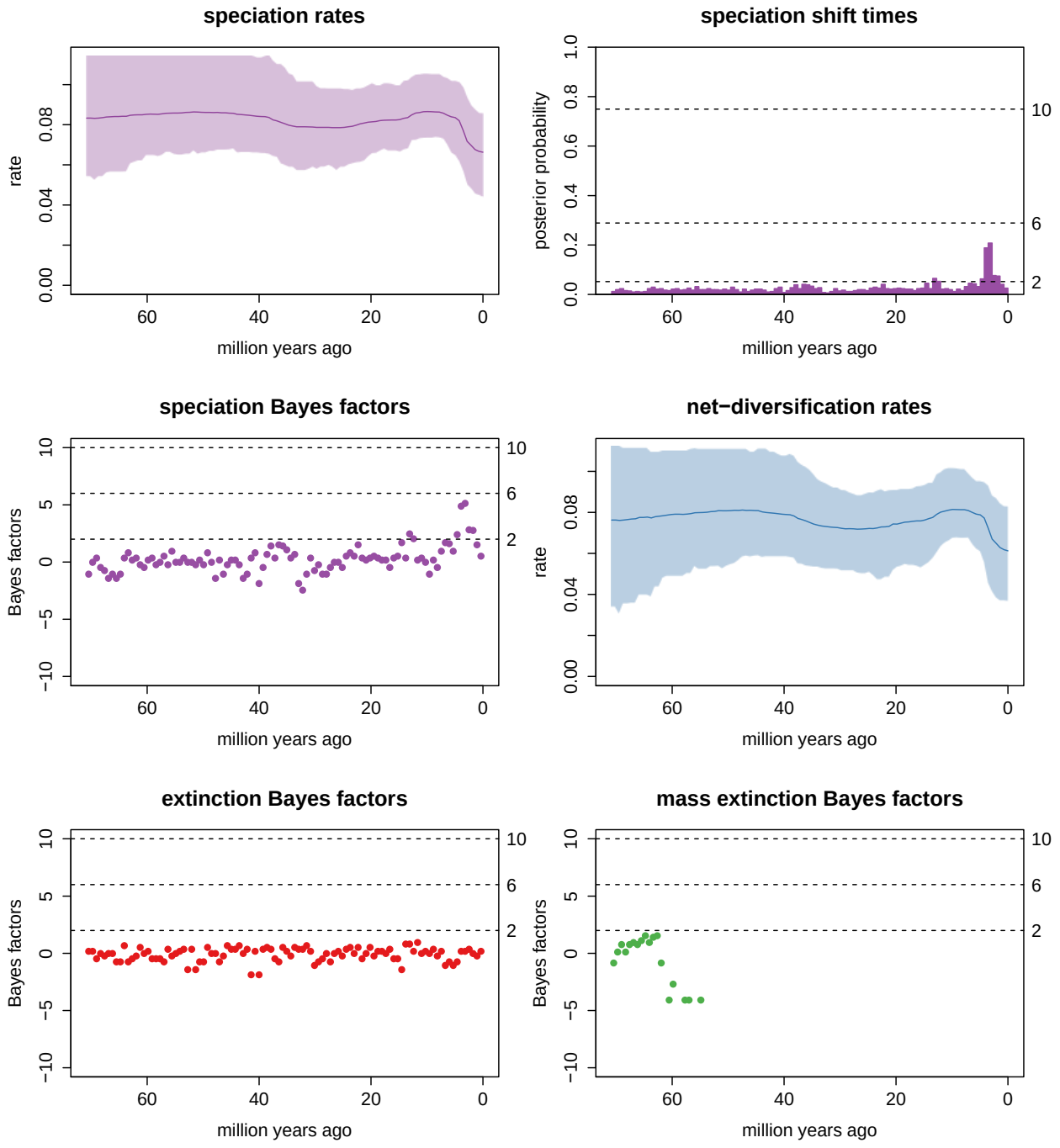


Fig. 1.9 Results of the TESS analyses performed on the family-wide phylogeny.

CHAPTER 2

Species delimitation with gene flow: a methodological comparison and population genomics approach to elucidate cryptic species boundaries in Malaysian Torrent Frogs

Chan, KO, Alana, MA, Grismer, LL, Su, YC, Grismer, JL, Quan, ESH, Brown, RM. 2017. Species delimitation with gene flow: a methodological comparison and population genomics approach to elucidate cryptic species boundaries in Malaysian Torrent Frogs. *Molecular Ecology* 2017;00:1–16. https://doi.org/10.1111/mec.14296

2.1 Abstract

Accurately delimiting species boundaries is a non-trivial undertaking that can have significant effects on downstream inferences. We compared the efficacy of commonly-used species delimitation methods (SDMs) and a population genomics approach based on genome-wide single nucleotide polymorphisms (SNPs) to assess lineage separation in the Malaysian Torrent Frog Complex currently recognized as a single species (*Amolops larutensis*). First, we used morphological, mitochondrial DNA and genome-wide SNPs to identify putative species boundaries by implementing non-coalescent and coalescent-based SDMs (mPTP, iBPP, BFD*). We then tested the validity of putative boundaries by estimating spatiotemporal gene flow (fastsimcoal2, ABBA-BABA) to assess the extent of genetic isolation among putative species.

Our results show that the *A. larutensis* complex runs the gamut of the speciation continuum from highly divergent, genetically isolated lineages (mean F_{st} = 0.9) to differentiating populations involving recent gene flow (mean F_{st} = 0.05; N_m > 5). As expected, SDMs were effective at delimiting divergent lineages in the absence of gene flow but overestimated species in the presence of marked population structure and gene flow. However, using a population genomics approach and the concept of species as separately evolving metapopulation lineages as the only necessary property of a species, we were able to objectively elucidate cryptic species boundaries in the presence of past and present gene flow. This study does not discount the utility of SDMs but highlights the danger of violating model assumptions and the importance of carefully considering methods that appropriately fit the diversification history of a particular system.
Keywords: Amolops, migration rate, fastsimcoal2, site frequency spectrum, gene flow, single-nucleotide polymorphism

2.2 Introduction

Delimiting species boundaries is a fundamental component of systematic biology that forms the framework for understanding the evolutionary processes that generate biodiversity (Mayr 1968). As such, accurately delimiting species boundaries is a non-trivial step that can have cascading ramifications (Veach *et al.* 2017). Species delimitation is usually performed using certain properties of a species as criteria for assessing lineage independence. The most commonly used criteria include phenotypic distinctiveness, molecular divergence and phylogenetic placement (Brown & Stuart 2012; Leavitt *et al.* 2015; Leliaert *et al.* 2014; de Queiroz 2007; Tobias *et al.* 2010; Wiens & Penkrot 2002). These “traditional” properties can be useful in delimiting allopatric (Chan *et al.* 2011), phenotypically distinct (Grismer *et al.* 2010), and genetically

distant lineages where barriers to gene flow or sufficient time has passed for fixed character differences to accumulate (Chan *et al.* 2016; Grismer *et al.* 2013). However, cryptic lineages that occur in sympatry, have similar niches, and are not readily distinguishable phenotypically such as those characterized by recent/rapid radiations, can be harder to diagnose because divergent lineages no longer connected by gene flow cannot be easily distinguished from the local population structure within such lineages, forming a hierarchy of genetic differentiation and divergence (Barley *et al.* 2013; Carstens *et al.* 2013; Rannala 2015; Sukumaran & Knowles 2017). In such cases, traditional criteria are limited in utility for assessing lineage separation (de Queiroz 2005) and if not implemented with caution, can lead to erroneous results (Carstens *et al.* 2013).

Advances in genomic sequencing and bioinformatics have led to the ability to detect population structure between closely-related populations at unprecedented resolution (Benestan *et al.* 2015; Candy *et al.* 2015; Larson *et al.* 2014; Leslie *et al.* 2015). One of the challenges with such datasets is to distinguish between structure that is associated with intraspecific variation from that resulting from speciation (Sukumaran & Knowles 2017). Model-based methods make simplifying assumptions about certain parameters (e.g. gene flow, population size) during the speciation process, and range in complexity from non-coalescent, sequence-based methods that model speciation in terms of number of substitutions (Zhang *et al.* 2013) to highly parameterized Bayesian models based on the multispecies coalescent (Yang & Rannala 2010) that allow for the integration of multiple data types into a single model-based framework (Solís-Lemus *et al.* 2015). The efficacy of each method depends on how well the model fits the data, and processes that violate model assumptions such as gene flow (Burbrink & Guiher 2015; Sousa & Hey 2013;

Streicher *et al.* 2014) or spatial autocorrelation (Meirmans 2012; Reeves & Richards 2011) can yield inaccurate species delimitations.

Implicit within most species definitions— that species are separately evolving metapopulation lineages (de Queiroz 2007; Simpson 1961; Wiley 1978)— is the expectation that populations within a metapopulation lineage are connected by gene flow, but remain distinct from other such lineages (Frost & Hillis 1990; Petit & Excoffier 2009; de Queiroz 2005). Levels of gene flow among populations are not only influenced by intrinsic traits (e.g., dispersal ability) but also extrinsic spatial and temporal processes that shape genetic patterns across a landscape (Richardson *et al.* 2016). If these processes are overlooked, inferences of lineage boundaries may fail to recognize historical population associations (Knowles & Carstens 2007) or may be unable to distinguish true discontinuities (i.e., lineage separation) from variation that occurs within a species as a result of other phenomena such as continuous geographical clines or isolation-by-distance (de Queiroz 2007; Medrano *et al.* 2014; Sexton *et al.* 2014). Moreover, standard phylogenetic estimation methods have been shown to produce highly supported, erroneous topologies when gene flow is present, thereby invalidating downstream inferences that are based on these phylogenies, such as the identification of terminal monophyletic groups (Reeves & Richards 2007; Rosenberg 2007).

Here, we compared a wide range of commonly-used species delimitation methods and a population genomics approach to elucidate cryptic species boundaries in an understudied Southeast Asian frog complex. The Southeast Asian Sundaland Biodiversity Hotspot harbors one of the highest concentrations of endemic plants and vertebrates on the planet (Mittermeier *et al.* 1998; Myers *et al.* 2000). Unfortunately, only 7.8% of Sundaland's original primary forest remains and some estimates suggest that up to 42% of its biodiversity could be lost by 2100

(Myers *et al.* 2000; Sodhi *et al.* 2004). Consequently, systematic research in this region has focused heavily on discovering and describing new species before they are lost. This is epitomized by the rapid surge of new amphibian and reptile descriptions over the last 15 years, resulting in more than a 20% increase in species richness (Brown & Stuart 2012; Giam *et al.* 2012; Grismer 2011). In virtually every one of these descriptions, species boundaries were delimited based on morphology and/or mitochondrial DNA (mtDNA) (e.g. Chan *et al.* 2014a; b; Chan & Grismer 2010; Grismer *et al.* 2012, 2014). Given the present Southeast Asian biodiversity crisis (Miettinen *et al.* 2011; Sodhi *et al.* 2004; Wilcove *et al.* 2013) and the need to rapidly inventory the region's diversity, it is important to tackle this problem using all available methods, including not only traditional morphology and mtDNA-based approaches but also species delimitation approaches that are better suited to elucidating cryptic lineage diversity. Such methods, which importantly can take gene flow into consideration, are best implemented using genomic-scale data — increasingly available even for non-model organisms. Torrent frogs of the genus *Amolops* are represented by 51 species that collectively range from Tibet, northeastern India, southern China, southward throughout Indochina and the Thai-Malay Peninsula (Frost 2015). The bulk of species diversity lies in southern China and Indochina, yet only one species, *A. larutensis*, occurs in extreme southern Thailand and Peninsular Malaysia. This species has never been studied outside of the context of higher-level phylogenetics where it was represented by samples from only two named localities (Hasan *et al.* 2014; Matsui *et al.* 2006; Stuart 2008) and as a result, little is known of its intraspecific phenotypic and genetic diversity.

Our sampling of *Amolops larutensis* from new localities throughout Peninsular Malaysia revealed subtle phenotypic variation among geographic populations, leading us to hypothesize

that *A. larutensis* constitutes a complex of cryptic lineages. Because no prior data were available, we used a two-step approach to species delimitation by applying widely-used species delimitation analyses (Carstens *et al.* 2013; Rannala 2015) using a variety of data types including morphology, mtDNA and genome-wide SNPs to form preliminary hypotheses of species boundaries. We then tested these putative boundaries using a rigorous population genomics framework. Specifically, we diagnose lineage separation by assessing spatiotemporal gene flow under the general concept of species as a separately evolving metapopulation lineage and treat this as the only necessary property of a species (de Queiroz 2007). Therefore, the objectives of this study are two-fold: (1) evaluate the efficacy of commonly-used species delimitation methods in assessing lineage separation in cryptic species; (2) determine whether population genomics can be an effective tool in elucidating cryptic species boundaries.

2.3 Methods

2.3.1 Sampling, data collection and accessibility

Our total data set consisted of 225 samples for which some combination of morphological, mtDNA and SNP data were available (Appendix Tables 2, 3). Morphological data were obtained from a subset of 141 vouchered museum specimens examined from collections at La Sierra University Herpetological Collection (LSUHC), Riverside, California; Zoological Reference Collection (ZRC), Lee Kong Chian Natural History Museum, Singapore; and University of Kansas Natural History Museum (KU), Lawrence, Kansas (Appendix Table 2).

DNA for mitochondrial and genomic sequencing was extracted from liver tissue using the Qiagen DNeasy Blood & Tissue Kit. A total of 117 samples (including 79 of the 141 samples scored for morphology) were sequenced for mtDNA and genome-wide SNPs (Appendix Table

3). These samples were chosen from populations that maximized geographic coverage and altitudinal variation across all major mountain ranges. For mtDNA, we sequenced the 16S rRNA-encoding gene using primers from Evans *et al.* (2003) and sequencing protocol from McLeod (2010). Raw sequence data were aligned using the MUSCLE algorithm and resulting alignments were subsequently refined by eye in Geneious Pro version 5.3 (Kearse *et al.* 2012). In addition to these samples, 49 16S rRNA *Amolops* sequences were obtained from GenBank to assess the monophyly and phylogenetic placement of Peninsular Malaysian populations. Samples and corresponding GenBank accession numbers are listed in Appendix Table 3.

A subset of 95 samples (including 67 samples scored for morphology, 18 with mtDNA data; Appendix Table 3) were selected for genomic sequencing of nuclear DNA in the form of genome-wide SNPs using a single-end multiplexed shotgun genotyping protocol (Andolfatto *et al.* 2011). Briefly, 500 ng of DNA from each sample were digested with *NdeI* (New England Biosystems), ligated with a sample-specific barcode and then pooled in sets of 48 samples and run through a Pippin Prep™ (Sage) to select fragments between 400–500 bp. Genomic samples were sequenced in one lane of the Illumina Hiseq 2500 platform at the Genome Sequencing Core Facility at the University of Kansas. Loci were subsequently assembled and filtered using the program pyRAD v.3.0.5 (Eaton 2014). The maximum number of low quality, undetermined sites ("N") in filtered sequences was set to 4 and proportion of shared polymorphic sites in a locus was set at 10% (Eaton 2014). Because overly stringent similarity thresholds have been shown to cause "over-splitting" by splitting orthologous reads into multiple loci (Catchen *et al.* 2013; Harvey *et al.* 2015; Ilut *et al.* 2014), whereas more liberal thresholds were found to have minimal bias effects on inference (Ilut *et al.* 2014; Rubin *et al.* 2012), we employed a relatively relaxed threshold of 88% similarity between reads when clustering loci. We then used two different

settings for minimum depth of coverage (min. read depth=5 and 10) and minimum taxon coverage (30% and 50% missing samples per locus) to produce four SNP datasets (Table 2.1). To avoid linkage across sites within the same locus, the single SNP with the highest sample coverage was selected from each locus.

2.3.2 Establishing putative species boundaries

Putative species boundaries were established using the following species delimitation framework based on traditional and widely used criteria:

1. We estimated mtDNA phylogenies and identified monophyletic lineages. Each monophyletic lineage that corresponded to a discrete or recognizable geographic region was defined as an Operational Taxonomic Unit (OTU).
2. We calculated uncorrected genetic distance between OTUs based on the mtDNA sequence alignment.
3. We assessed morphological variation and distinctiveness between OTUs using multivariate analyses.
4. Finally, given the support of geographic data, clades recovered with our concatenated SNP phylogenies and the sNMF populations structure assignments (methods and results discussed below) for the OTUs described using mtDNA, we then performed non-coalescent and coalescent-based species delimitation analyses to establish putative species boundaries for downstream hypothesis testing.

Phylogenetic estimation.—Bayesian and maximum likelihood (ML) methods were used to infer phylogenetic relationships from mtDNA. Bayesian inference was implemented in the program

MrBayes 3.2.6 (Ronquist *et al.* 2012) using a reversible jump MCMC + Γ substitution model and default priors. Four independent MCMC runs (10,000,000 generations and four chains per run) were combined and assessed for convergence using the program Tracer v1.6 (Rambaut *et al.* 2014) and a 50% majority rule consensus tree was produced by excluding the first 25% of sampled trees as burn-in. The program IQ-TREE (Nguyen *et al.* 2014) was used to perform ML analyses. The Bayesian Information Criterion (BIC) was used to select the most appropriate substitution model and branch support was assessed using 10,000 ultrafast bootstrap approximation replicates (Minh *et al.* 2013).

For SNP data, a ML phylogeny was also estimated using IQ-TREE. We applied an ascertainment bias correction using the ASC model (Lewis 2001) and branch support was assessed using 10,000 ultrafast bootstrap approximation replicates. Additionally, we estimated a species tree under the Bayesian multispecies coalescent framework implemented in the program SNAPP (Bryant *et al.* 2012) through BEAST v.2.3.1 (Drummond & Bouckaert 2015). We used the previously identified OTUs as *a priori* species assignments and the following parameter settings: mutation rates (*u* and *v*) and the shape parameter for the gamma distribution prior on population sizes (alpha) were set at 1.0; the beta scale parameter was set at 333 (calculated from the mean value of the total number of polymorphic sites); and the speciation rate prior (lambda) was sampled from a broad gamma distribution of alpha = 2 and beta = 200. The MCMC chain was run for 10,000,000 generations, sampling every 1,000 states, and stationarity was assessed in the program Tracer v1.6. The posterior distribution was considered adequately sampled when effective sample size (ESS) values for parameters were greater than 200.

Morphological variation and species delimitation.—Nine continuous morphological characters were measured from adult specimens: snout-vent-length (SVL), head length (HL), head width (HW), internarial distance (IND), snout length (SNL), forearm length (FAL), femur length (FL), tibia length (TBL), and third finger disc width (Fin3DW) following Chan *et al.* (2016). Due to pronounced sexual size dimorphism, male and female measurements were analyzed separately. Characters were adjusted for allometric growth using the following equation: $X_{adj}=X-\beta(SVL-SVL_{mean})$, where X_{adj}=adjusted value; X=measured value; β=unstandardized regression coefficient for each OTU; SVL=measured snout-vent-length; SVL_{mean} =overall average SVL of all OTUs (Lleonart *et al.* 2000; Thorpe 1975, 1983; Turan 1999). Adjusted variables were then log-transformed prior to downstream analyses. We performed a principal components analysis (PCA) on this adjusted morphological dataset to find the best low-dimensional representation of variation in the data. Components with eigenvalues above 1.0 were retained in accordance to Kaiser's criterion (Kaiser 1960). To further characterize population clustering, a discriminant analysis of principal components (DAPC) was performed to find the linear combinations of variables that have the largest between-group variance and the smallest within-group variance. The DAPC analysis relies on data transformation using PCA as a prior step to a discriminant analysis (DA), ensuring that variables submitted to the DA analysis are uncorrelated and that their number is less than that of analyzed individuals (Jombart *et al.* 2010)

Uncorrected pairwise p-distances were calculated from the mitochondrial sequence alignments using the program PAUP* (Swofford 2002). We then carried out species delimitation based on mtDNA using a method that has been shown to perform well with single-locus data (Tang *et al.* 2014). The multi-rate Poisson tree processes (mPTP) is a non-coalescent, maximum likelihood, sequence-based method that models speciation in terms of number of substitutions

(Zhang *et al.* 2013). This method identifies changes in the tempo of branching events, where the number of substitutions between species is assumed to be significantly higher than the number of substitutions within species. Additionally, the model incorporates different levels of intraspecific genetic diversity deriving from differences in either evolutionary history or sampling of each species (Kapli *et al.* 2017; Zhang *et al.* 2013). During phylogenetic inference, identical sequences are assigned very short non-zero branch lengths to retain the binary shape of the tree. Because this program requires a binary phylogeny, the 50% majority consensus tree estimated from the Bayesian analysis was used as the input tree. Confidence of the delimitation scheme was assessed using two independent MCMC chains at 5,000,000 generations each. Support values indicate the fraction of sampled delimitations in which a node was part of the speciation process.

We jointly analyzed morphological and mtDNA in a common coalescent Bayesian framework using the program iBPP. This method has been shown to improve the accuracy of species delimitation by integrating phenotypic and genetic data (Solís-Lemus *et al.* 2015). The iBPP analysis was performed with and without mtDNA data to maximize the signal derived from phenotypic variation and to evaluate the influence of mtDNA sequence data on this integrated analysis. Male and female datasets were analyzed separately to avoid biases from sexual dimorphism. We used three different combinations of priors for ancestral population size (□) and root age (t_0) drawn from a gamma distribution specified as G(□, β), where □ is the shape and β is the rate parameter. All other divergence time parameters were assigned the uniform Dirichlet prior (Fujita *et al.* 2012; Pyron *et al.* 2016; Yang & Rannala 2010). We chose a diffuse shape parameter (□ = 1 or 2) and parameterized β for large ancestral populations and deep divergences, □ ~ G(1, 10), t_0~ G(1, 10); small ancestral populations and shallow divergences □ ~ G(2, 2000),

t_0~ G(2, 2000); and large ancestral populations with shallow divergences □ ~ G(1, 10), t_0~ G(2, 2000). Both rjMCMC algorithms were implemented: Algorithm 0 with □ = 5; Algorithm 1 with □ = 2 and *m* = 1. Two independent runs were performed for each algorithm and prior combination with a chain length of 50,000 sampled every fifty generations, discarding 1,000 generations as burn-in. MCMC convergence was assumed when results were the same between multiple runs using the two algorithms (Yang 2015).

Species delimitation analysis on genomic SNPs was performed using the Bayes factor delimitation method (BFD*; Leaché *et al.* 2014). Different species delimitation models were constructed by lumping and splitting OTUs based on plausible biogeographic scenarios and phylogenetic topologies derived from prior phylogenetic analyses (Table 2). The marginal likelihood of each model was estimated via path sampling using 48 steps, an alpha of 0.3, and a MCMC chain length of 100,000 with a pre-burnin of 100,000 (Leaché *et al.* 2014). Natural log Bayes factors (BF) were used to compare the log marginal likelihoods (MLE) of competing models using the equation BF = 2[MLE(model1) – MLE(model2)], where model 1 was the model with the largest number of species. A positive BF value indicates support for model 1 and a negative value support for model 2.

2.3.3 Validation of putative species boundaries using genome-wide SNPs

Population structure and differentiation.—Population structure was characterized by estimating individual ancestry coefficients that represent the proportions of an individual genome that originate from multiple ancestral gene pools. Calculations were implemented in the program sNMF based on sparse non-negative matrix factorization and least-squares optimization (Frichot *et al.* 2014; Kim & Park 2007). Ancestry coefficients estimated using the sNMF method have

been shown to produce results that are comparable to other widely used programs such as ADMIXTURE and STRUCTURE, but have the advantage of estimating homozygote and heterozygote frequencies and avoiding Hardy-Weinberg equilibrium assumptions (Frichot *et al.* 2014). We calculated ancestry coefficients for 1–16 ancestral populations (K) using 100 replicates for each K. The preferred number of K was chosen using a cross entropy criterion based on the prediction of masked genotypes to evaluate the error of ancestry estimation. The sNMF method was implemented in the R package *LEA* (Frichot & François 2015).

To determine whether genetic structure was spatially autocorrelated, we conducted a Mantel test by examining the correlation between genetic distance and Euclidean geographic distance. Correlation values were compared against a distribution of permuted values based on 1,000 replicates simulated under the absence of spatial structure. The Mantel test was performed using the R package *adegenet* 2.0.1 (Jombart 2008).

Genetic distances between population pairs were estimated using Wright's F_{st} and Jost's *D* (Jost 2008; Meirmans & Hedrick 2011; Whitlock 2011; Wright 1951). Population differentiation was tested with Analysis of Molecular Variance, AMOVA (Excoffier *et al.* 1992) using the number of different alleles (F_{st}) based on the infinite allele model (Weir & Cockerham 1984), nesting individuals within populations, and populations within the eastern (East) versus central + western (West) mountain ranges. Significance was assessed using 1,000 permutations. These calculations were performed using the program GENODIVE v2.0b27 (Meirmans & Van Tienderen 2004).

Hybridization and demographic analyses.— Hybridization at the contact zone was investigated by calculating the hybrid index (Buerkle 2005). East 1 and Larutensis were selected as parental populations while West 1 was designated as the putative hybrid population.

Population connectivity was assessed by estimating the effective number of migrants exchanged between populations per generation (N_m) using fastsimcoal2 v.52.21 (Excoffier *et al.* 2013). Due to computational constraints, we only analyzed populations form the western clade. A folded site frequency spectrum (SFS) was obtained with custom *R*-code (Alexander 2017) and δaδi v1.7.0 (Gutenkunst *et al.* 2009), projecting down population sizes to maximize the number of segregating sites using custom *R*-code (Alexander 2017). Four scenarios were examined: contemporary and historical migration, contemporary migration only, historical migration only, no migration (Appendix Figs. 9–12; Appendix Table 6). Population divergence times followed Chan & Brown (2017), with the exception of the timing of the divergence of West 2 from West 3/West 4, which was set as halfway between the coalescence of all populations and the divergence of West 3/West 4 (as the SNP topology differs from the mtDNA for this lineage). For each scenario, 50 replicate fastsimcoal2 runs were carried out with the following settings: -n 100,000 -N 100,000 -m –multiSFS -q -M 0.001 -l 10 -L 40. Initial prior distributions followed a uniform distribution based on the population-specific theta (distribution range: one order of magnitude lower and higher than theta estimate) estimated using the program GENEPOP (Rousset 2008), and the data were modeled as FREQ, with the number of independent chromosomes equal to the number of non-monomorphic SNPs in the SFS (4018). We used a mutation rate of 1.91×10^{-8} following Barker et al. (2011), and assumed vicariant splits between populations (i.e. the number of simulated individuals remained constant through time). The range in parameter estimates for the initial fifty runs were used as the prior distributions for the next run, and this

process continued until no further increase in likelihood was detected. Using the parameter values from the run with the highest likelihood, an additional run with -n/-N = 1,000,000 was carried out to more accurately estimate the likelihood. The best fitting scenario was then assessed by Akaike's information criterion (AIC) score (Akaike 1974). The parameter estimates for the best-fitting scenario were used to simulate 100 parametric bootstraps of the SFS. The data-type was changed to DNA, with a mutation rate of 1.91×10^{-8}, and the number of chromosomal segments equaling the total number of sites in the SFS (including monomorphic sites: 5695). The length of the chromosomal segments was set at 100 bp, and the mutation rate (to three significant figures) was adjusted by trial and error until the closest match to the number of non-monomorphic in the observed SFS was obtained. After the bootstrap replicates were generated, the *.tpl and *.est files that led to the run with the highest likelihood in the initial screening runs of the best scenario were then used with the bootstrap replicates to obtain confidence intervals for the parameter estimates, discarding the 2.5% lowest and highest estimates for each parameter.

To differentiate between introgression and incomplete lineage sorting, we used Patterson's D-statistic (ABBA-BABA test), based on the frequencies of discordant SNP genealogies in a pectinate four-taxon tree [(((P1,P2),P3),O)]. This test assumes that two SNP patterns, "ABBA" and "BABA" should be equally frequent under a scenario of incomplete lineage sorting without gene flow, where "A" denotes the ancestral allele and "B", the derived allele. An excess of ABBA or BABA patterns would therefore be indicative of introgression (Durand *et al.* 2011; Patterson *et al.* 2012). We calculated *D* over combinations of four taxa that fitted the four-taxon tree configuration across all plausible topologies inferred from our phylogenetic analyses. Population pairs that did not conform to any plausible relationships were not included in the test. Four samples from each population were randomly chosen to form taxon

sets. Ingroup taxa (P1–P3) were then iterated over all possible combinations of individuals that were chosen, while samples were pooled into groups for the outgroup population (O). This approach allows the use of any locus shared by the three sampled ingroup taxa and at least one outgroup, effectively down-weighting *D* if the ancestral allele was not fixed across multiple outgroup samples, thus making it a more conservative test. The standard deviation of *D* was calculated from 200 bootstrap replicates and the observed *D* was converted to a Z-score measuring the number of standard deviations it deviated from 0. Significance was assessed using a *P*-value at $\alpha = 0.01$ after the Holm-Bonferroni correction for multiple testing (number of possible combinations fitting the given species tree hypothesis; Eaton & Ree 2013; Eaton *et al.* 2015). The D-statistic test was implemented in PyRAD.

2.4 Results

2.4.1 Phylogenetic relationships

Mitochondrial DNA.—Both Bayesian and ML phylogenetic analyses on mtDNA produced congruent topologies at most major nodes, inferred Peninsular Malaysian *Amolops* as a monophyletic clade sister to *A. cremnobatus* from Indochina, and had identical topologies within the Peninsular Malaysian subclade (Appendix Fig. 5). Within the Peninsular Malaysian subclade, all nodes were highly supported in the ML tree (bootstrap > 90%, Fig. 2.1A) whereas in the Bayesian tree, one node received relatively low support (posterior probability 0.4; Appendix Fig. 5). Two highly divergent (14–16% p-distance, 16S; Appendix Fig. 6), reciprocally monophyletic Peninsular Malaysian *Amolops* clades were recovered with high support by both methods. These corresponded to populations from the eastern vs. western + central mountain ranges (hereafter referred to simply as eastern and western clades; Fig. 2.1). In the eastern clade, we defined two

genetically distinct and reciprocally monophyletic OTU's (separated by 7–8% p-distance, 16S) that corresponded to populations from the northeastern mountain range (East 1) and southeastern mountain range (East 2). In the western clade, five subclades were recovered (1–5% p-distance, 16S; Appendix Fig. 6). We designated these OTU's as Larutensis (type locality of *A. larutensis*), West 1, West 2, West 3 and West 4 (Fig. 2.1).

Genome-wide SNPs.—After quality control filtering of the initial 153 million reads obtained across all samples, a total of ~130 million reads were retained. The total number of unlinked SNPs in the final datasets that were used for downstream analyses ranged from 4,744 to 17,123 (Table 2.1).

Maximum likelihood analyses on concatenated SNP datasets led to four different phylogenies depending on how loci were filtered (Appendix Fig. 7). At a minimum depth of 5 and 50% missing data, all major splits were highly supported, however a topology differing from the mtDNA tree was produced (Fig. 2.1B). The SNP dataset at a minimum depth of 5 and less missing data (30%) inferred a similar phylogeny, albeit with low support for the relationships among populations within the western clade (Appendix Fig. 7). Phylogenies constructed from the SNP datasets with a minimum depth of 10 failed to recover the West 2, West 3 and West 4 populations as monophyletic groups. Furthermore, support for deeper nodes was significantly lower. Topological placement of the East 1 and East 2 populations were congruent and highly supported throughout all phylogenetic analyses and datasets, including mtDNA. One sample from the contact zone (denoted by an asterisk in Fig. 2.1) was embedded within West 1 in the mtDNA phylogeny but recovered as a distinct lineage within the eastern clade across all SNP phylogenies, indicating a putative hybrid. The SNAPP analysis failed to converge when

including populations from both eastern and western clades. Since the relationships of populations in the eastern clade were highly supported in all other analyses, we performed a separate analysis on a dataset that only included populations from the western clade. This analysis converged and produced a maximum clade credibility tree topology similar to the concatenated SNP ML phylogeny (min. depth=5, 50% missing data) with 1.0 posterior probability at each node (not shown).

2.4.2 Morphological variation and putative species boundaries

In both the male and female morphological datasets, the first three principal components (PCs) had eigenvalues above 1.0 and were retained for subsequent analyses. These PCs captured 69% (males) and 78% (females) of the total variance (Appendix Table 2). In males, East 2 showed some separation along the first and third (but not second) axes, whereas in females, the East 2 formed a distinct, non-overlapping cluster along the first axis but was undifferentiated along the second and third axes (Fig. 2.2). When variances between OTUs were maximized, the DAPC analysis also isolated the East 2 as a distinct cluster in both males and females but showed no separation for the other populations. No clear separation was detected in either sex across the other populations in the PCA or DAPC analysis (Figs. 2.2, Appendix Fig. 8).

A total of five species were delimited using the mPTP species delimitation method. East 1, East 2, West 3 and West 4 were delimited as separate species with maximum average support values of 1.0, whereas Larutensis, West 1 and West 2 received low support (0.003), suggesting that these OTUs should be lumped as a single species (Fig. 2.2).

All independent iBPP runs under both rjMCMC algorithms and all combinations of priors produced the same results, indicating convergence (Yang 2015). Species delimitation results

were similar regardless of whether sequence data were included or excluded in the analyses. In the male dataset (with sequences included), all OTUs were highly supported as distinct species (pp = 1.0). For the female dataset, all OTUs were supported as distinct species (posterior probability = 1.0) with the exception of the split between West 1 and Larutensis, which was moderately supported (pp = 0.7, Fig. 2.2).

Due to the previous lack of convergence in SNAPP analyses including both western and eastern clade individuals, and because relationships were unambiguous for the eastern clade, we restricted the BFD* analysis to populations from the western clade only. Marginal likelihood estimates improved as the number of species increased and favored the model that defined each population as a distinct species. The second-ranked model favored four species by lumping West 3 and West 4 as a single species. However, when compared to the five-species model, the BF value was high (+3978), indicating strong support for the five-species model (Table 2.2).

2.4.3 Validation of putative species boundaries

Population structure and differentiation.—The population structure analysis (sNMF) on both SNP datasets at a minimum depth of five inferred similar patterns of population structure and admixture, where K=2 split individuals into eastern and western clusters. For the dataset with 50% missing data, K=7 had the lowest cross-entropy value (Fig. 2.3), whereas the dataset with 30% missing data inferred K=6 as the preferred number of genetic clusters. Signatures of admixture were detected among populations from the western clade. At the contact zone, the putative hybrid sample appeared admixed between East 1 and West 1 genotypes. Apart from the putative hybrid (further investigated below), no further admixture was detected between the eastern and western clades. This analysis also inferred an additional population at the central

region of the eastern mountain range (the southernmost East 1 population in Fig 2.3). We refer to this subpopulation as East 1.2 in subsequent analyses, to distinguish it from the north East 1.1 subpopulation.

Within the western clade, F_{st} and Jost's *D* values based on SNP data were low among populations, ranging from 0.03–0.09 (mean=0.053) and 0.001–0.003 (mean=0.002) respectively. These values were higher among populations within the eastern clade, ranging from 0.57–0.93 (mean=0.7) and 0.02–0.14 (mean=0.08) for F_{st} and Jost's *D* respectively. Similarly, F_{st} and Jost's *D* values were also high when western and eastern populations were compared with each other (Table 2.3). Results of the AMOVA analyses on populations from the western clade showed that most of the variation (74%) occurred within individuals whereas in the eastern clade, most of the variation (65%) occurred among populations. When populations were nested within the western and eastern clades, most of the variation (53%) was attributed to the eastern versus western groupings (Table 2.4).

Hybridization and demographic analyses.—The hybrid index analysis showed that one out of seven samples (sample ID 21011, previously identified as a putative hybrid above) within the West 1 population was a hybrid between Larutensis and the combined East 1 parental populations (h = 0.549). The other six samples had h-values close to zero, indicating a strong affinity with Larutensis and that it was very unlikely they were of hybrid origin (Appendix Table 5).

For the fastsimcoal2 analysis, the full migration model was the best fit according to AIC. We examined a version of this model where migration rates between populations were constrained to be symmetrical, but it had a poorer fit to the data than the full migration model

that allowed for asymmetrical migration rates (Appendix Table 6). We therefore restrict our discussion to the full asymmetrical migration model only. Contemporary migration rates between Larutensis and all other populations were low (Nm = 0.2–0.6), suggesting reproductive isolation. Gene flow was highest between West 1 and West 2 (Nm = 5.5) and West 3 and West 4 (Nm = 5.8), whereas relatively low levels of gene flow were detected between West 2 and West 3 (Nm = 1.0; Fig. 2.4, Table 2.3). However, it should be pointed out that the confidence intervals of all point estimates associated with this model were wide (Appendix Table 6), suggesting denser sampling of the genome would be needed to accurately estimate parameters of this parameter-rich model. Among the historical migration rates, an outlier was the very high rate of migration from West 2 into the ancestor of the Larutensis/West 1 populations (Appendix Table 6). This high inferred gene flow could explain the discrepancy between the SNP and mtDNA phylogenies, with the sister relationship of West 2 and Larutensis/West 1 in the latter due to this introgression event.

The D-statistic was used to differentiate between introgression and incomplete lineage sorting among adjacent populations within the western clade, eastern clade, and between both western and eastern clades (Fig. 2.4, Table 2.3). Within the western clade, high levels of introgression (significant for all 63/63 combinations) were detected between West 1 and West 2, while low levels of introgression were detected between West 2 and West 3 (significant for 28/63 combinations). No introgression was detected between Larutensis and West 1 or Larutensis and West 2. These results are congruent with estimates from the fastsimcoal2 analysis (Table 2.3). Introgression between the sister lineages Larutensis–West 1 and West 3–West 4 were not assessed because the D-statistic is unable to test for gene flow between sister lineages P1 and P2 in a pectinate four-taxon tree [(((P1,P2),P3),O)].

Within the eastern clade, low levels of introgression were detected between East 1.1 and East 1.2 (significant in 8/23 combinations) whereas introgression was not detected between East 1.2 and East 2. Introgression was also absent among adjacent populations from the western and eastern clade, even between syntopic populations at the contact zone, excluding the hybrid sample (Table 2.3; Fig. 2.4).

Spatial autocorrelation was not detected when the Mantel test was performed on the entire SNP dataset (p=0.242), but was significant when the test was performed separately on the eastern (p=0.014) and western (p=0.009) clades (Appendix Fig. 13). Although spatial autocorrelation can result in a correlation of genetic and geographic distances, distant and divergent populations can also result in such a pattern. To distinguish between these two scenarios, we used a non-parametric approach by plotting both genetic and geographic distances and using two-dimensional kernel density estimation (KDE) to measure local densities. Continuous genetic clines such as those caused by spatial autocorrelation would result in a single cloud of points without discontinuities, whereas distant and divergent populations would be represented by separate high density patches. The KDE plots show that the western clade consists mostly of a single cloud, with a few outliers (samples from the Larutensis populations located on a different mountain range). The eastern clade was represented by two distinct patches (Appendix Fig. 13), indicating that the East 1 and East 2 populations are both distant and divergent and are not spatially autocorrelated.

2.5 Discussion

Our results show that commonly-used species delimitation methods were effective at assessing lineage separation in highly divergent lineages where gene flow was absent (East 1 and East 2)

but overestimated the number of species in younger lineages where gene flow was prevalent but populations were markedly structured genetically. “Splitting” of lineages within a metapopulation occurred even when genomic data were used. We attribute this to the violation of the underlying assumptions of the models implemented by these programs: the guide tree is assumed to be correct (Zhang *et al.* 2013); speciation is modeled as an instantaneous event (Nee 2006; Sukumaran & Knowles 2017); and divergence is assumed to occur without gene flow (Yang & Rannala 2010). Using these methods on a system that violated these assumptions led to model misspecification and inaccurate estimation of species boundaries (Camargo *et al.* 2012; Carstens *et al.* 2013; Ence & Carstens 2011; Jackson *et al.* 2016; Sukumaran & Knowles 2017). On the other hand, we showed that a population genomics approach can be an effective tool at delimiting species boundaries both when gene flow is absent and when it is present at varying levels. By considering lineage independence as the only necessary property of a species, we can shift our focus away from traditional criteria (phenotypic distinctness, monophyly, genetic divergence, etc.) and re-cast the species delimitation framework as one that strictly focuses on assessing lineage cohesion/separation. Using this approach, we demonstrate that Peninsular Malaysian *Amolops* are comprised of at least three species, the true *A. larutensis* (i.e. the western clade) and two unnamed lineages from the eastern clade (East 1 and East 2).

2.5.1 Support for lineage separation

All analyses unanimously supported at least three separately evolving lineages. The western and eastern lineages were separated by very large uncorrected mitochondrial distances and F_{st} values. The differentiation between these two lineages was also supported by the majority of AMOVA variance being explained by these lineages as opposed to populations within these lineages.

Furthermore, the D-statistic test showed no evidence of introgression between the western and eastern lineages (with the exception of a single hybrid sample discussed below). Similar results were obtained when comparing the populations East 1 and East 2 within the eastern lineage, thereby supporting the divergence and isolation of these two species.

Despite the presence of a single hybrid sample, our analyses indicated that all other samples from the eastern/western contact zone (n=21) consisted of either eastern or western genotypes with no genetic intermediates. We view this as evidence of strong reproductive isolation and hypothesize that hybridization events between these separately evolving lineages are rare and produce hybrids of low fitness that do not subsequently reproduce successfully. However, denser sampling will be required to better understand the extent and viability of hybrids at this contact zone.

These multiple lines of congruent evidence from different sources of data provides strong support for the recognition of at least three distinct species of *Amolops* in Peninsular Malaysia: the true *A. larutensis*, consisting of populations from the western lineage; and two undescribed species represented by the lineages East 1 and East 2.

2.5.2 Support for population cohesion

Within the highly structured western clade, the relatively high mitochondrial distances were consistent with inter-specific distances among other amphibian species (Fouquet *et al.* 2007; Vences *et al.* 2005a; b) and were identified as separate species based on traditional species delimitation methods. However, we reject this hypothesis based on results from population genomic analyses (sNMF, F_{st}, AMOVA, H-index, fastsimcoal2, D-statistic, Mantel test) that detected different levels of gene flow among these populations (discussed in further detail

below). Disturbingly, the populations that had the highest mitochondrial divergences (West 1, West 2, West 3 and West 4) were also the populations that were most undifferentiated and showed the highest levels of gene flow based on genomic data, potentially as a result of sex-biased gene flow. Genetic variation within the western clade is more reflective of intra-specific population structure than divergence associated with speciation events. As such, we consider the entire western lineage to be a single, cohesive metapopulation lineage represented by the taxon name *Amolops larutensis*.

Within the eastern clade, the D-statistic showed low levels of gene flow between the subpopulations East 1.1 and East 1.2 but not between East 1.2 and East 2. We attribute the low levels of gene flow and migration between the subpopulations East 1.1 and East 1.2 to the lack of samples from the region spanning those populations. We hypothesize that as samples from that area become available, populations from the entire northeastern mountain range will form a cohesive metapopulation lineage (East 1), separate from the southeastern population East 2 due to the lack of contiguous habitat between East 1 and East 2.

2.5.3 Biogeography and the speciation continuum

The different levels of genetic differentiation within Peninsular Malaysian *Amolops* illustrates the complex nature of speciation, ranging from the presence of continuous variation within a group without reproductive isolation, to complete and irreversible reproductive isolation between groups (Hendry *et al.* 2009). Deep divergence coupled with strong reproductive isolation could be caused by divergent selection (McKinnon *et al.* 2004; Rundle & Nosil 2005; Schluter 2009) or allopatric speciation (Coyne & Orr 2004; Wiley 1978). In this study, ecological conservatism

in *Amolops* and the discontinuous genetic variation between western and eastern lineages are more indicative of the latter as opposed to the former. Additionally,

At the other end of the divergence spectrum, populations from the western lineage were highly structured and showed varying levels of historical and contemporary migration consistent with a complex history involving gene flow between recently diverging lineages. Because speciation with gene flow can occur in nature (Niemiller *et al.* 2008; Nosil 2008; Zarza *et al.* 2016), we applied a migration threshold for genetic isolation of 1 individual per 10 generations as a cut-off to determine the level of gene flow below which we consider populations to be separately evolving lineages (Rannala 2015; Zhang *et al.* 2011). Using this threshold, gene flow among populations of the western clade has not been sufficiently reduced to be considered genetically isolated enough to represent distinct species. However, it is worth noting that gene flow between Larutensis on the northwestern mountain range and the geographically proximate W1 and W2 populations on the central range were the most reduced (N_m = 0.3). We interpret this as an indication of incipient speciation triggered by recent and rapid human development and the disruption of habitat corridors along the Bintang-Kledang range, a small mountain range situated between the northwestern and central mountain ranges (Jamaluddin *et al.* 2011; Khoo & Lubis 2005). Given sufficient time, Larutensis's lower long-term migration rates could lead this population to qualify as a species separate from the other western populations. However at this point, the data does not support this split and we therefore consider these populations as belonging to a single species.

2.5.4 Effects of genomic filtering parameters

Filtering parameters for SNP assembly can have a significant impact on downstream analyses and inferences. The correlation between including sites with more missing taxa and better bipartition support is consistent with previous simulation (Huang & Knowles 2016) and empirical studies (Eaton & Ree 2013; Wagner *et al.* 2013). Conversely, allowing large amounts of missing data can also result in high bootstrap support for incorrect clades (Leaché *et al.* 2015). Since a consensus has yet to be reached on the best ways to process large datasets, our preference for the dataset with lower minimum read depth and higher allowance for missing data should be interpreted with caution, especially as our preferred phylogeny might not actually represent the true species tree. However, our use of population genomic methods to delimit the number of Peninsular Malaysian *Amolops* species means our conclusions are relatively robust to errors in reconstructing the true topology of lineages included in this study.

In a separate study on the trade-off between coverage depth and the number of individuals in a sample, Buerkle & Gompert (2013) showed that low coverage sequencing (as low as 1X coverage) is not only sufficient, but also could be optimal to accurately estimate population parameters, as this allows the inclusion of greater numbers of individuals or sites in the genome. Lower coverage was also optimal for phylogenetic estimation in our datasets, as our higher minimum depth datasets had low branch support, and failed to recover some monophyletic groups. However, these recommendations are dependent on the overall level of sequencing depth in our project: our study supports previous research in that general rules of thumb for SNP filtering are unlikely but instead may depend on the properties of the dataset and species biology, which should be evaluated on a case-by-case basis (Huang & Knowles 2014; Leaché *et al.* 2015).

2.6 Conclusions

This study does not discount the utility of traditional species delimitation methods but instead highlights the importance of choosing the right tool for the right task. Using methods that do not account for gene flow to delimit cryptic species boundaries where gene flow occurs will inherently yield erroneous results. We therefore caution against using these methods to delimit recent and rapidly diverging populations where gene flow may be prevalent. For such cases, we demonstrate that a population genomics approach can be used to objectively assess lineage separation in line with the general lineage concept of species.

Our findings are especially significant for systematic research in regions where new species are being described at a high rate. Malaysia stands as a particularly relevant test case (i.e., a potential future study system for evaluating the performance of species delimitation procedures) in that numerous newly-described species, co-distributed throughout the range of localities studied here, have been split into multiple, formally named species using traditional species delimitation methods. This study does not invalidate those descriptions but provides evidence that gene flow is present among co-occurring populations in one taxonomic group (*Amolops*). Our findings suggest that other co-distributed and taxonomically diverse taxa could provide compelling examples for future genomic species delimitation studies.

2.7 Tables

Table 2.1 Summaries of the four different SNP datasets generated using different values for minimum read depth coverage and percentage of missing data. PIS = Parsimony Informative sites. Min. # and Max. # loci give the minimum and maximum number of loci observed within an individual for each dataset.

Min. Depth	% missing	Total variable sites	Total PIS	Min. # loci	Max. # loci	Unlinked SNPs
5	50	94,313	70,833	1,761	17,831	17,123
5	30	53,208	43,191	1,572	7,572	7,544
10	50	65,541	49,821	112	12,478	11,951
10	30	32,253	26,122	97	4,826	4,744

Table 2.2 Results of the BFD* analysis based on nine species delimitation models ranging from 2–5 species within the western clade. Models were split or lumped according to plausible biogeographic scenarios and phylogenetic topologies. Competing models were compared and ranked using log marginal likelihood estimates (MLE) and Bayes factors (BF) following the equation: BF = 2 x (MLE of model 1 – MLE of model 2), where model 1 was the model with five species. A positive BF value indicates support for model 1 over model 2 and vice versa. Model abbreviations are L=Larutensis, W=West.

# Species	Model	MLE	BF	Rank
5	(L) (W1) (W2) (W3) (W4)	-99906.30527	-	1
4	(L) (W1) (W2) (W3 + W4)	-103884.8744	3978.57	2
4	(L +W1) (W2) (W3) (W4)	-107029.7769	7123.47	3
3	(L) (W1) (W2 + W3 + W4)	-107736.139	7829.83	4
3	(L) (W1 +W2) (W3+W4)	-108921.1603	9014.86	5
3	(L + W1) (W2) (W3 + W4)	-111397.7147	11491.41	6
2	(L) (W1 + W2 + W3 + W4)	-113053.1408	13146.84	7
2	(L + W1) (W2 + W3 + W4)	-115914.2886	16007.98	8
2	(L + W1 + W2) (W3 + W4)	-116802.3151	16896.01	9

Table 2.3 Pairwise comparisons of demographic parameters within and between eastern and western populations. Examined parameters include genetic distance (F_{st} and Jost's *D*), migration rates (N_m) and D-statistic scores represented by the range of Z-scores followed by the number of significant location comparisons assessed using a *P*-value at $\alpha = 0.01$ after the Holm-Bonferroni correction. NT= not tested.

Populations		Genetic distances		Migration rates	D-statistic
Pop 1	**Pop 2**	**Fst**	**D**	**(Nm)**	**Z-range (nSig.)**
Within West					
Larutensis	West 1	0.0860	0.0030	0.5814	NT
Larutensis	West 2	0.0670	0.0020	0.3331	0.0 - 3.0 (0/63)
West 1	West 2	0.0630	0.0020	5.4939	2.8 - 7.3 (63/63)
West 2	West 3	0.0290	0.0010	1.0291	0.9 - 5.6 (28/63)
West 3	West 4	0.0340	0.0010	5.7762	NT
	Mean	**0.0530**	**0.0017**		
Within East					
East 1	East 1.2	0.5700	0.0150	NT	0.8 - 5.3 (8/23)
East 1.2	East 2	0.8880	0.1380	NT	0.4 - 3.4 (0/47)
	Mean	**0.7290**	**0.0765**		
Between West/East					
East 1	West 1	0.8970	0.1260	NT	0.0 - 4.0 (0/191)
East 1	West 2	0.9140	0.1350	NT	0.0 - 1.9 (0/63)
East 1.2	West 3	0.8820	0.1380	NT	0.0 - 2.7 (0/71)
East 1.2	West 4	0.8890	0.1380	NT	0.0 - 2.7 (0/71)
East 3	West 4	0.9320	0.2210	NT	0.0 - 1.6 (0/63)
	Mean	**0.9028**	**0.1516**		

Table 2.4 AMOVA results showing the proportion of variation and F_{st} analogues calculated for different hierarchical levels of population structure under the infinite allele model. P-values were assessed using 1,000 permutations.

Source of Variation	Nested in	%var	F-stat	F-value	Std.Dev.	P-value	F'-value
Within West							
Within Individual	--	**0.744**	F_{it}	0.256	0.011	--	--
Among Individual	Population	0.12	F_{is}	0.139	0.008	0.001	--
Among Population	--	0.136	F_{st}	0.136	0.009	0.001	0.138
Within East							
Within Individual	--	0.203	F_{it}	0.797	0.007	--	--
Among Individual	Population	0.15	F_{is}	0.425	0.012	0.001	--
Among Population	--	**0.647**	F_{st}	0.647	0.01	0.001	0.655
Between West/East							
Within Individual	--	0.262	F_{it}	0.738	0.005	--	--
Among Individual	Population	0.114	F_{is}	0.303	0.004	0.001	--
Among Population	East/West	0.091	F_{sc}	0.195	0.005	0.001	0.2

2.8 Figures

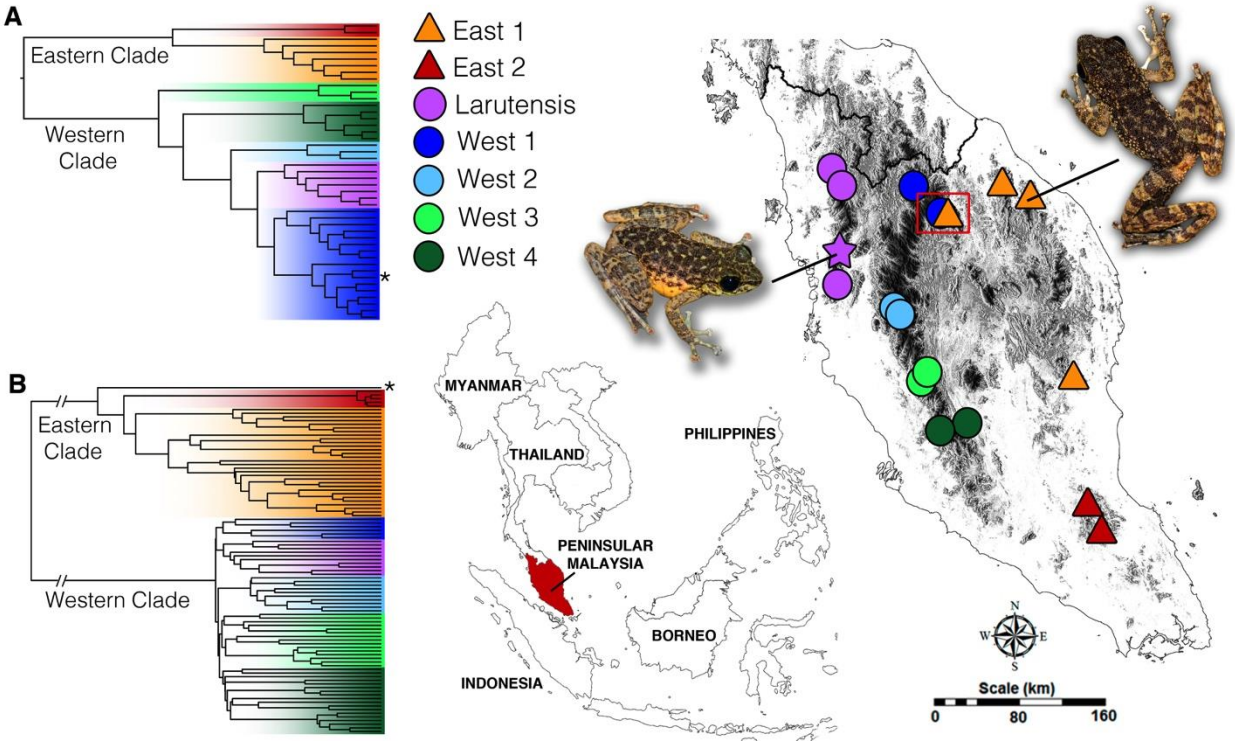


Figure 2.1 Ultrametric maximum likelihood phylogenies inferred from **A)** 1,466 bp of the 16S rRNA-encoding mitochondrial gene. All major nodes were highly supported with >90% bootstrap; **B.** 17,123 unlinked SNP loci filtered at a minimum depth of 5 and allowing for 50% missing data. All major nodes were highly supported with >90% bootstrap. The asterisk (*) denotes the putative hybrid sample that was placed in the western clade in the mtDNA phylogeny and the eastern clade in the SNP phylogeny. The distribution map (right) shows sampling localities and examples of phenotypic differences between populations from the western (circles) and eastern (triangles) clades. The star represents the type locality of *Amolops larutensis* and the red box indicates the location of the putative contact zone between the eastern and western clades. **Inset:** location of Peninsular Malaysia within Southeast Asia.

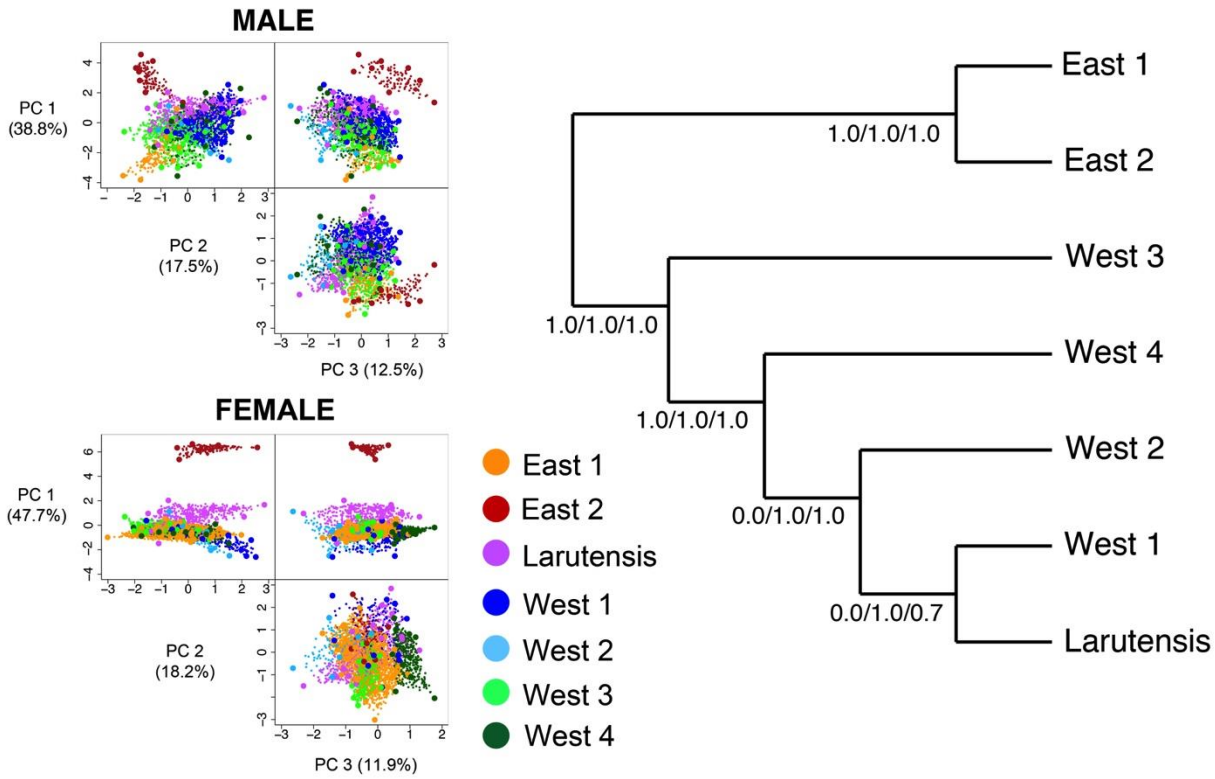


Figure 2.2 Left: Principal components scores of morphological variables visualized as three-dimensional hypervolumes constructed using multidimensional kernel density estimation. Geometry of hypervolumes are based on minimum convex polytopes and axes show the first three principal components and their proportion of variance. **Right:** Results of the mPTP and iBPP species delimitation analyses depicted on an mtDNA cladogram. Values at internal nodes denote the average support value for the mPTP analysis followed by posterior probabilities from the iBPP analysis for males and females, respectively.

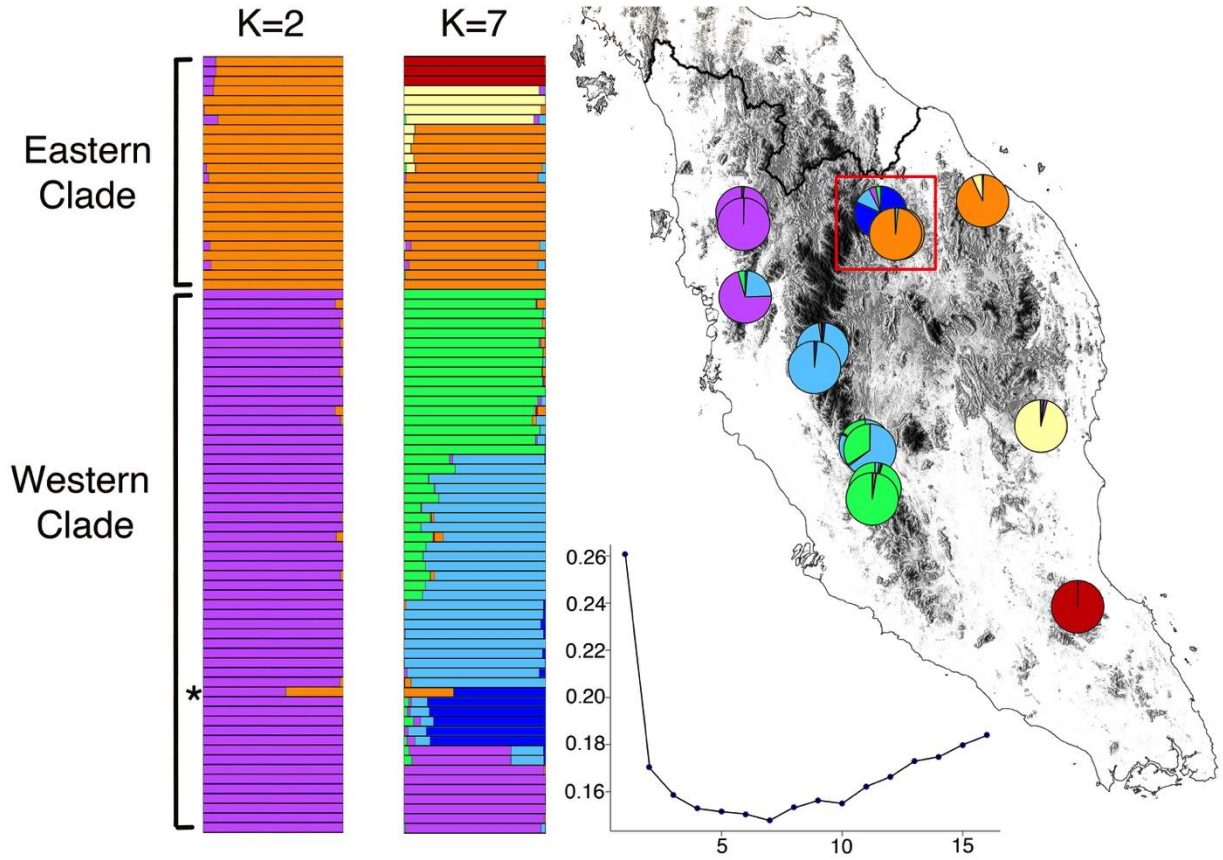


Figure 2.3 Estimated population structure as inferred by the sNMF analysis. Each individual is partitioned into K colored segments that represent the proportions of an individual's genome that originate from one or multiple inferred genetic clusters colored consistently with the other figures (note the light yellow cluster was not detected using morphological PCA). Asterisk (*) indicates the putative hybrid sample at the inferred contact zone between eastern and western clades (red box). The inset graph plots cross-entropy values (y-axis) versus number of ancestral populations (x-axis). K=2 splits individuals into eastern and western genotypes. Population structure for K=7 (the number of clusters with the lowest cross-entropy value) is plotted on the map using the average ancestry coefficient values for each estimated population.

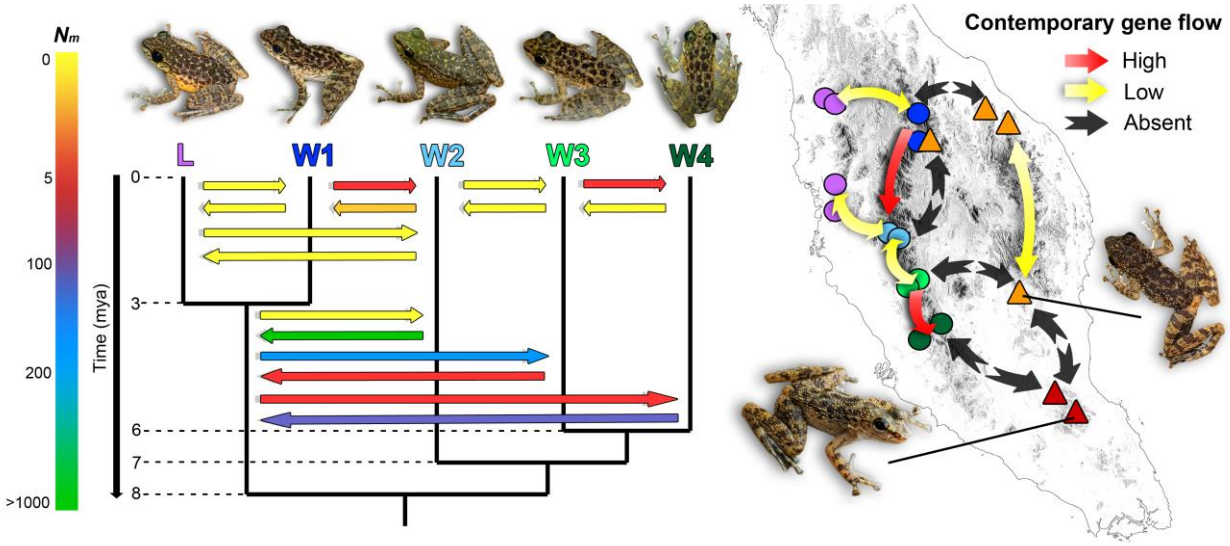


Figure 2.4 Left: A phylogram depicting contemporary and historical migration rates (N_m) for populations from the western clade estimated using fastsimcoal2 under a full asymmetrical migration model. **Right**: Contemporary gene flow scenarios based on plausible phylogenetic relationships, population structure, and geography. High gene flow: $N_m > 1$ and all sample location comparisons significant for D-statistic; low gene flow: $0.1 < N_m < 1$ and some comparisons significant for D-statistic; no gene flow: $N_m < 0.1$ or no significant comparisons for D-statistic.

CHAPTER 3

Drivers of genetic differentiation and modes of speciation in Peninsular Malaysian *Amolops*

3.1 Introduction

Genetic variation is the building block on which evolution acts on and plays an important role in the survival and adaptability of a species. An important evolutionary process that underlies genetic variation among populations is gene flow. Barriers to gene flow can be influenced by intrinsic factors such as reproductive compatibility (Seehausen *et al.* 2014), extrinsic factors such as physical barriers (Edwards *et al.* 2012; Geissler *et al.* 2015), or a combination of both whereby local adaptation to different environments affects reproductive compatibility (Glor & Warren 2011; Nosil 2012; Savolainen *et al.* 2013). Understanding the factors that affect gene flow among populations can therefore provide valuable insights into how biodiversity is generated and maintained and can be important to answer many biological questions related to speciation (Chan *et al.* 2017; Jackson *et al.* 2017), adaptation (Nadeau *et al.* 2016), hybridization (Payseur & Rieseberg 2016), or conservation (Castillo *et al.* 2014; Cushman *et al.* 2006).

The field of landscape genetics focuses on the influence of landscape characteristics on gene flow and consequently, genetic differentiation (Manel *et al.* 2003; Sork *et al.* 1999; Storfer *et al.* 2007). Populations can diverge through selection or local adaptation to different environments, a process generally referred to as isolation-by-environment, IBE (Wang &

Bradburd 2014; Wang & Summers 2010). Isolation-by-environment can be caused by abiotic factors such as climatic variables (Leamy *et al.* 2016), or biotic factors such as vegetation density and host (Via & Hawthorne 2002). However, genetic differentiation can also be generated by neutral processes that are influenced by landscape features and configurations. Spatial arrangements of populations can affect gene flow due to limited dispersal where individuals that are geographically close to each other tend to be more genetically similar than individuals that are far apart, a phenomenon known as isolation-by-distance, IBD (Wright 1943). Isolation-by-resistance, IBR (Mcrae 2006) on the other hand, produces genetic differentiation as a result of landscape features such as habitat corridors, road networks, mountains and rivers that form physical barriers or resistances to migration (Cushman *et al.* 2014; Cushman & Landguth 2012; Ozerov *et al.* 2012; Spear *et al.* 2010). Additionally, historical events such as vicariance or founder events/colonization (isolation-by-colonization, IBC) can also generate genetic structure that persists through time and are unaffected by landscape/environmental characteristics (Dewoody *et al.* 2015; Lanier *et al.* 2015; Nadeau *et al.* 2016; Nason *et al.* 2002; Orsini *et al.* 2013b). Assessing the relative contribution of historical, environment and geographic factors in shaping genetic variation has been a key focus of landscape genetics with a number of studies suggesting that environmental adaptation may an important role in population divergence (Leamy *et al.* 2016; Sexton *et al.* 2014).

Disentangling adaptive (IBE) from neutral processes (IBD, IBC) can be challenging (Wang & Bradburd 2014). For example, patterns of genetic differentiation generated by IBD can be similar to that of IBE when geography is correlated with environmental variation (Meirmans 2012). Distinguishing patterns generated by founder events (IBC) from IBD and IBE can also be complicated because they can be determined by several processes. Local adaptation can reinforce

founder events resulting in patterns similar to IBE (De Meester *et al.* 2002), and colonization can produce allele frequency gradients similar to IBD or IBE because colonization routes often covary with environmental gradients (Nadeau *et al.* 2016; Orsini *et al.* 2013a). Because selective environmental gradients, geography and colonization routes are often spatially correlated, decoupling the relative effects of adaptive from neutral processes can be complicated. However, separating the factors that contribute to genetic differentiation is not only important for land management (Cushman *et al.* 2014; Ruiz-Lopez *et al.* 2016) but is also crucial for assessing population responses to environmental change and pressures (Castillo *et al.* 2014; Reusch & Wood 2007).

Mantel tests are one of the most widely-used approaches to assess spatial processes that drive population structure, particularly to detect IBD and to partial out the relative contributions of IBD and IBE (Meirmans 2012). However, numerous studies have shown that Mantel tests do not provide an accurate decomposition of the genetic variation and therefore is unable to detect spatial structures or control for spatial autocorrelation in the relationship between genetic variation (Diniz-Filho *et al.* 2013; Guillot & Rousset 2013; Legendre *et al.* 2015; Meirmans 2012). As an alternative, ordination techniques such as redundancy analysis (RDA) have been shown to be an improvement over the Mantel test because they do not require distance-based metrics and can overcome underlying assumptions of the Mantel test (e.g. linear relationships between variables). Additionally, distance-based RDA (dbRDA) can utilize transformations such as principal coordinates analysis (PCoA) to linearize genetic variables, thus removing any potential violations of linearity observed in Mantel tests (Guillot & Rousset 2013; Kierepka & Latch 2015). Furthermore, RDA is able to provide ANOVA-like statistics such as variance around F-ratios and variance explained by each dependent variable, which allows for more

robust interpretation of results compared to Mantel tests that only provide a correlation coefficient and P-value (Kierepka & Latch 2015; Legendre *et al.* 2015). Given these advantages, constrained ordination methods such as dbRDAs represent improved alternatives to the Mantel tests.

Peninsular Malaysia is one of the most biodiverse yet highly threatened regions in Southeast Asia (Myers *et al.* 2000). Conversion and fragmentation of large tracts of rainforests for timber, infrastructure and agricultural development are main contributors to the loss of biodiversity in this region (Sodhi *et al.* 2004; Wilcove *et al.* 2013). Despite this imminent crisis, no studies have ever been performed on any vertebrate group to investigate the processes involved in shaping genetic diversity across the highly heterogeneous landscape in Peninsular Malaysia. Recent phylogenetic studies have shown that numerous species of amphibians and reptiles are endemic to specific mountain ranges (Chan *et al.* 2014c; Grismer *et al.* 2012, 2013, 2015; Sumarli *et al.* 2016; Wood *et al.* 2009), indicating that landscape characteristics could play an important role in shaping genetic diversity. This study utilizes restriction site-associated DNA sequencing (RADSeq) data generated by Chan *et al.* (2017) to investigate the factors that shape the genetic diversity of Malaysian Torrent Frogs (genus *Amolops*). The specific objectives of this study are: (i) test the proposed scenario that spatial configuration of mountain ranges plays an important role in genetic diversification; (ii) identify the relative contributions of other factors that may shape genetic diversity including geographic distance (IBD), river basins, forest cover, habitat suitability (IBE) and historical vicariance/founder events (IBC). This represents the first landscape genomic study in any Peninsular Malaysian vertebrate and is therefore a crucial step towards understanding the role of adaptive and neutral processes in generating biodiversity in Malaysia.

3.2 Methods

3.2.1 Spatial gene flow and demographic history

The genetic dataset consisted of 95 individuals of *Amolops* genotyped at 17,123 unlinked SNP loci sampled from 20 unique locations across Peninsular Malaysia (see Chan *et al.* 2017 for a complete description of molecular techniques). Individual samples were assigned to eight geographic/genetic populations following results from Chan *et al.* (2017) corresponding to the populations Larutensis, W1, W2, W3, W4, E1, E1.2, E2 (Fig 3.1A). A comprehensive assessment of temporal gene flow among populations was provided by Chan *et al.* (2017). However, in order to investigate the effects of landscape features on genetic variation, gene flow has to be evaluated within a spatial context. Effective migration was used to model the relationship between genetics and geography and visualized using an estimated effective migration surface, EEMS (Petkova *et al.* 2016). This approach uses a population genetic model to relate effective migration rates to expected genetic dissimilarities where effective migration is low in regions where genetic similarity decays quickly. To capture population structure, the study area was covered with a dense regular grid composed of triangular demes. Under a stepping-stone model, each deme exchanges migrants with only its neighbors and expected genetic dissimilarities depend on sample locations and migration rates. The expected genetic dissimilarities between two individuals is computed by integrating over all possible migration histories in their genetic ancestry which is approximated using resistance distance that integrates all possible migration routes between two demes. The estimation procedure adjusts the migration rates for all edges so that genetic differences expected under the model closely matches the genetic differences observed in the data. The estimates are then interpolated across the habitat

area to produce an estimated effective migration surface that provides a visual summary of the observed genetic dissimilarities and how they relate to geographic location (Petkova *et al.* 2016).

The genetic dissimilarity matrix was computed using the mean allele frequency imputation method implemented in the R script 'str2diffs' available from the EEMS Github repository (Petkova 2017). We used two different grid densities (200 and 300 demes) to assess the effect of grid resolution on the interpolation of effective migration rates across a continuous landscape. Four independent MCMC chains (500,000 generations/chain) were initiated from different randomly initialized parameter state. Chains were combined and assessed for convergence by plotting the trace of log posterior probabilities against MCMC iterations after a 100,000 generation burn-in period.

Demographic history was estimated using a stairway plot approach which infers changes in population size over time using SNP frequency spectra, SFS (Liu & Fu 2015). To reduce the effects of missing data, the SNP dataset was further filtered to include only loci that contained no more than 10% missing data. A folded SFS was then generated using δaδi v1.7.0 (Gutenkunst *et al.* 2009) for input into the stairway plot analysis. The total number of observed nucleic sites (including polymorphic and monomorphic) was calculated from raw pyRAD outputs and set at 10,000,000. Mutation rate was set at 1.9×10^{-8} (Chan *et al.* 2017) with an assumed generation time of 4 years. We used 67% percent of sites for training with four random break points for each try (18, 36, 54, 70) and stairway plot estimations were generated using 200 bootstrap SFs.

To estimate lineage specific changes in contemporary and ancestral population sizes, we implemented a Generalized Phylogenetic Coalescent Sampler (G-PhoCS) approach (Gronau *et al.* 2011) using as input, 17,995 SNP loci filtered at a minimum depth of 5 and 50% missing data (Chan *et al.* 2017). Demographic parameters were estimated for the three major lineages (west,

(east, south)) and priors for all population size (θ) and divergence time (τ) parameters used a gamma distribution with α = 1, β = 10,000 and priors for all migration (*m*) parameters used a gamma distribution with α = 0.002 and β = 0.00001 following recommendations from Gronau *et al.* (2011). Because gene flow was already estimated in the previous chapter and are not the focus of this study, only two migration bands were assessed: east to south and east to west. A constant rate was used to model rate variation across loci and sampling was performed using a total of 300,000 MCMC iterations with the first 100,000 discarded as burn-in. The automatic fine-tune procedure was invoked and fine-tune settings were dynamically searched using the first 10,000 samples, with fine-tunes being updated every 100 samples. MCMC sampling output was assessed for convergence using the program Tracer (Rambaut *et al.* 2014).

3.2.2 Isolation by distance, environment, colonization

Genetic distances were represented by pairwise population F_{st} values that were calculated using GENODIVE v2.0b27 (Meirmans & Van Tienderen 2004). Due to the high divergence between western and eastern lineages, missing data were imputed using population allele frequencies prior to F_{st} calculations. Geographic distances were transformed into spatial eigenfunctions in the form of distance-based Moran's eigenvector maps (dbMEM) for use as an independent spatial variable in subsequent ordination analyses (Legendre 2013; Legendre *et al.* 2015). The dbMEM analyses was performed in the R package *adespatial* using the function *dbmem*. This function first calculates a Euclidean distance matrix among all populations. The threshold value *thresh* for the truncation of the geographic distance matrix was set as the length of the longest edge of the minimum spanning tree. All distances that were larger than the truncation threshold were modified to 4**thresh*. A PCoA was then performed on the modified matrix and eigenfunctions

that model positive spatial correlation (Moran's *I* larger than expected value of *I*) of the dbMEMs are retained as spatial variables. A forward selection procedure was then performed to reduce the number of dbMEMs (Blanchet *et al.* 2008).

Environmental variables were generated using categorical (mountains and rivers; Fig. 3.1A, B) and continuous raster data (forest cover and habitat suitability; Fig. 3.1C, D). To test the Mountain Range Hypothesis, populations from the same mountain ranges were assumed to exchange genes freely and were coded with a low resistance value of 1. Populations from different but adjacent mountain ranges were given a value of 10 and populations from non-adjacent mountain ranges (e.g. between western and eastern ranges) were given a value of 100. Because Torrent Frogs are strictly associated with streams, an independent landscape variable based on river basin configurations was also constructed. River basins were calculated from a 90 m SRTM digital elevation model using the QGIS GRASS plugin *r.watershed* (QGIS Development Team 2017) with a minimum size of exterior watershed basin set at 10,000. Similarly, populations that occurred within the same major river basin were given a low resistance value of 1 and populations that occurred in separate basins were given a resistance value of 10 (Fig. 3.1B).

For continuous landscape variables, raster data for forest cover was derived from a 2015 land cover map of Southeast Asia at 250 m spatial resolution (Miettinen *et al.* 2016) obtained from the Centre of Remote Imaging, Sensing and Processing (CRISP) at the National University of Singapore (Fig. 3.1C). Habitat suitability was represented by an ecological niche model generated using 19 WorldClim (version 2) bioclimatic variables at a resolution of 30 seconds (Fick & Hijmans 2017). To reduce the dimensionality and correlation within the dataset, variables were subjected to a principal components analysis using the R function *iPCARaster*

from the package 'ENMGadgets' (Barve & Barve 2014). The first three principal components that accounted for 99% of the total variance were retained and used to generate niche models in the program Maxent (Phillips *et al.* 2006). The final niche model was constructed using the median values of 10 independent Maxent runs (Fig. 3.1D). All categorical and continuous landscape variables were subsequently transformed into resistance matrices using CIRCUITSCAPE 4.0 (McRae & Shah 2009). To test for IBC, an east-west ancestry variable that corresponded to the major diverging split of the eastern and western lineages was used. This variable was represented by the Q-values from the sNMF population structure analysis at K=2 (Chan *et al.* 2017).

3.2.3 Statistical analyses

A dbRDA analysis was implemented to assess the significance and relative contribution of IBD, IBE and IBC to genetic variation. Because dbRDA requires independent variables to be site-specific, the pairwise resistance matrices from CIRCUITSCAPE were summarized by population using a negative exponential distribution kernel that represents a connectivity index based on the Incidence Function model:

$$S_i = \sum \exp\left(-\alpha d_{ij}\right)$$

S_i is the connectivity index for the cell i, α is a scalar correlated with average dispersal distance of the species, and d is the resistance value between sites i and j (Kierepka *et al.* 2016; Moilanen & Nieminen 2002). Because no information is available on the average dispersal distance of *Amolops*, α was set at 1. This value is considered reasonable given the similarity in ecology/behavior of *Amolops* that does not suggest different average dispersal distances among

populations. All independent variable matrices were scaled to a mean of zero and a variance of one prior to dbRDA analyses. The dependent variable (genetic distance) was left untransformed because the dbRDA analysis transforms this variable using PcoA and extracts all PcoA vectors that have positive eigenvalues. To inspect the potential correlation between independent variables, a correlation analysis was also performed. For brevity, the independent variables will be referred to as Distance (dbMEMs of geographic distance), Ancestry (east-west Q-matrix), and Mountains, Rivers, Forest and Habitat for the resistance matrices representing mountain ranges, river basins, forest cover and habitat suitability respectively.

The dbRDA analysis was first performed on the entire dataset that included both the western and eastern lineages. A potential problem of analyzing the entire dataset that includes a high range of genetic divergences could arise from a strong signal (e.g. diversification of the western and eastern lineages) potentially masking out weaker signals that could be acting independently on populations within the western and eastern lineages. Additionally, assessing the contribution of specific variables responsible for within or between lineage diversification would be problematic due to non-identifiability issues in a hierarchical dataset. To circumvent these problems, the dbRDA analysis was also performed on separate datasets filtered to eastern and western populations.

The dbRDA analyses were first performed separately on each independent variable using the R function *capscale* in the package 'vegan' (Oksanen *et al.* 2017). To partial out the effects of spatial autocorrelation, a partial ordination analysis was then performed by conditioning each variable on geographic distance. Statistical significance of models was assessed using 999 permutations. In addition to the R^2 statistic, an adjusted R^2 (R^2_{adj}) was also calculated to adjust for multiple predictors in the model. The best overall model was assessed using a forward and

backward selection procedure (Blanchet *et al.* 2008) to select the most significant variables that explained the observed genetic variation. To avoid including correlated variables in the model, a statistic called the variance inflation factor (VIF) was calculated where a value of 1 represents completely independent variables, and values above 10 or 20 are regarded as highly multicollinear (Oksanen 2012). Finally, variance partitioning (Borcard *et al.* 1992) was performed using the R function *varpart* to provide a neutral decomposition of variation into unique and shared components using environmental, spatial and ancestral components as sources of genetic variation.

3.3 Results

3.3.1 EEMS and demographic history

For the EEMS analysis, combined MCMC chains ran using 200 and 300 demes converged (Appendix Fig. 14). Plots of observed versus fitted dissimilarities between and within demes showed a strong linear relationship (Fig. 3.2, 3.3), indicating that the EEMS model fitted the data well. As expected, the analysis using 300 demes showed better resolution and detected finer-scale differences in migration and diversity estimates and therefore will be retained for subsequent discussions. The estimated effective migration surface derived from posterior mean migration rates (*m*) showed that effective migration rates were higher than the overall average (highlighted in blue) among populations within the western and eastern lineages. Conversely, effective migration rates were lower than the overall average (highlighted in orange) in areas separating eastern from western populations and also separated the southern population from both eastern and western populations (Fig. 3.4A). Areas where effective migration rates are significantly higher or lower than the overall average can be visualized by highlighting areas

where the posterior probability Pr{$m > 0$} or Pr{$m < 0$} exceeds 90% and can be interpreted as areas representing corridors and barriers to gene flow (Fig. 3.4B). Diversity estimates were congruent with migration rate estimates where corridor areas along the northwestern mountain range showed low diversity estimates (orange). Areas along the eastern and southern mountain ranges showed higher diversity estimates (blue) (Fig. 3.4C).

The stairway plot analysis that was performed on the entire dataset (east + west) revealed two bottleneck events that severely reduced the overall effective population size of *Amolops*. Effective populations size increased after the first bottleneck event before being reduced again as a result of a second bottleneck event. Effective population size increased rapidly following the second bottleneck event to an even larger size (Fig. 3.5). A separate analysis was conducted on populations of the western lineage only but no significant changes in effective population size were detected (Fig. 3.5).

Bottleneck events were also detected using the G-PhoCS analysis. The ancestor of the eastern and western lineage had a large population size ($\theta_{root} = 120$) that was drastically reduced when the western lineage diverged ($\theta_{west} = 44$). Similarly, the ancestral population size of the eastern lineage ($\theta_{ES} = 107$) was severely reduced when East and South populations diverged ($\theta_{east} = 17$; $\theta_{South} = 5$). Most parameters converged with high ESS values (>500) with the exception of θ_{west} and θ_{east} that had moderate ESS values (200 > ESS > 150; Table 3.1).

3.3.2 IBD, IBE, and IBC

Results of the dbRDA analyses are summarized in Table 3.2: Distance and Ancestry were the only variables that showed significant associations with genetic distance ($p = 0.001$ and 0.008 respectively). This was also reflected in high R^2_{adj} values ($R^2_{adj} = 0.67$ and 0.55 respectively).

The significant relationship between Ancestry and genetic distance was still significant after separating out the effect of Distance [p(Ancestry | Distance) = 0.04]. None of the environmental variables showed significant relationships with genetic distance with the exception of Rivers that was marginally significant [p(Rivers) = 0.05]. However, Rivers were insignificant after controlling for Distance [p(Rivers | Distance) = 0.4]. When the dbRDA analysis was performed separately on the western dataset, Distance remained significant (p = 0.001) while Ancestry was not (p = 0.08). There was no relationship between genetic distance and Ancestry after controlling for Distance [p($Ancestry_{West}$ | Distance = 0.9; R^2_{adj} = -0.04). For the eastern dataset, Distance and Ancestry were significant (p = 0.008 and 0.05 respectively) and Ancestry remained significant after controlling for Distance [p($Ancestry_{East}$ | Distance = 0.02; R^2_{adj} = 0.6). Both forward and backward selection procedures selected Distance and Ancestry as the only variables that significantly contributed to explaining genetic variation in the dataset, while Habitat contributed the least (Table 3.3).

A Pearson's correlation test (at $\alpha = 0.01$) showed that Rivers and Mountains were significantly positively correlated while Ancestry and Distance were negatively correlated (Fig. 3.6). A further assessment of multicollinearity among independent variables within the RDA model testing framework was performed using the variance inflation factor (VIF). Results showed that Ancestry and Distance were the only variables that were independent (VIF < 10) while the variables Rivers, Mountains, Forest and Habitat were highly collinear (VIF > 20; Table 3.3). Because multicollinearity of independent variables increases estimates of parameter variance that can result in failure to detect significance in the model, collinear variables were combined into a single variable representing "Environment". The dimensionality of the Environment variable was reduced using PCA and the first two PCs that accounted for 91.3% of

the total variance were retained. The dbRDA analysis was then performed on the retained PCs but the results remained insignificant [p(Environment) = 0.15; R^2_{adj} = 0.18) (Table 3.3).

Using RDA's, genetic differentiation was partitioned into three components: (1) Distance (IBD) represented by "Distance"; (2) Environment (IBE) represented by the combined variable "Environment"; (3) Ancestry (IBC) represented by the east-west ancestry variable "Ancestry". A total of 78% of the variation in Peninsular Malaysian *Amolops* can be explained by the three components and their various combinations. Ancestry contributed 23% after constraining by Environment and Distance while Environment and Distance by themselves did not contribute to overall variation. A total of 70.5% of the explained variation was confounded between the effects of Distance, Environment and Ancestry (Fig. 3.7A). When the western and eastern lineages were analyzed separately, Distance and Ancestry explained 79% of the total variation within the western lineage. However, Distance contributed to most of the variation (53%) while Ancestry contributed 0%. A total of 16% was confounded by a shared effect between Distance and Ancestry (Fig 3.7B). Conversely for the eastern lineage, Distance only explained 12% while most of the variation was explained by Ancestry (53%) (Fig 3.7C).

3.4 Discussion

3.4.1 Drivers of genetic differentiation

Results from this study strongly rejects the hypothesis that mountain ranges play a direct role in in the genetic differentiation of Malaysian *Amolops*. Instead, the distribution of river basins explained significantly more variation. This was not unexpected given the ecology of *Amolops* that is strictly associated with stream systems. However, Mountains and Rivers were also shown to be highly correlated, most likely due to the fact that the river basin analysis was predicated on

an elevation model. These results show that mountain ranges should not be considered strictly as a physical barrier to gene flow, but instead be viewed as part of an interactive complex of environmental variables that can have confounding effects on genetic variation. This provides insight for future hypotheses testing on other co-distributed taxa that show similar patterns of distribution and genetic diversification.

When western and eastern lineages were jointly analyzed, geographic distance was shown to contribute significantly to the overall genetic variation (Table 3.2). However, when the effects of Environment and Ancestry were controlled for, Distance by itself did not uniquely contribute to genetic variation but produced shared confounding effects with the other variables (Fig 3.7A). Further insights into the confounding effects were gained when western and eastern lineages were analyzed separately: for the western lineage, Distance was the main source of genetic variation (53%) whereas for the eastern lineage, Distance only contributed 12%, indicating that Distance had contrasting effects on genetic variation on separate lineages. The differential effects of Distance could be due to the physiographic differences of the different mountain ranges that these lineages occur on. Populations of the western lineage mainly occur along the contiguous central range that could facilitate a more continuous dispersal gradient compared to the highly fragmented eastern range. However, an alternative explanation could be the bias in sampling density with the western lineage being represented by denser population sampling. Based on the current data, IBD explains most of the genetic variation within the western lineage, has a significantly smaller effect on eastern populations, and produces confounding effects when both lineages are jointly analyzed. This highlights the importance of hierarchical analyses on subsets of data that could potentially yield misleading results when analyzed as a whole.

When environmental variables were analyzed separately, Rivers contributed the most (albeit a weak significant *p*-value of 0.05), while Habitat play no role in shaping genetic diversity (R^2_{adj} = -0.12; Table 3.2). When the effects of ancestry and geographic distance were partialled out, none of the environmental variables, either considered separately or combined, contributed to the overall genetic variance (0%), indicating that IBE played no role in shaping the genetic variation in *Amolops.* These results are congruent with the conservative ecology of *Amolops* that shows little to no environmental adaptation across species.

Ancestry was significant component before and after accounting for spatial autocorrelation (Table 3.2) and contributed 23% of the genetic variation when western and eastern lineages were analyzed together (Fig. 3.7A). However when the dataset was analyzed separately, Ancestry was insignificant and did not contribute towards the genetic variation of western populations. Conversely, Ancestry was the main source of variation (53%) for the eastern lineage. Because IBC was detected in the combined western and eastern lineage analyses but not in the western lineage, suggests that IBC was responsible for the initial diversification event that formed the western and eastern lineages and subsequently for the diversification of the southern population from the rest of the eastern populations. This was followed by secondary contact between eastern and western populations at the contact zone (Fig. 3.1A) that could explain the significant negative correlation between Ancestry and Distance (Fig. 3.6). These patterns are consistent with results from the EEMS analysis that detected barriers to gene flow between eastern and western lineages and also between the southern lineage from the rest of the populations (Fig. 3.4).

3.4.2 Modes of speciation

Chan *et al.* (2017) showed that populations from the western lineage belonged to the species *Amolops larutensis* and that the eastern and southern lineages represent distinct, undescribed species. In this study, strong signatures of IBC associated with the diversification of eastern, western and southern lineages, coupled with insignificant contributions from IBE and IBD indicate that these major lineages diverged via vicariant allopatric or founder effect/peripatric speciation (Lawson *et al.* 2015; Mayr 1942; Runemark *et al.* 2012; Yeung *et al.* 2011). These modes of speciation can be hard to distinguish from each other as both processes involve an ancestral distribution, which is split into discontinuous populations and subsequently prevented from exchanging genes (Mayr 1963). However, one of the theoretical expectations of a founder event is that effective population size changes through time (Orsini *et al.* 2013b). In a classic scenario, a small number of individuals are isolated from the source population, resulting in a bottleneck which causes increased genetic drift that accelerates the formation of novel allelic combinations (Mayr 1954). The stairway plot analysis detected two bottleneck events consistent with a founder effect scenario (Fig. 3.5). Because a separate analysis of the western lineage showed no decrease in effective population size, the bottleneck events were most likely associated with founder speciation event that first separated the eastern and western lineages, followed but a second founder event that separated the southern lineage. This was corroborated by the G-PhoCS analysis that showed severe reductions in population sizes during the diversification of the western, eastern and southern lineages. Unfortunately, we were unable to associate directly, the timing of the bottleneck events with the timing of lineage diversification as parameter estimates only represent relative values. Absolute estimates of the timing and changes in population size requires reliable prior information on mutation rates and generation times,

both of which are not available. However, the strong signatures of IBC, insignificant contributions of IBD and IBD, and severe reductions in population sizes that were concomitant with major lineage diversification events, strongly suggests that interspecies diversification is a result of founder effect speciation and intraspecies population structure is caused by IBD. The negligible effect of environmental factors indicates that selection and local adaptation play an insignificant role in the evolution of *Amolops* and founder effects can result in high genetic differentiation (and speciation) without being reinforced by local adaptation.

3.5 Tables

Table 3.1. Summary statistics of the population size parameter (θ) estimated for the contemporary western (θ_{west}), eastern (θ_{east}), and southern (θ_{south}) as well we the ancestor of the east + south populations (θ_{ES}) and ancestor of the entire clade (θ_{root}). Parameter estimates are relative values that not scaled to represent effective or consensus population sizes.

	θ_{west}	θ_{east}	θ_{south}	θ_{ES}	θ_{root}
Mean	43.7375	16.6117	4.7316	102.6266	122.6828
SE mean	0.0227	0.0456	0.0156	0.0562	0.1299
Std Dev	0.5523	0.5466	0.1926	2.1342	4.2757
Variance	0.305	0.2988	0.0371	4.555	18.282
Median	43.7346	16.6038	4.7261	102.5888	122.7072
Geometric mean	43.734	16.6027	4.7277	102.6045	122.6082
95% HPD	[42.6385, 44.8003]	[15.5494, 17.6498]	[4.357, 5.1135]	[98.534, 106.8886]	[114.2699, 130.9011]
ESS	590.226	151.5438	191.9313	1441.688	1083.9253

Table 3.2. Results of the dbRDA and partial dbRDA analyses.

	R^2	R^2_{adj}	F	P-value
West + East				
Distance*	0.766	0.666	7.653	0.001
Habitat	0.042	-0.118	0.263	0.815
Mountains	0.183	0.046	1.340	0.206
Rivers	0.382	0.279	3.710	0.050
Forest	0.190	0.055	1.410	0.266
Ancestry	0.614	0.549	9.533	0.008
Habitat | Distance	0.060	-0.036	0.662	0.591
Mountains | Distance	0.044	-0.059	0.462	0.746
Rivers | Distance	0.087	0.001	1.747	0.398
Forest | Distance	0.134	0.067	1.068	0.181
Ancestry | Distance	0.201	0.161	3.184	0.041
Environment[§]	0.417	0.183	1.787	0.152
Environment | Distance	0.119	-0.094	0.596	0.743
West only				
GeoDist	0.766	0.666	7.653	0.001
Ancestry	0.187	0.096	2.066	0.083
Ancestry | Distance	0.012	-0.035	0.333	0.886
East only				
Distance	0.537	0.382	3.473	0.008
Ancestry	0.878	0.838	21.617	0.050
Ancestry | Distance	0.449	0.590	63.664	0.017

* *The Distance variable is represented by Moran's second eigenvector derived from a distance-based Moran's Eigenvector Map analysis (dbMEM). The second eigenvector was selected based on a forward selection procedure.*

[§] *Environment is a combination of the variables Rivers, Mountains, Forest and Habitat that were found to be highly collinear. These variables were summarized using the first two principal components that accounted for 91.3% of the total variance.*

Table 3.2. Model selection using forward selection. VIF = variance inflation factor; AIC = Akaike information criterion..

	VIF	AIC	F	P-value
Distance	6.627374	1.388	5.6069	0.01
Ancestry	1	-0.9429	9.533	0.015
Rivers	27.293737	2.8157	3.7098	0.09
Mountains	48.221059	5.0538	1.3402	0.23
Forest	44.321195	4.9784	1.4098	0.265
Habitat	22.736636	6.3231	0.2633	0.805

3.6 Figures

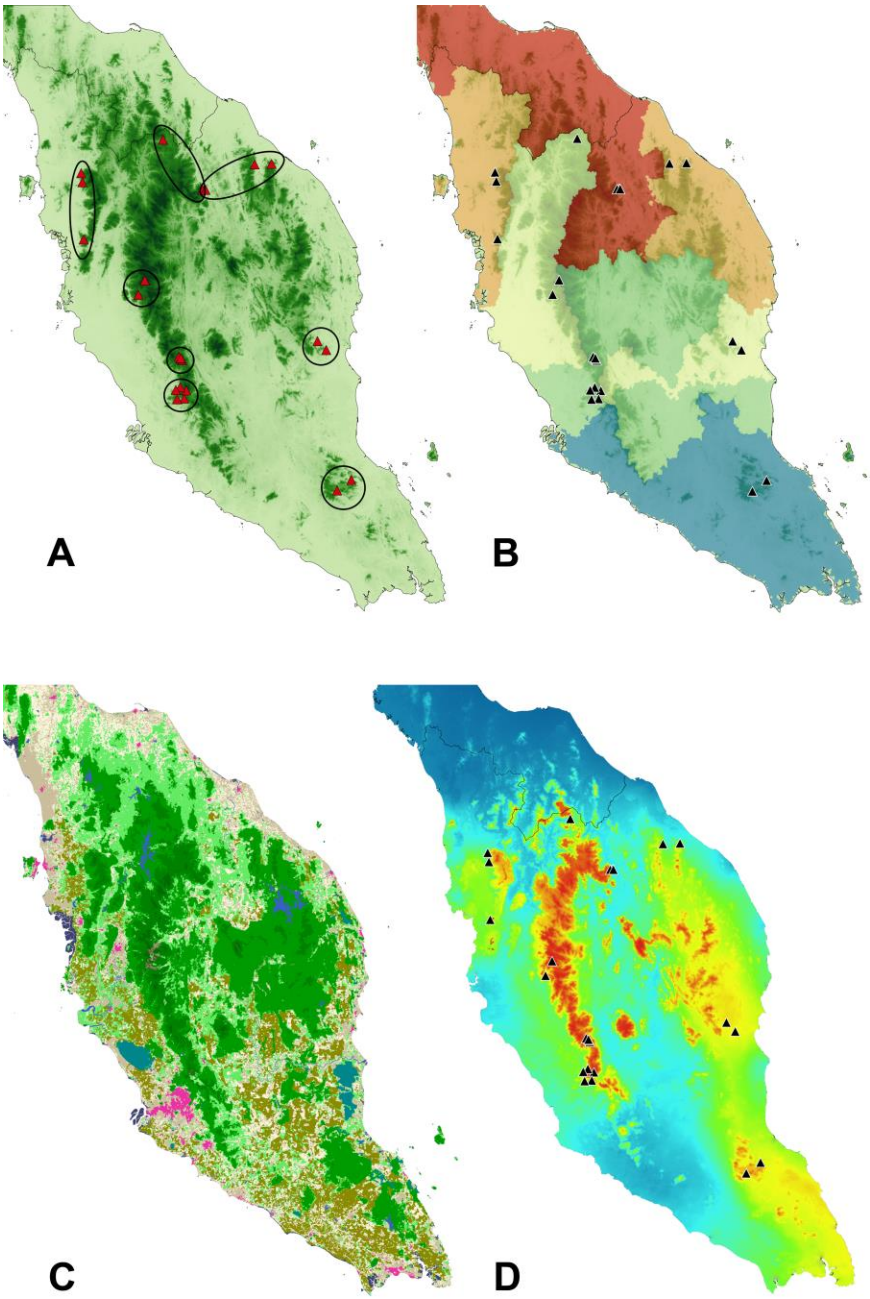


Fig. 3.1 Environmental layers used to generate resistance matrices in Circuitscape. A) Mountains: Ellipses demarcate populations following Chan *et al.* (2017); B) River basins: colors represent major watersheds estimated using QGIS; C) Forest cover: raster layer of landuse; D) Habitat suitability: environmental niche model derived from the first 3 principal components of 19 Bioclim variables.

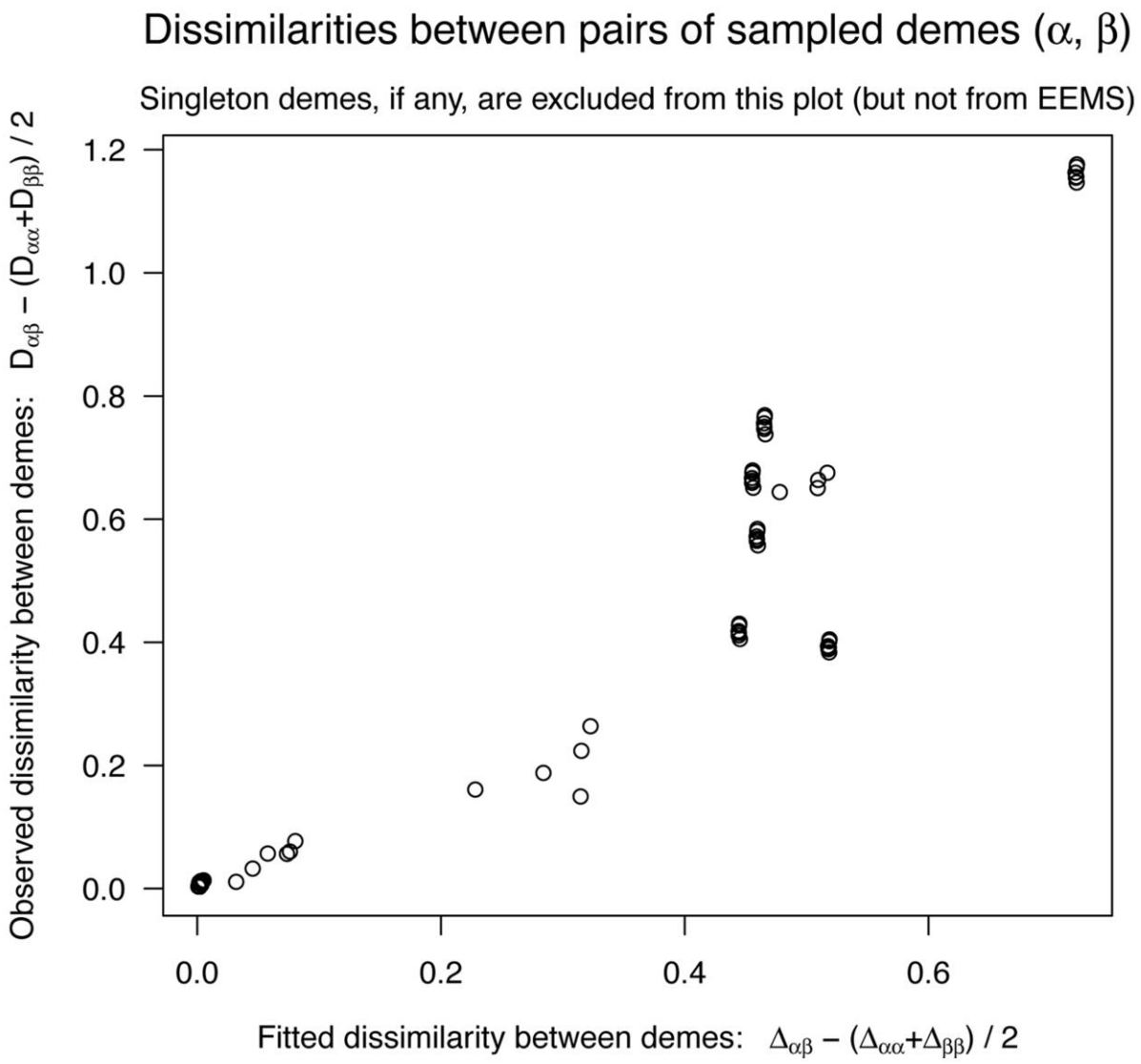


Fig. 3.2 Fitted vs. observed dissimilarities between pairs of sampled demes from the EEMS analysis.

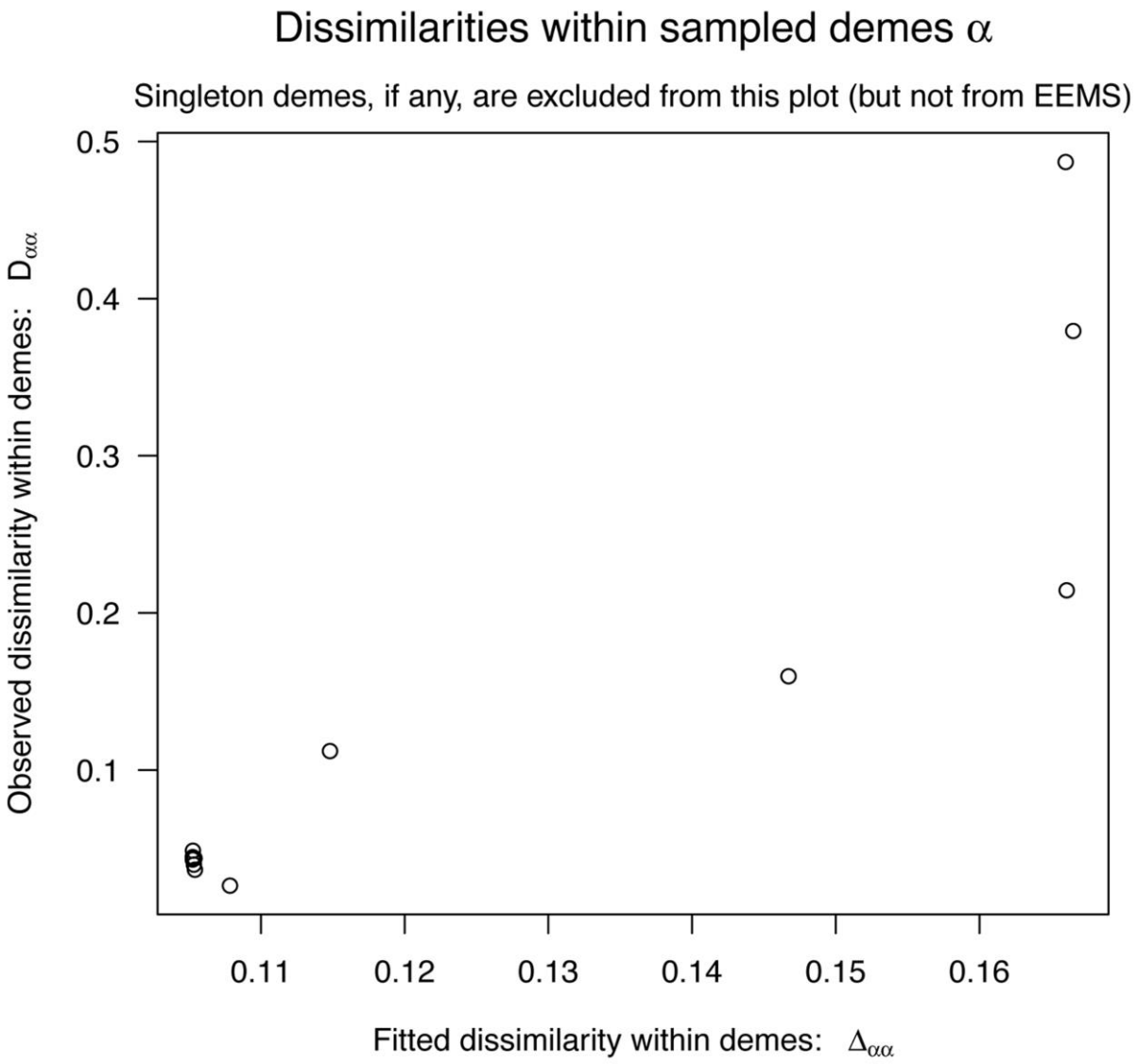


Fig. 3.3 Fitted vs. observed dissimilarities within sampled demes from the EEMS analysis.

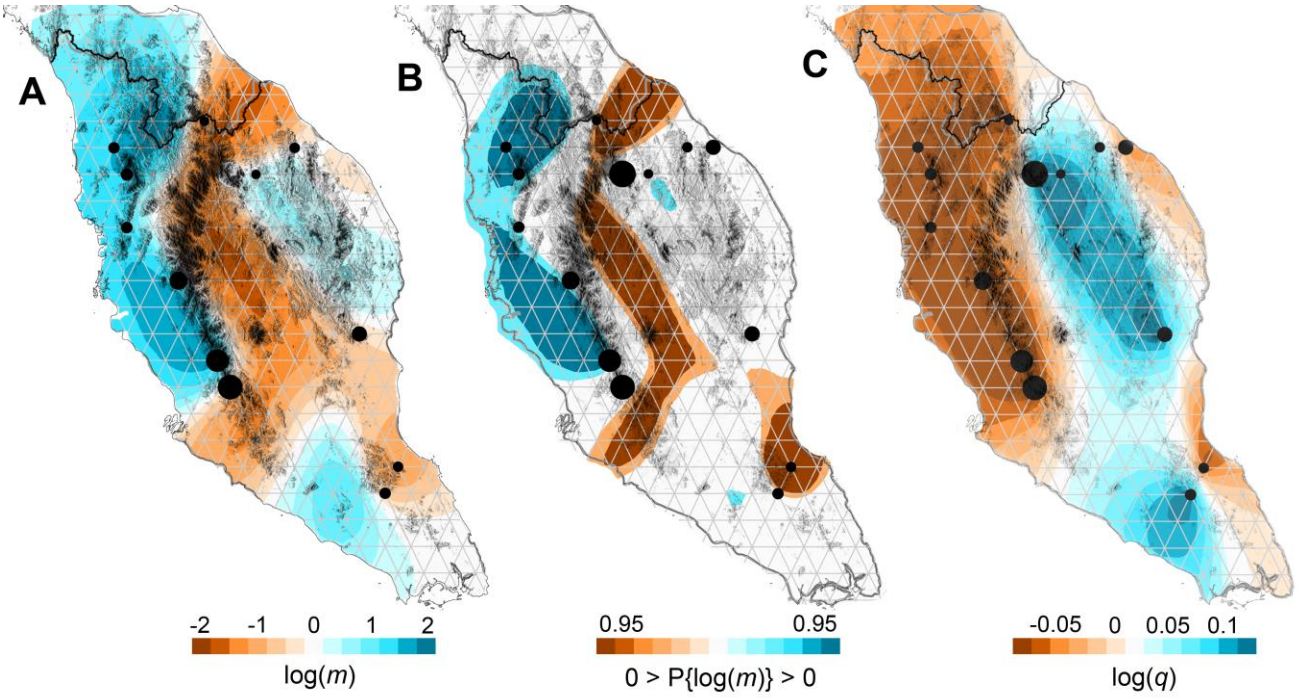


Fig. 3.4 Estimated effective migration surfaces (EEMS) at 300 demes. A) Posterior mean migration rates (m) on a $\log_{10}$ scale after mean centering where 0 corresponds to the overall mean migration rate; B) Plot that emphasizes regions where the effective migration rates are significantly higher/lower than the overall average. Areas where the posterior probability Pr{m > 0} exceeds 90% are highlighted in blue, areas where Pr{m < 0} exceeds 90% are highlighted in orange; C) Posterior mean diversity rates (q) on a $\log_{10}$ scale after mean centering.

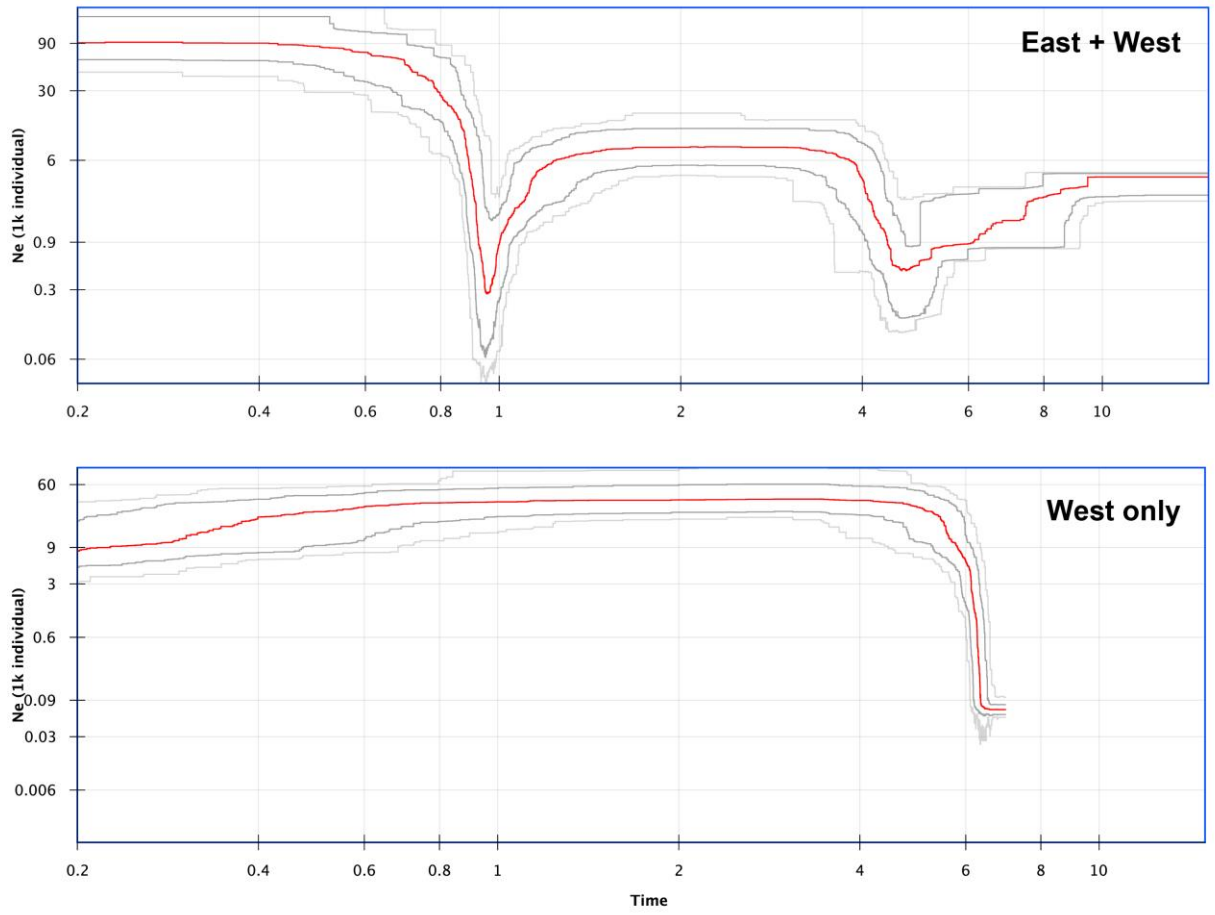


Fig. 3.5 Stairway plots showing changes in effective population size (N_e) through relative time for the combined East + West and West only datasets.

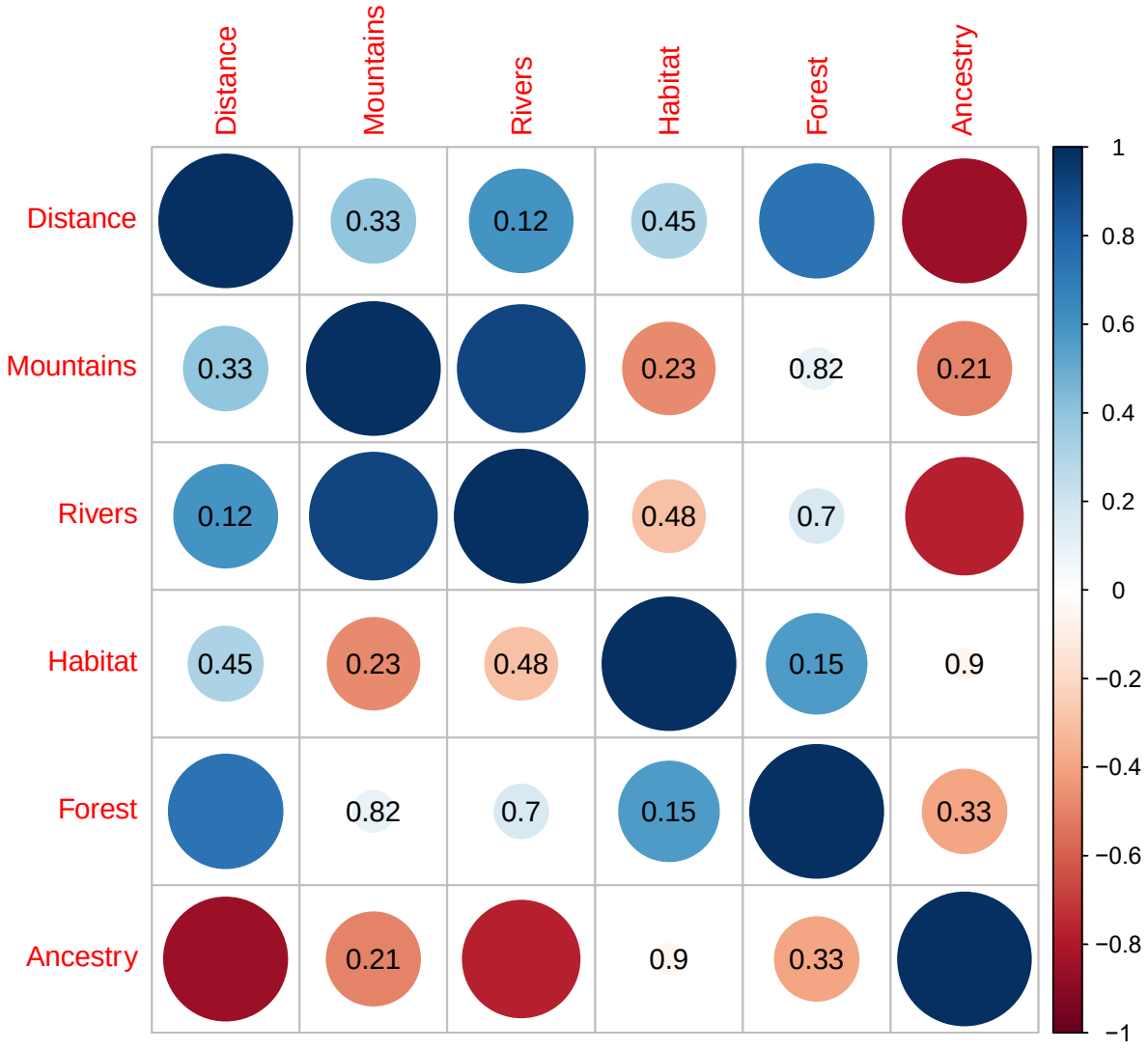


Fig. 3.6 Pairwise correlation plot of independent variables and corresponding *p*-values. Circles without values represent -0.01 > *p* < 0.01.

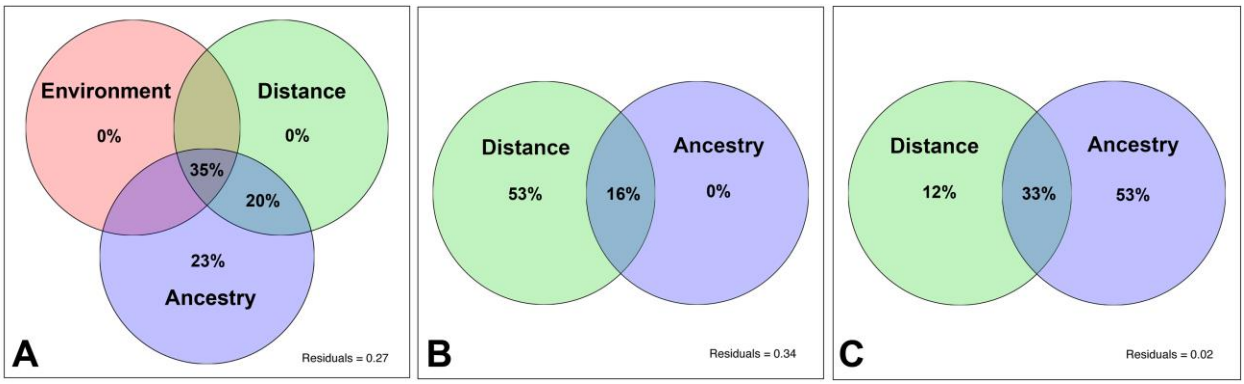


Fig. 3.7 Venn diagrams showing the relative contributions of Environment, Distance and Ancestry to genetic distance. A) west + east combined; B) west only; C) east only.

REFERENCES

Andolfatto, P., Davison, D., Erezyilmaz, D., Hu, T.T., Mast, J., Sunayama-morita, T. & Stern, D.L. (2011) Multiplexed shotgun genotyping for rapid and efficient genetic mapping. *Genome research* 21, 610–617.

Barley, A.J., White, J., Diesmos, A.C. & Brown, R.M. (2013) The challenge of species delimitation at the extremes: diversification without morphological change in Philippine Sun Skinks. *Evolution* 67, 3556–3572.

Barve, N. & Barve, V. (2014) ENMGadgets: tools for pre and post processing in ENM workflow.

Beard, K.C., Qi, T., Dawson, M.R., Wang, B. & Chuankuei, L. (1994) A diverse new primate fauna from middle Eocene fissure-fillings in southeastern China. *Nature* 368, 604–609.

Benestan, L., Gosselin, T., Perrier, C., Sainte-Marie, B., Rochette, R. & Bernatchez, L. (2015) RAD genotyping reveals fine-scale genetic structuring and provides powerful population assignment in a widely distributed marine species, the American lobster (Homarus americanus). *Molecular Ecology* 24, 3299–3315.

Blackburn, D.C., Bickford, D.P., Diesmos, A.C., Iskandar, D.T. & Brown, R.M. (2010) An ancient origin for the enigmatic flat-headed frogs (Bombinatoridae: Barbourula) from the Islands of Southeast Asia. *PLoS ONE* 5.

Blanchet, G., Legendre, P. & Borcard, D. (2008) Forward selection of spatial explanatory variables. *Ecology* 89, 2623–2632.

Bohaty, S.M. & Zachos, J.C. (2003) Significant Southern Ocean warming event in the late middle Eocene. *Geology* 31, 1017–1020.

Borcard, D., Legendre, P. & Drapeau, P. (1992) Partialling out the Spatial Component of

Ecological Variation. *Ecological Society of America* 73, 1045–1055.

Bouckaert, R. & Drummond, A. (2017) bModelTest: Bayesian phylogenetic site model averaging and model comparison. *BMC Evolutionary Biology* 17, 1–11.

Bouilhol, P., Jagoutz, O., Hanchar, J.M. & Dudas, F.O. (2013) Dating the India-Eurasia collision through arc magmatic records. *Earth and Planetary Science Letters* 366, 163–175.

Brocklehurst, N., Ruta, M., Müller, J. & Fröbisch, J. (2015) Elevated extinction rates as a trigger for diversification rate shifts: early amniotes as a case study. *Scientific Reports* 5, 17104.

Brown, R.M., Siler, C.D., Oliveros, C.H., Esselstyn, J. a., Diesmos, A.C., Hosner, P. a., Linkem, C.W., Barley, A.J., Oaks, J.R., Sanguila, M.B., Welton, L.J., Blackburn, D.C., Moyle, R.G., Townsend Peterson, a. & Alcala, A.C. (2013) Evolutionary processes of diversification in a model island archipelago. *Annual Review of Ecology, Evolution, and Systematics* 44, 411–435.

Brown, R.M. & Stuart, B.L. (2012) Patterns of biodiversity discovery through time: an historical analysis of amphibian species discoveries in the Southeast Asian mainland and adjacent island archipelagos. In: D. J. Gower, K. Johnson, J. Richardson, B. Rosen, L. Ruber, and S. Williams (Eds), *Biotic Evolution and Environmental Change in Southeast Asia*. Cambridge University Press, Cambridge, pp. 348–389.

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

Buerkle, A.C. & Gompert, Z. (2013) Population genomics based on low coverage sequencing: How low should we go? *Molecular Ecology* 22, 3028–3035.

Buerkle, C.A. (2005) Maximum-likelihood estimation of a hybrid index based on molecular

markers. *Molecular Ecology Notes* 5, 684–687.

Burbrink, F.T. & Guiher, T.J. (2015) Considering gene flow when using coalescent methods to delimit lineages of North American pitvipers of the genus Agkistrodon. *Zoological Journal of the Linnean Society* 173, 505–526.

Camargo, A., Morando, M., Avila, L.J. & Sites, J.W. (2012) Coalescent-based methods with ABC and other coalescent-based methods : a yest of accuracy with simulations and an empirical example with lizards of the Liolaemus darwinii complex (Squamata : Liolaemidae). *Evolution* 66, 2834–2849.

Candy, J.R., Campbell, N.R., Grinnell, M.H., Beacham, T.D., Larson, W.A. & Narum, S.R. (2015) Population differentiation determined from putative neutral and divergent adaptive genetic markers in Eulachon (Thaleichthys pacificus, Osmeridae), an anadromous Pacific smelt. *Molecular Ecology Resources* 15, 1421–1434.

Carstens, B.C., Pelletier, T.A., Reid, N.M. & Satler, J.D. (2013) How to fail at species delimitation. *Molecular Ecology* 22, 4369–4383.

Castillo, J.A., Epps, C.W., Davis, A.R. & Cushman, S.A. (2014) Landscape effects on gene flow for a climate-sensitive montane species, the American pika. *Molecular Ecology* 23, 843–856.

Catchen, J., Hohenlohe, P.A., Bassham, S., Amores, A. & Cresko, W.A. (2013) Stacks: An analysis tool set for population genomics. *Molecular Ecology* 22, 3124–3140.

Chaimanee, Y., Chavasseau, O., Beard, K.C., Kyaw, a. a., Soe, a. N., Sein, C., Lazzari, V., Marivaux, L., Marandat, B., Swe, M., Rugbumrung, M., Lwin, T., Valentin, X., Zin-Maung-Maung-Thein & Jaeger, J.-J. (2012) Late Middle Eocene primate from Myanmar and the initial anthropoid colonization of Africa. *Proceedings of the National Academy of*

Sciences 109, 10293–10297.

Chan, K.O., Alexander, A.M., Grismer, L.L., Su, Y.-C., Grismer, J.L., Quah, E.S.H. & Brown, R.M. (2017) Species delimitation with gene flow: a methodological comparison and population genomics approach to elucidate cryptic species boundaries in Malaysian Torrent Frogs. *Molecular Ecology* 38, 42–49.

Chan, K.O., Brown, R.M., Lim, K.K.P. & Grismer, L.L. (2014)a) A new species of frog (Amphibia: Anura: Ranidae) of the Hylarana signata Complex from Peninsular Malaysia. *Herpetologica* 70, 228–240.

Chan, K.O. & Grismer, L.L. (2010) Re-assessment of the Reinwardt's Gliding Frog, Rhacophorus reinwardtii (Schlegel 1840) (Anura: Rhacophoridae) in Southern Thailand and Peninsular Malaysia and its re-description as a new species. *Zootaxa* 2505, 40–50.

Chan, K.O., Grismer, L.L. & Brown, R.M. (2014)b) Reappraisal of the Javanese Bullfrog complex, Kaloula baleata (Müller, 1836) (Amphibia: Anura: Microhylidae), reveals a new species from Peninsular Malaysia. *Zootaxa* 3900, 569–580.

Chan, K.O., Grismer, L.L., Zachariah, A., Brown, R.M. & Abraham, R.K. (2016) Polyphyly of Asian Tree Toads, genus Pedostibes Günther, 1876 (Anura: Bufonidae), and the description of a new genus from Southeast Asia. *Plos One* 11, e0145903.

Chan, K.O., Lee Grismer, L. & Grismer, J. (2011) A new insular, endemic frog of the genus Kalophrynus Tschudi, 1838 (Anura: Microhylidae) from Tioman Island, Pahang, Peninsular Malaysia. *Zootaxa* 68, 60–68.

Chan, K.O., Wood, P.L., Anuar, S., Muin, M.A., Quah, E.S.H., Sumarli, A.X.Y. & Grismer, L.L. (2014)c) A new species of upland Stream Toad of the genus Ansonia Stoliczka, 1870 (Anura: Bufonidae) from northeastern Peninsular Malaysia. *Zootaxa* 3764, 427–440.

Che, J., Pang, J., Zhao, H., Wu, G.F., Zhao, E.M. & Zhang, Y.P. (2007) Phylogeny of Raninae (Anura: Ranidae) inferred from mitochondrial and nuclear sequences. *Molecular Phylogenetics and Evolution* 43, 1–13.

Coyne, J.A. & Orr, H.A. (2004) Sunderland, MA: Sinauer Associates *Speciation*. Sinauer Associates, Sunderland, MA.

Cracraft, J. (1985) Biological diversification and its causes. 72, 794–822.

Cushman, S.A. & Landguth, E.L. (2012) Multi-taxa population connectivity in the Northern Rocky Mountains. *Ecological Modelling* 231, 101–112.

Cushman, S.A., Lewis, J.S. & Landguth, E.L. (2014) Why did the bear cross the road? Comparing the performance of multiple resistance surfaces and connectivity modeling methods. *Diversity* 6, 844–854.

Cushman, S.A., McKelvey, K.S., Hayden, J. & Schwartz, M.K. (2006) Gene flow in complex landscapes: Testing multiple hypotheses with causal modeling. *American Naturalist* 168, 486–499.

Dewoody, J., Trewin, H. & Taylor, G. (2015) Genetic and morphological differentiation in Populus nigra L.: isolation by colonization or isolation by adaptation? *Molecular Ecology* 24, 2641–2655.

Diniz-Filho, J.A.F., Soares, T.N., Lima, J.S., Dobrovolski, R., Landeiro, V.L., Telles, M.P. de C., Rangel, T.F. & Bini, L.M. (2013) Mantel test in population genetics. *Genetics and Molecular Biology* 36, 475–485.

Drummond, A.J. & Bouckaert, R.R. (2015) Bayesian evolutionary analysis with BEAST. In: Cambridge University Press, pp. 260.

Drummond, C.S., Eastwood, R.J., Miotto, S.T.S. & Hughes, C.E. (2012) Multiple continental

radiations and correlates of diversification in lupinus (leguminosae): Testing for key innovation with incomplete taxon sampling. *Systematic Biology* 61, 443–460.

Durand, E.Y., Patterson, N., Reich, D. & Slatkin, M. (2011) Testing for ancient admixture between closely related populations. *Molecular Biology and Evolution* 28, 2239–2252.

Eaton, D.A.R. (2014) PyRAD: assembly of de novo RADseq loci for phylogenetic analyses. *Bioinformatics* 30, 1844–1849.

Eaton, D.A.R., Hipp, A.L., González-Rodríguez, A. & Cavender-Bares, J. (2015) Historical introgression among the American live oaks and the comparative nature of tests for introgression. *Evolution* 69, 2587–2601.

Eaton, D.A.R. & Ree, R.H. (2013) Inferring phylogeny and introgression using RADseq data: an example from glowering plants (Pedicularis: Orobanchaceae). *Systematic Biology* 62, 689–706.

Edwards, D.L., Keogh, J.S. & Knowles, L.L. (2012) Effects of vicariant barriers, habitat stability, population isolation and environmental features on species divergence in the south-western Australian coastal reptile community. *Molecular Ecology* 21, 3809–3822.

Ence, D.D. & Carstens, B.C. (2011) SpedeSTEM: A rapid and accurate method for species delimitation. *Molecular Ecology Resources* 11, 473–480.

Evans, B.J., Brown, R.M., McGuire, J. a, Supriatna, J., Andayani, N., Diesmos, A., Iskandar, D., Melnick, D.J. & Cannatella, D.C. (2003) Phylogenetics of fanged frogs: testing biogeographical hypotheses at the interface of the asian and Australian faunal zones. *Systematic biology* 52, 794–819.

Excoffier, L., Smouse, P.E. & Quattro, J.M. (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: Application to human mitochondrial DNA

restriction data. *Genetics* 131, 479–491.

Feng, Y.-J., Blackburn, D.C., Liang, D., Hillis, D.M., Wake, D.B., Cannatella, D.C. & Zhang, P. (2017) Phylogenomics reveals rapid, simultaneous diversification of three major clades of Gondwanan frogs at the Cretaceous–Paleogene boundary. *Proceedings of the National Academy of Sciences*, 201704632.

Fick, S.E. & Hijmans, R.J. (2017) Worldclim2: New 1-km spatial resolution climate surfaces for global land areas. *International Journal of Climatology*.

Fouquet, A., Gilles, A., Vences, M., Marty, C., Blanc, M. & Gemmell, N.J. (2007) Underestimation of species richness in neotropical frogs revealed by mtDNA analyses. *PLoS ONE* 2.

Frichot, E. & François, O. (2015) LEA: An R package for landscape and ecological association studies. *Methods in Ecology and Evolution* 6, 925–929.

Frichot, E., Mathieu, F., Trouillon, T., Bouchard, G. & François, O. (2014) Fast and efficient estimation of individual ancestry coefficients. *Genetics* 196, 973–983.

Frost, D.R. (2015) Amphibian Species of the World: an Online Reference. Version 6.0 (accessed 10 Oct 2015). *Electronic Database accessible at http://research.amnh.org/herpetology/amphibia/index.html. American Museum of Natural History, New York, USA.*

Frost, D.R. & Hillis, D.M. (1990) Species in concept and practice: Herpetological applications. *Herpetologica* 46, 86–104.

Fujita, M.K., Leache, A.D., Burbrink, F.T., McGuire, J.A. & Moritz, C. (2012) Coalescent-based species delimitation in an integrative taxonomy. *Trends in Ecology and Evolution* 27, 480–488.

Geissler, P., Hartmann, T., Ihlow, F., Rödder, D., Poyarkov, N.A., Nguyen, T.Q., Ziegler, T. & Böhme, W. (2015) The Lower Mekong: An insurmountable barrier to amphibians in southern Indochina? *Biological Journal of the Linnean Society* 114, 905–914.

Giam, X., Scheffers, B.R., Sodhi, N.S., Wilcove, D.S., Ceballos, G. & Ehrlich, P.R. (2012) Reservoirs of richness: least disturbed tropical forests are centres of undescribed species diversity. *Proceedings of the Royal Society B: Biological Sciences* 279, 67–76.

Glor, R.E. & Warren, D. (2011) Testing ecological explanations for biogeographic boundaries. *Evolution* 65, 673–683.

Grismer, L.L. (2011) *Lizards of Peninsular Malaysia, Singapore and their Adjacent Archipelagos*. Edition Chimaira, Frankfurt.

Grismer, L.L., Anuar, S., Quah, E.S.H., Muin, M.A., Chan, K.O., Grismer, J.L. & Ahmad, N. (2010) A new spiny, prehensile-tailed species of Cyrtodactylus (Squamata: Gekkonidae) from Peninsular Malaysia with a preliminary hypothesis of relationships based on morphology. *Zootaxa* 52, 40–52.

Grismer, L.L., Wood, P.L., Anuar, S., Muin, M.A., Quah, E.S.H., McGuire, J.A., Brown, R.M., Ngo, V.T., Hong Thai, P. & Pham, H.T. (2013) Integrative taxonomy uncovers high levels of cryptic species diversity in Hemiphyllodactylus Bleeker, 1860 (Squamata: Gekkonidae) and the description of a new species from Peninsular Malaysia. *Zoological Journal of the Linnean Society* 169, 849–880.

Grismer, L.L., Wood, P.L., Lee, C.H., Quah, E.S.H., Anuar, S., Ngadi, E. & Sites, J.W. (2015) An integrative taxonomic review of the agamid genus Bronchocela (Kuhl, 1820) from Peninsular Malaysia with descriptions of new montane and insular endemics. 3948, 1–23.

Grismer, L.L., Wood, P.L.J., Anuar, S., Riyanto, A., Ahmad, N., Muin, M.A., Sumontha, M.,

Grismer, J.L., Chan, K.O., Quah, E.S.H. & Pauwels, O.S.G. (2014) Systematics and natural history of Southeast Asian Rock Geckos (genus Cnemaspis Strauch, 1887) with descriptions of eight new species from Malaysia, Thailand and Indonesia. *Zootaxa* 3880, 1–147.

Grismer, L.L., Wood, P.L.J., Quah, E.S.H., Anuar, S., Muin, M.A., Sumontha, M., Ahmad, N., Bauer, A.M., Wangkulangkul, S., Grismer, J.L. & Pauwels, O.S.G. (2012) A phylogeny and taxonomy of the Thai-Malay Peninsula Bent-toed Geckos of the Cyrtodactylus pulchellus complex (Squamata: Gekkonidae): combined morphological and molecular analyses with descriptions of seven new species. *Zootaxa* 3520, 1–55.

Gronau, I., Hubisz, M.J., Gulko, B., Danko, C.G. & Siepel, A. (2011) Bayesian inference of ancient human demography from individual genome sequences. *Nature Genetics* 43, 1031–1034.

Guillot, G. & Rousset, F. (2013) Dismantling the Mantel tests. *Methods in Ecology and Evolution* 4, 336–344.

Harvey, M.G., Duffie-Judy, C., Seeholzer, G.F., Maley, J.M., Graves, G.R. & Brumfield, R.T. (2015) Similarity threshholds used in short read assembly reduce the comparability of population histories across species. *PeerJ* 3:e895, DOI 10.7717/peerj.895.

Hasan, M., Islam, M.M., Khan, M.M.R., Igawa, T., Alam, M.S., Djong, H.T., Kurniawan, N., Joshy, H., Sen, Y.H., Belabut, D.M., Kurabayashi, A., Kuramoto, M. & Sumida, M. (2014) Genetic divergences of South and Southeast Asian frogs: A case study of several taxa based on 16S ribosomal RNA gene data with notes on the generic name Fejervarya. *Turkish Journal of Zoology* 38, 389–411.

Hendry, A.P., Bolnick, D.I., Berner, D. & Peichel, C.L. (2009) Along the speciation continuum in sticklebacks. *Journal of Fish Biology* 75, 2000–2036.

Hodges, S.A. (1997) Floral nectar spurs and diversification. *International Journal of Plant Sciences* 158, S81–S88.

Hohna, S., May, M.R. & Moore, B.R. (2015) Phylogeny Simulation and Diversification Rate Analysis with TESS. , 1–98.

Höhna, S., May, M.R. & Moore, B.R. (2015) TESS: An R package for efficiently simulating phylogenetic trees and performing Bayesian inference of lineage diversification rates. *Bioinformatics* 32, 789–791.

Hooker, J.J., Collinson, M.E. & Sille, N.P. (2004) Eocene-Oligocene mammalian faunal turnover in the Hampshire Basin, UK: calibration to the global time scale and the major cooling event. *Journal of the Geological Society* 161, 161–172.

Huang, H. & Knowles, L.L. (2016) Unforeseen consequences of excluding missing data from next-generation sequences: simulation study of RAD sequences. *Systematic Biology* 65, 357–365.

Huang, Z., Yang, C. & Ke, D. (2014) DNA barcoding and molecular phylogeny in Ranidae. *Mitochondrial DNA*, 1–5.

Ilut, D.C., Nydam, M.L. & Hare, M.P. (2014) Defining loci in restriction-based reduced representation genomic data from nonmodel species: Sources of bias and diagnostics for optimal clustering. *BioMed Research International* 2014.

Jackson, N.D., Carstens, B.C., Morales, A.E. & O'Meara, B.C. (2016) Species delimitation with gene flow. *Systematic Biology*.

Jackson, N.D., Carstens, B.C., Morales, A.E. & O'Meara, B.C. (2017) Species Delimitation with Gene Flow. *Systematic Biology* 0, syw117.

Jamaluddin, J.A.F., Pau, T.M. & Siti-Azizah, M.N. (2011) Genetic structure of the Snakehead

Murrel, Channa striata (Channidae) based on the Cytochrome-c Oxidase Subunit I gene: Influence of historical and geomorphological factors. *Genetics and Molecular Biology* 34, 152–160.

Jombart, T. (2008) Adegenet: A R package for the multivariate analysis of genetic markers. *Bioinformatics* 24, 1403–1405.

Jombart, T., Devillard, S. & Balloux, F. (2010) Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. *BMC genetics* 11, 94.

Jost, L. (2008) GST and its relatives do not measure differentiation. *Molecular Ecology* 17, 4015–4026.

Kaiser, H.F. (1960) The application of electronic computers to factor analysis. *Educational and Psychological Measurement* 20, 141–151.

Kapli, P., Lutteropp, S., Zhang, J., Kobert, K., Pavlidis, P., Stamatakis, A. & Flouri, T. (2017) Multi-rate Poisson tree processes for single-locus species delimitation under maximum likelihood and Markov chain Monte Carlo. *Bioinformatics* 33, 1630–1638.

Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., Buxton, S., Cooper, A., Markowitz, S., Duran, C., Thierer, T., Ashton, B., Meintjes, P. & Drummond, A. (2012) Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* 28, 1647–9.

Khoo, S.N. & Lubis, A.R. (2005) *Kinta Valley: pioneering Malaysia's modern development*. Areca Books.

Kierepka, E.M., Anderson, S.J., Swihart, R.K. & Rhodes, O.E. (2016) Evaluating the influence of life-history characteristics on genetic structure: a comparison of small mammals inhabiting complex agricultural landscapes. *Ecology and Evolution* 6, 6376–6396.

Kierepka, E.M. & Latch, E.K. (2015) Performance of partial statistics in individual-based landscape genetics. *Molecular Ecology Resources* 15, 512–525.

Kim, H. & Park, H. (2007) Sparse non-negative matrix factorizations via alternating non-negativity-constrained least squares for microarray data analysis. *Bioinformatics* 23, 1495–1502.

Knowles, L.L. & Carstens, B.C. (2007) Estimating a geographically explicit model of population divergence. *Evolution* 61, 477–493.

Ksepka, D.T., Stidham, T.A. & Williamson, T.E. (2017) Early Paleocene landbird supports rapid phylogenetic and morphological diversification of crown birds after the K–Pg mass extinction. *Proceedings of the National Academy of Sciences* 114, 201700188.

Lanier, H.C., Massatti, R., He, Q., Olson, L.E. & Lacey Knowles, L. (2015) Colonization from divergent ancestors: Glaciation signatures on contemporary patterns of genomic variation in Collared Pikas (Ochotona collaris). *Molecular Ecology* 24, 3688–3705.

Larson, W.A., Seeb, L.W., Everett, M. V., Waples, R.K., Templin, W.D. & Seeb, J.E. (2014) Genotyping by sequencing resolves shallow population structure to inform conservation of Chinook salmon (Oncorhynchus tshawytscha). *Evolutionary Applications* 7, 355–369.

Laurent, S., Robinson-Rechavi, M. & Salamin, N. (2015) Detecting patterns of species diversification in the presence of both rate shifts and mass extinctions. *BMC evolutionary biology* 15, 157.

Lawson, L.P., Bates, J.M., Menegon, M. & Loader, S.P. (2015) Divergence at the edges: peripatric isolation in the montane spiny throated reed frog complex. *BMC Evolutionary Biology* 15, 128.

Leaché, A.D., Banbury, B.L., Felsenstein, J., de Oca, A. nieto-M. & Stamatakis, A. (2015) Short

tree, long tree, right tree, wrong tree: new acquisition bias corrections for inferring SNP phylogenies. *Systematic Biology* 64, 1032–1047.

Leaché, A.D., Fujita, M.K., Minin, V.N. & Bouckaert, R.R. (2014) Species delimitation using genome-wide SNP Data. *Systematic Biology* 63, 534–542.

Leamy, L.J., Lee, C.R., Song, Q., Mujacic, I., Luo, Y., Chen, C.Y., Li, C., Kjemtrup, S. & Song, B.H. (2016) Environmental versus geographical effects on genomic variation in wild soybean (Glycine soja) across its native range in northeast Asia. *Ecology and Evolution* 6, 6332–6344.

Leavitt, S.D., Moreau, C.S. & Lumbsch, H.T. (2015) The dynamic discipline of species delimitation: Progress toward effectively recognizing species boundaries in natural populations. In: D. K. Upreti, P. K. Divakar, V. Shukla, and R. Bajpai (Eds), *Recent Advances in Lichenology: Modern Methods and Approaches in Lichen Systematics and Culture Techniques, Volume 2*. Springer, India, pp. 11–44.

Legendre, P. (2013) Spatial eigenfunction methods : new developments. , 1–23.

Legendre, P., Fortin, M.-J. & Borcard, D. (2015) Should the Mantel test be used in spatial analysis? *Methods in Ecology and Evolution* 6, 1239–1247.

Leliaert, F., Verbruggen, H., Vanormelingen, P., Steen, F., López-Bautista, J.M., Zuccarello, G.C. & De Clerck, O. (2014) DNA-based species delimitation in algae. *European Journal of Phycology* 49, 179–196.

Leslie, S., Winney, B., Hellenthal, G., Davison, D., Boumertit, A., Day, T., Hutnik, K., Royrvik, E.C., Cunliffe, B., Wellcome Trust Case Control Consortium 2, International Multiple Sclerosis Genetics Consortium, Lawson, D.J., Falush, D., Freeman, C., Pirinen, M., Myers, S., Robinson, M., Donnelly, P. & Bodmer, W. (2015) The fine-scale genetic structure of the

British population. *Nature* 519, 309–314.

Lewis, P.O. (2001) A likelihood approach to estimating phylogeny from discrete morphological character data. *Systematic Biology* 50, 913–925.

Liu, X. & Fu, Y.-X. (2015) Exploring population size changes using SNP frequency spectra. *Nature genetics* 47, 555–9.

Lleonart, J., Salat, J. & Torres, G.J. (2000) Removing allometric effects of body size in morphological analysis. *Journal of Theoretical Biology* 205, 85–93.

Manel, S., Schwartz, M.K., Luikart, G. & Taberlet, P. (2003) Landscape genetics: Combining landscape ecology and population genetics. *Trends in Ecology and Evolution* 18, 189–197.

Matsui, M., Shimada, T., Liu, W.Z., Maryati, M., Khonsue, W. & Orlov, N. (2006) Phylogenetic relationships of Oriental torrent frogs in the genus Amolops and its allies (Amphibia, Anura, Ranidae). *Molecular Phylogenetics and Evolution* 38, 659–666.

Matzke, N.J. (2013) Probabilistic historical biogeography: new models for founder-event speciation, imperfect detection, and fossils allow improved accuracy and model-testing. *Frontiers of Biogeography* 5, 242–248.

Matzke, N.J. (2014) Model selection in historical biogeography reveals that founder-event speciation is a crucial process in island clades. *Systematic Biology* 63, 951–970.

May, M.R., Höhna, S., Moore, B.R. & Cooper, N. (2016) A Bayesian approach for detecting the impact of mass-extinction events on molecular phylogenies when rates of lineage diversification may vary. *Methods in Ecology and Evolution* 7, 947–959.

Mayr, E. (1942) *Systematics and the origin of Species: from the Viewpoint of a Zoologist*. Harvard University Press.

Mayr, E. (1954) Change of genetic environment and evolution. In: J. S. Huxley, A. C. Hardy,

and E. B. Ford (Eds), *Evolution as a Process*. Allen & Unwin, London, pp. 157–180.

Mayr, E. (1963) Animal species and evolution. *The Eugenics review* 55, 226–228.

Mayr, E. (1968) The role of systematics in biology. *Science* 159, 595–599.

McKinnon, J.S., Mori, S., Blackman, B.K., David, L., Kingsley, D.M., Jamieson, L., Chou, J. & Schluter, D. (2004) Evidence for ecology's role in speciation. *Nature* 429, 294–298.

McLeod, D.S. (2010) Of Least Concern? Systematics of a cryptic species complex: Limnonectes kuhlii (Amphibia: Anura: Dicroglossidae). *Molecular Phylogenetics and Evolution* 56, 991–1000.

McQuarrie, N. & Van Hinsbergen, D.J.J. (2013) Retrodeforming the Arabia-Eurasia collision zone: Age of collision versus magnitude of continental subduction. *Geology* 41, 315–318.

Mcrae, B.H. (2006) Isolation By Resistance. *Evolution* 60, 1551–1561.

McRae, B.H. & Shah, V.B. (2009) Circuitscape user's guide. ONLINE. The University of California, Santa Barbara.

Medrano, M., López-Perea, E. & Herrera, C.M. (2014) Population genetics methods applied to a species delimitation problem: endemic trumpet daffodils (Narcissus Section Pseudonarcissi) from the Southern Iberian Peninsula. *International Journal of Plant Sciences* 175, 501–517.

De Meester, L., Gómez, A., Okamura, B. & Schwenk, K. (2002) The Monopolization Hypothesis and the dispersal-gene flow paradox in aquatic organisms. *Acta Oecologica* 23, 121–135.

Meirmans, P.G. (2012) The trouble with isolation by distance. *Molecular Ecology* 21, 2839–2846.

Meirmans, P.G. & Hedrick, P.W. (2011) Assessing population structure: FST and related measures. *Molecular Ecology Resources* 11, 5–18.

Meirmans, P.G. & Van Tienderen, P.H. (2004) GENOTYPE and GENODIVE: Two programs for the analysis of genetic diversity of asexual organisms. *Molecular Ecology Notes* 4, 792–794.

Miettinen, J., Shi, C. & Liew, S.C. (2011) Deforestation rates in insular Southeast Asia between 2000 and 2010. *Global Change Biology* 17, 2261–2270.

Miettinen, J., Shi, C. & Liew, S.C. (2016) 2015 Land cover map of Southeast Asia at 250 m spatial resolution. *Remote Sensing Letters* 7, 701–710.

Minh, B.Q., Nguyen, M.A.T. & von Haeseler, A. (2013) Ultrafast approximation for phylogenetic bootstrap. *Molecular Biology and Evolution* 30, 1188–1195.

Mittermeier, R.A., Myers, N., Thomsen, J.B., da Fonseca, G.A.B. & Olivieri, S. (1998) Biodiversity hotspots and major tropical Wilderness areas: approaches to setting conservation priorities. *Conservation Biology* 12, 516–520.

Moilanen, A. & Nieminen, M. (2002) Simple connectivity measure in spatial ecology. *Ecology* 83, 1131–1145.

Moore, B.R. & Donoghue, M.J. (2007) Correlates of diversification in the plant clade Dipsacales: geographic movement and evolutionary innovations. *Am. Nat.* 170, S28–S55.

Moore, B.R., Höhna, S., May, M.R., Rannala, B. & Huelsenbeck, J.P. (2016) Critically evaluating the theory and performance of Bayesian analysis of macroevolutionary mixtures. *Proceedings of the National Academy of Sciences* 113, 201518659.

Morley, R.J. (2012) Biotic Evolution and Environmental Change in Southeast Asia *A review of the Cenozoic palaeoclimate history of Southeast Asia*.

Myers, N., Mittermeier, R.A., Mittermeier, C.G., da Fonseca, G.A.B. & Kent, J. (2000) Biodiversity hotspots for conservation priorities. *Nature* 403, 853–858.

Nadeau, S., Meirmans, P.G., Aitken, S.N., Ritland, K. & Isabel, N. (2016) The challenge of separating signatures of local adaptation from those of isolation by distance and colonization history: The case of two white pines. *Ecology and Evolution* 6, 8649–8664.

Nason, J.D., Hamrick, J.L. & Fleming, T.H. (2002) Historical Vicariance And Postglacial Colonization Effects On The Evolution Of Genetic Structure In Lophocereus, A Sonoran Desert Columnar Cactus. *Evolution* 56, 2214–2226.

Nee, S. (2006) Birth-Death Models in Macroevolution. *Annual Review of Ecology, Evolution, and Systematics* 37, 1–17.

Nguyen, L.-T., Schmidt, H. a., von Haeseler, A. & Minh, B.Q. (2014) IQ-TREE: A fast and effective stochastic algorithm for estimating maximum likelihood phylogenies. *Molecular Biology and Evolution* 32, 268–274.

Niemiller, M.L., Fitzpatrick, B.M. & Miller, B.T. (2008) Recent divergence with gene flow in Tennessee cave salamanders (Plethodontidae: Gyrinophilus) inferred from gene genealogies. *Molecular Ecology* 17, 2258–2275.

Nosil, P. (2008) Speciation with gene flow could be common. *Molecular ecology* 17, 2006–2008.

Nosil, P. (2012) *Ecological Speciation*. Oxford University Press.

Oksanen, J. (2012) Constrained ordination: tutorial with R and vegan. *R- packace Vegan*, 1–10.

Oksanen, J., Blanchet, F.G., M., F., Kindt, R., Legendre, P., McGlinn, D., Minchin, P.R., O'Hara, R.B., Simpson, G.L., Solymos, P., Henry, M., Stevens, H., Szoecs, E. & Wagner, H. (2017) Vegan: community ecology package. R package version 2.4-4.

Oliver, L.A., Prendini, E., Kraus, F. & Raxworthy, C.J. (2015) Systematics and biogeography of the Hylarana frog (Anura: Ranidae) radiation across tropical Australasia, Southeast Asia,

and Africa. *Molecular Phylogenetics and Evolution* 90, 176–192.

Orsini, L., Mergeay, J., Vanoverbeke, J. & De Meester, L. (2013)a) The role of selection in driving landscape genomic structure of the waterflea Daphnia magna. *Molecular Ecology* 22, 583–601.

Orsini, L., Vanoverbeke, J., Swillen, I., Mergeay, J. & De Meester, L. (2013)b) Drivers of population genetic differentiation in the wild: Isolation by dispersal limitation, isolation by adaptation and isolation by colonization. *Molecular Ecology* 22, 5983–5999.

Ostrom, J.H. (1979) Bird flight: how did it begin? *American Scientist* 67, 56–56.

Ozerov, M.Y., Veselov, A.E., Lumme, J., Primmer, C.R. & Moran, P. (2012) "Riverscape" genetics: river characteristics influence the genetic structure and diversity of anadromous and freshwater Atlantic salmon (Salmo salar) populations in northwest Russia. *Canadian Journal of Fisheries and Aquatic Sciences* 69, 1947–1958.

Patterson, N., Moorjani, P., Luo, Y., Mallick, S., Rohland, N., Zhan, Y., Genschoreck, T., Webster, T. & Reich, D. (2012) Ancient admixture in human history. *Genetics* 192, 1065–1093.

Payseur, B.A. & Rieseberg, L.H. (2016) A genomic perspective on hybridization and speciation. *Molecular Ecology*, 2337–2360.

Pearson, P.N., McMillan, I.K., Wade, B.S., Jones, T.D., Coxall, H.K., Bown, P.R. & Lear, C.H. (2008) Extinction and environmental change across the Eocene-Oligocene boundary in Tanzania. *Geology* 36, 179–182.

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

Petkova, D. (2017) Estimated Effective Migration Surface (EEMS). *Github*. Available from:

https://github.com/dipetkov/eems.

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

Phillips, S.B., Aneja, V.P., Kang, D. & Arya, S.P. (2006) Modelling and analysis of the atmospheric nitrogen deposition in North Carolina. *International Journal of Global Environmental Issues* 6, 231–252.

Pramuk, J.B., Robertson, T., Sites, J.W. & Noonan, B.P. (2008) Around the world in 10 million years: Biogeography of the nearly cosmopolitan true toads (Anura: Bufonidae). *Global Ecology and Biogeography* 17, 72–83.

Pyron, A.R. & Wiens, J.J. (2011) A large-scale phylogeny of Amphibia including over 2800 species, and a revised classification of extant frogs, salamanders, and caecilians. *Molecular Phylogenetics and Evolution* 61, 543–583.

Pyron, R.A., Hsieh, F.W., Lemmon, A.R., Emily, M. & Hendry, C.R. (2016) Integrating phylogenomic and morphological data to assess candidate species-delimitation models in brown and red-bellied snakes (Storeria). *Zoological Journal of the Linnean Society* 177, 937–949.

de Queiroz, K. (2005) Ernst Mayr and the modern concept of species. *Proceedings of the National Academy of Sciences of the United States of America* 102, 6600–6607.

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

Rabosky, D.L. (2014) Automatic detection of key innovations, rate shifts, and diversity-dependence on phylogenetic trees. *PLoS ONE* 9.

Rabosky, D.L., Mitchell, J.S. & Chang, J. (2017) Is BAMM flawed? Theoretical and practical

concerns in the analysis of multi-rate diversification models. *Systematic Biology* 66, 477–498.

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

Rannala, B. (2015) The art and science of species delimitation. *Current Zoology* 61, 846–853.

Reeves, P.A. & Richards, C.M. (2007) Distinguishing terminal monophyletic groups from reticulate taxa: performance of phenetic, tree-based, and network procedures. *Systematic Biology* 56, 302–20.

Reeves, P.A. & Richards, C.M. (2011) Species delimitation under the general lineage concept: An empirical example using wild North American hops (Cannabaceae: Humulus lupulus). *Systematic Biology* 60, 45–59.

Reusch, T.B.H. & Wood, T.E. (2007) Molecular ecology of global change. *Molecular Ecology* 16, 3973–3992.

Richardson, J.L., Brady, S.P., Wang, I.J. & Spear, S.F. (2016) Navigating the pitfalls and promise of landscape genetics. *Molecular Ecology* 25, 849–863.

Roček, Z. & Rage, J.-C. (2003) Evolution of anuran assemblages in the Tertiary and Quaternary of Europe, in the context of palaeoclimate and palaeogeography. *Amphibia-Reptilia* 24, 133–167.

Roelants, K., Gower, D.J., Wilkinson, M., Loader, S.P., Biju, S.D., Guillaume, K., Moriau, L. & Bossuyt, F. (2007) Global patterns of diversification in the history of modern amphibians. *Proceedings of the National Academy of Sciences of the United States of America* 104, 887–892.

Rolland, J., Condamine, F.L., Jiguet, F. & Morlon, H. (2014) Faster Speciation and Reduced

Extinction in the Tropics Contribute to the Mammalian Latitudinal Diversity Gradient. *PLoS Biology* 12.

Ronquist, F., Teslenko, M., van der Mark, P., Ayres, D.L., Darling, A., Hohna, S., Larget, B., Liu, L., Suchard, M. a. & Huelsenbeck, J.P. (2012) MrBayes 3.2: Efficient Bayesian phylogenetic inference and model choice across a large model space. *Systematic Biology* 61, 539–542.

Rosenberg, N.A. (2007) Statistical tests for taxonomic distinctiveness from observations of monophyly. *Evolution* 61, 317–323.

Rousset, F. (2008) GENEPOP'007: A complete re-implementation of the GENEPOP software for Windows and Linux. *Molecular Ecology Resources* 8, 103–106.

Rubin, B.E.R., Ree, R.H. & Moreau, C.S. (2012) Inferring phylogenies from RAD sequence data. *PLoS ONE* 7, 1–12.

Ruiz-Lopez, M.J., Barelli, C., Rovero, F., Hodges, K., Roos, C., Peterman, W.E. & Ting, N. (2016) A novel landscape genetic approach demonstrates the effects of human disturbance on the Udzungwa red colobus monkey (Procolobus gordonorum). *Heredity* 116, 167–176.

Rundle, H.D. & Nosil, P. (2005) Ecological speciation. *Ecology Letters* 8, 336–352.

Runemark, A., Hey, J., Hansson, B. & Svensson, E.I. (2012) Vicariance divergence and gene flow among islet populations of an endemic lizard. *Molecular Ecology* 21, 117–129.

Ruta, M., Pisani, D., Lloyd, G.T. & Benton, M.J. (2007) A supertree of temnospondyli: cladogenetic patterns in the most species-rich group of early tetrapods. *Proceedings of the royal society B.* 274, 3087–3095.

Savolainen, O., Lascoux, M. & Merilä, J. (2013) Ecological genomics of local adaptation. *Nature reviews. Genetics* 14, 807–20.

Schluter, D. (2009) Evidence for ecological speciation and its alternative. *Science* 323, 737–741.

Seehausen, O., Butlin, R.K., Keller, I., Wagner, C.E., Boughman, J.W., Hohenlohe, P.A., Peichel, C.L., Saetre, G.-P., Bank, C., Brännström, Å., Brelsford, A., Clarkson, C.S., Eroukhmanoff, F., Feder, J.L., Fischer, M.C., Foote, A.D., Franchini, P., Jiggins, C.D., Jones, F.C., Lindholm, A.K., Lucek, K., Maan, M.E., Marques, D.A., Martin, S.H., Matthews, B., Meier, J.I., Möst, M., Nachman, M.W., Nonaka, E., Rennison, D.J., Schwarzer, J., Watson, E.T., Westram, A.M. & Widmer, A. (2014) Genomics and the origin of species. *Nature Reviews Genetics* 15, 176–192.

Sexton, J.P., Hangartner, S.B. & Hoffmann, A.A. (2014) Genetic isolation by environment or distance: Which pattern of gene flow is most common? *Evolution* 68, 1–15.

Simpson, G.G. (1961) *Principles of animal taxonomy*. Columbia University Press, New York.

Sodhi, N.S., Koh, L.P., Brook, B.W. & Ng, P.K.L. (2004) Southeast Asian biodiversity: an impending disaster. *Trends in Ecology and Evolution* 19, 654–660.

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

Sork, V.L., Nason, J., Campbell, D.R. & Fernandez, J.F. (1999) Landscape approaches to historical and contemporary gene flow in plants. *Trends in Ecology and Evolution* 14, 219–224.

Sousa, V. & Hey, J. (2013) Understanding the origin of species with genome-scale data: modelling gene flow. *Nature reviews. Genetics* 14, 404–14.

Spear, S.F., Balkenhol, N., Fortin, M.J., McRae, B.H. & Scribner, K. (2010) Use of resistance surfaces for landscape genetic studies: Considerations for parameterization and analysis. *Molecular Ecology* 19, 3576–3591.

Stadler, T. & Smrckova, J. (2016) Estimating shifts in diversification rates based on higher-level phylogenies. *Biology Letters* 12, 20160273.

Storfer, A., Murphy, M.A., Evans, J.S., Goldberg, C.S., Robinson, S., Spear, S.F., Dezzani, R., Delmelle, E., Vierling, L. & Waits, L.P. (2007) Putting the 'landscape' in landscape genetics. *Heredity* 98, 128–142.

Streicher, J.W., Devitt, T.J., Goldberg, C.S., Malone, J.H., Blackmon, H. & Fujita, M.K. (2014) Diversification and asymmetrical gene flow across time and space: Lineage sorting and hybridization in polytypic barking frogs. *Molecular Ecology* 23, 3273–3291.

Stuart, B.L. (2008) The phylogenetic problem of Huia (Amphibia: Ranidae). *Molecular Phylogenetics and Evolution* 46, 49–60.

Sukumaran, J. & Knowles, L.L. (2017) Multispecies coalescent delimits structure, not species. *Proceedings of the National Academy of Sciences* 114, 1607–1612.

Sumarli, A.X., Grismer, L.L., Wood, P.L.J., Ahmad, A., Rizal, S., Ismail, L.H., Izam, N.A.M., Ahmad, N. & Linkem, C.W. (2016) The first riparian skink (Genus: Sphenomorphus Strauch, 1887) from Peninsular Malaysia and its relationship to other Indochinese and Sundaic species. *Zootaxa* 4173, 29.

Sun, J., Ni, X., Bi, S., Wu, W., Ye, J., Meng, J. & Windley, B.F. (2014) Synchronous turnover of flora, fauna, and climate at the Eocene–Oligocene Boundary in Asia. *Scientific Reports* 4, 7463.

Swofford, D.L. (2002) *PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods)*. Sinauer Associates, Sunderland, Massachusetts.

Tang, C.Q., Humphreys, A.M., Fontaneto, D. & Barraclough, T.G. (2014) Effects of phylogenetic reconstruction method on the robustness of species delimitation using single-

locus data. *Methods in Ecology and Evolution* 5, 1086–1094.

Team, Q.D. (2017) QGIS Geographic Information System. Open Source Geospatial Foundation Project.

Thorpe, R.S. (1975) Quantitative handling of characters useful in snake systematics with particular reference to intraspecific variation in the Ringed Snake Natrix natrix. *Biological Journal of the Linnean Society* 7, 27–43.

Thorpe, R.S. (1983) A review of the numerical methods for recognizing and analyzing racial differentiation. In: J. Felsenstein (Ed), *Numerical Taxonomy: Proceedings of a NATO Advanced Studies Institute NATO ASI series*. Springer Verlag, Berlin, Heidelberg, pp. 404–423.

Tobias, J.A., Seddon, N., Spottiswoode, C.N., Pilgrim, J.D., Fishpool, L.D.C. & Collar, N.J. (2010) Quantitative criteria for species delimitation. *Ibis* 152, 724–746.

Turan, C. (1999) A note on the examination of morphometric differentiation among fish populations: The Truss System. *Turkish Journal of Zoology* 23, 259–263.

Uribe-Convers, S. & Tank, D.C. (2015) Shifts in diversification rates linked to biogeographic movement into new areas: An example of a recent radiation in the andes. *American Journal of Botany* 102, 1854–1869.

Veach, V., Di Minin, E., Pouzols, F.M. & Moilanen, A. (2017) Species richness as criterion for global conservation area placement leads to large losses in coverage of biodiversity. *Diversity and Distributions* 23, 715–726.

Vences, M., Thomas, M., Bonett, R.M. & Vieites, D.R. (2005)a) Deciphering amphibian diversity through DNA barcoding: chances and challenges. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* 360, 1859–68.

Vences, M., Thomas, M., van der Meijden, A., Chiari, Y. & Vieites, D.R. (2005)b) Comparative performance of the 16S rRNA gene in DNA barcoding of amphibians. *Frontiers in zoology* 2, 5.

Via, S. & Hawthorne, D.J. (2002) The genetic architecture of ecological specialization: correlated gene effects on host use and habitat choice in pea aphids. *The American Naturalist* 159 Suppl, S76–S88.

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

Wang, I.J. & Bradburd, G.S. (2014) Isolation by environment. *Molecular Ecology* 23, 5649–5662.

Wang, I.J. & Summers, K. (2010) Genetic structure is correlated with phenotypic divergence rather than geographic isolation in the highly polymorphic strawberry poison-dart frog. *Molecular Ecology* 19, 447–458.

Weir, B.S. & Cockerham, C.C. (1984) Estimating F-Statistics for the analysis of population structure. *Evolution* 38, 1358–1370.

Whitlock, M.C. (2011) G'ST and D do not replace FST. *Molecular Ecology* 20, 1083–1091.

Wiens, J.J. & Donoghue, M.J. (2004) Historical biogeography, ecology and species richness. *Trends in Ecology and Evolution* 19, 639–644.

Wiens, J.J. & Penkrot, T.A. (2002) Delimiting species using DNA and morphological variation and discordant species limits in spiny lizards (Sceloporus). *Systematic Biology* 51, 69–91.

Wiens, J.J., Sukumaran, J., Pyron, R.A. & Brown, R.M. (2009) Evolutionary and biogeographic

origins of high tropical diversity in old world frogs (ranidae). *Evolution* 63, 1217–1231.

Wilcove, D.S., Giam, X., Edwards, D.P., Fisher, B. & Koh, L.P. (2013) Navjot's nightmare revisited: Logging, agriculture, and biodiversity in Southeast Asia. *Trends in Ecology and Evolution* 28, 531–540.

Wiley, E.O. (1978) The evolutionary species concept reconsidered. *Systematic Zoology* 27, 17–26.

Wood, P.L., Grismer, J.L., Grismer, L.L., Ahmad, N., Chan, K.O. & Bauer, A.M. (2009) Two new montane species of Acanthosaura gray, 1831 (Squamata: Agamidae) from peninsular Malaysia. *Zootaxa* 46, 28–46.

Woodburne, M.O., Rich, T.H. & Springer, M.S. (2003) The evolution of tribospheny and the antiquity of mammalian clades. *Molecular Phylogenetics and Evolution* 28, 360–385.

Wright, S. (1943) Isolation by distance. *Genetics* 28, 114–138.

Wright, S. (1951) The genetical structure of populations. *Annals of Eugenics* 15, 322–354.

Yang, Z. (2015) A tutorial of BPP for species tree estimation and species delimitation. *Current Zoology* 61, 854–865.

Yang, Z. & Rannala, B. (2010) Bayesian species delimitation using multilocus sequence data. *Proceedings of the National Academy of Sciences of the United States of America* 107, 9264–9269.

Yeung, C.K.L., Tsai, P.W., Chesser, R.T., Lin, R.C., Yao, C. Te, Tian, X.H. & Li, S.H. (2011) Testing founder effect speciation: Divergence population genetics of the Spoonbills Platalea regia and Pl. minor (Threskiornithidae, Aves). *Molecular Biology and Evolution* 28, 473–482.

Yuan, Z.Y., Zhou, W.W., Chen, X., Poyarkov, N.A., Chen, H.M., Jang-Liaw, N.H., Chou, W.H.,

Matzke, N.J., Iizuka, K., Min, M.S., Kuzmin, S.L., Zhang, Y.P., Cannatella, D.C., Hillis, D.M. & Che, J. (2016) Spatiotemporal Diversification of the True Frogs (Genus Rana): A Historical Framework for a Widely Studied Group of Model Organisms. *Systematic Biology* 65, 824–842.

Yumul, G.P., Dimalanta, C.B., Marquez, E.J. & Queaño, K.L. (2009) Onland signatures of the Palawan microcontinental block and Philippine mobile belt collision and crustal growth process: A review. *Journal of Asian Earth Sciences* 34, 610–623.

Zarza, E., Faircloth, B.C., Tsai, W.L.E., Bryson, R.W., Klicka, J. & McCormack, J.E. (2016) Hidden histories of gene flow in highland birds revealed with genomic markers. *Molecular Ecology* 25, 5144–5157.

Zhang, C., Zhang, D., Zhu, T. & Yang-, Z. (2011) Evaluation of a Bayesian Coalescent Method of Species Delimitation. *Systematic Biology* 60, 747–761.

Zhang, J., Kapli, P., Pavlidis, P. & Stamatakis, A. (2013) A general species delimitation method with applications to phylogenetic placements. *Bioinformatics* 29, 2869–2876.

APPENDICES

Appendix Table 1. Genetic samples used in Chapter 1 and their corresponding Genbank accession numbers

Taxa	Gene			
	16S	Cytochrome-b	RAG-1	Tyrosinase
Abavorana_luctuosa_kubah	DQ861313			
Abavorana_luctuosa_sabah	KF477635			
Abavorana_luctuosa_sedim	KF739008			
Amnirana_albolabris177	KR264110	KR264201		KR264501
Amnirana_albolabris664	KR264122	KR264209	KR264431	KR264507
Amnirana_albolabris861	KR264106	KR264197	KR264422	KR264499
Amnirana_albolabris863	KR264107	KR264198	KR264435	KR264500
Amnirana_albolabris991	KR264062	KR264152		
Amnirana_amnicola606		KR264153	KR264359	KR264437
Amnirana_amnicola621	KR264035	KR264125	KR264360	KR264438
Amnirana_amnicola818	KR264036	KR264126	KR264361	KR264439
Amnirana_darlingi	KR264121	KR264127	KR264430	KR264506
Amnirana_galamensis148	KR264092	KR264128		KR264486
Amnirana_galamensis225	KR264096	KR264129	KR264414	KR264490
Amnirana_galamensis262	KF991278	KR264130	KF991339	
Amnirana_galamensis740	AY341682	KR264131		AY341749
Amnirana_lepus220	KR264112		KR264423	KR264503
Amnirana_lepus732	AY014377		AY571641	
Amnirana_lepus985	KR264067	KR264157	KR264390	KR264458
Amnirana_lepus987	KR264068	KR264158	KR264391	KR264459
Amnirana_nicobariensis020	DQ347332		DQ347274	DQ347181
Amnirana_nicobariensis177	KR264094	KR264186	KR264412	KR264488

Amnirana_nicobariensis773	KU840574		KU840721	KU840777
Amnirana_nicobariensis938	KR264097	KR264189	KR264415	KR264491
Amnirana_nicobariensis_bario	DQ810283			
Amnirana_nicobariensis_gasing	KF738959			
Amnirana_nicobariensis_java	KF477631			
Amnirana_nicobariensis_sedim	KF738960			
Amnirana_nicobariensis_siberut	AB530581			
Amnirana_nicobariensis_sumatra				KR264472
Amnirana_nicobariensis_sumatra2	AB200965			
amolops_20981_baha		MF061713		
amolops_20988_baha		MF061715		
amolops_20989_baha		MF061716		
amolops_21008_rivendell		MF061717		
amolops_21011_rivendell		MF061719		
amolops_21012_rivendell		MF061720		
amolops_21077_tebu		MF061721		
amolops_21080_tebu		MF061722		
amolops_21130_jeriau		MF061723		
amolops_21138_belum		MF061726		
amolops_21141_belum		MF061728		
amolops_21152_sglong		MF061730		
amolops_21153_sglong		MF061731		
amolops_21237_jeriau		MF061733		
amolops_21250_genting		MF061734		
amolops_21263_robinson		MF061735		
amolops_21266_iskandar		MF061736		
amolops_21290_gombak		MF061737		
amolops_4894_lembing		MF061738		
amolops_5662_temengor		MF061739		
amolops_5663_temengor		MF061740		
amolops_574_bubu		MF061741		
Amolops_594_lentang		MF061742		
amolops_622_parit		MF061743		
amolops_7673_peta		MF061745		
amolops_8101_selai		MF061746		
amolops_9805_sendat		MF061747		
Amolops_akhaorum	FJ417160			FJ417324
Amolops_albispinus	KX507312			
Amolops_archotaphus	FJ417124			EU076756
Amolops_bellulus	FJ417126		FJ417266	FJ417299
amolops_BHF2_bkthijau				

amolops_BLF2_larut				
amolops_BLF3_larut				
Amolops_chunganensis	KU840605	KJ008460		
Amolops_compotrix	FJ417133		EF088235	EU076757
Amolops_cremnobatus_laos	FJ417143		FJ417282	FJ417315
Amolops_cremnobatus_vietnam	DQ204477			
Amolops_cremnobatus_vietnam2	AB211483		EF088236	
Amolops_cucae	FJ417146		EF088237	EU076759
Amolops_daiyunensis	DQ204479			
Amolops_daorum	FJ417148		EF088238	EU076760
Amolops_granulosus	AB211481	KJ008444		
Amolops_hainanensis	DQ204481		DQ019495	
Amolops_hongkongensis	AF206453			
Amolops_indoburmanensis	JF794458			
Amolops_iriodes	FJ417154			EU076761
Amolops_jinjiangensis	EF453741	KJ008379		
Amolops_kangtingensis	EF453742	KJ008423		
Amolops_liangshanensis	EF453743			
Amolops_lifanensis	DQ204482	KJ008458		DQ360065
Amolops_loloensis	AF206493	KJ008431		DQ360039
Amolops_mantzorum	AF315148	KJ008424	EF088240	EU076762
Amolops_marmoratus	DQ204485			EU076763
Amolops_mengyangensis	KR827704			
amolops_palongM1_palong				
Amolops_panhai	AB211487			
Amolops_ricketti	DQ204486		EF088242	EU076764
Amolops_spinapectoralis	DQ204488		EF088243	EU076765
amolops_SSM1_sedim				
Amolops_torrentis	DQ204489			
Amolops_tuberodepressus	FJ417162	KJ008426	FJ417294	FJ417327
Amolops_viridimaculatus	DQ204490	KJ008459		
Amolops_vitreus	FJ417164		FJ417295	
Amolops_wuyiensis	DQ204491		KP191587	KP191590
Babina_adenopleura		AB826434	KU840714	KU840783
Babina_chapaensis	KR264073	KR264163	KR264393	
Babina_daunchina	KU840597	KR264164	KU840723	KU840782
Babina_holsti		AB826407	AB777229	
Babina_lini	KF185066			
Babina_okinavana	AB058879	AB826435		

Babina_pleuraden	KR264059	KR264150	KR264384	
Babina_subaspera		AB826432		
Buergeria_buergeri	AY880444		AB612031	AB612033
Buergeria_oxycephala	EF564514	EU924592	EU924508	EU924564
Chalcorana_cf_eschatia_belatan	KF738983			
Chalcorana_cf_megalonesa_jerai	DQ650421			
Chalcorana_cf_megalonesa_tioman	KF738987			
Chalcorana_chalconota	KR264095	KR264187	KR264413	KR264489
Chalcorana_eschatia_myan	KR264060	KR264151		
Chalcorana_eschatia_thai	KR264084		KR264402	KR264476
Chalcorana_macrops	KR264098	KR264190	KR264416	KR264492
Chalcorana_megalonesa	KR264069	KR264176	KR264434	KR264478
Chalcorana_mocquardi	KR264099	KR264191	KR264417	KR264493
Chalcorana_parvacola	EF487455	KR264173	KR264401	KR264475
Chalcorana_raniceps	KF052067			
Chalcorana_rufipes	KR264081	KR264172	KR264400	
Clinotarsus_alticola424			EU076751	EU076789
Clinotarsus_alticola530	AB200961			
Clinotarsus_alticola633	KR827723			
Clinotarsus_alticola668	KR869788			
Clinotarsus_curtipes029	GU136111	AF249079	DQ347209	AF249180
Glandirana_emeljanovi	AF315155			AY322362
Glandirana_minima003	DQ359998			DQ360052
Glandirana_minima932	AF315153			
Glandirana_rugosa	JQ815306	JQ798757		
Glandirana_tientaiensis	KF185063			
Gracixalus_gracilipes	EF564523	EU924593	GQ285764	GQ285807
Huia_cavitympanum171	KF052041			
Huia_cavitympanum565	AB211489		KU840702	KU840768
Huia_masonii	DQ347313		EF088247	EU076770
Huia_sumatrana	AB211491		EF088249	EU076772
Humerana_humeralis	KU589224			
Humerana_lateralis	KR827775		EF088273	EU076800

Humerana_miopus	KR827778			
Humerana_sp_myan		KR264182	KR264409	KR264484
Hydrophylax_bahuvistara	KP867063			
Hydrophylax_leptoglossa	KR264065	KR264155	KR264388	
Hydrophylax_malabaricus	AB530579			
Hylarana_erythraea	KR264061	KR264156	KR264385	KR264509
Hylarana_macrodactyla	KR264071	KR264161	KR264392	KR264504
Hylarana_taipehensis	KR264089	KR264162	KR264408	KR264483
Indosylvirana_aurantiaca	AB530574			
Indosylvirana_caesari	KM068916			
Indosylvirana_doni	KM068926			
Indosylvirana_flavescens	KM068931			
Indosylvirana_indica	KM068951			
Indosylvirana_intermedia	KM068960			
Indosylvirana_magna	KM068965			
Indosylvirana_milleti_chantaburi	KR827825			
Indosylvirana_milleti_laos	KR827824			
Indosylvirana_milleti_viet	AF206490			
Indosylvirana_montana	KM068971			
Indosylvirana_serendipi	KM068979			
Indosylvirana_sreeni	KM068982			
Indosylvirana_temporalis	AB530578			
Indosylvirana_urbis	KM069005			
Liuixalus_romeri	EF564535	EU924598	EU924514	EU924570
Meristogenys_amoropalamus	KU840613	AB360073	AB360191	KU840772
Meristogenys_dyscritus	AB526616	AB360063	AB360187	
Meristogenys_jerboa	AB526608	AB526620	AB526662	

Meristogenys_kinabaluensis	AB526618	AB526630	AB526672	
Meristogenys_maryatiae	AB526611	AB526623	AB526665	
Meristogenys_orphnocnemis	DQ283147	AB360144	AB360205	EU076774
Meristogenys_phaeomerus	EU604210	AB526622		
Meristogenys_poecilus	AB526610	AB526717	EF088252	EU076775
Meristogenys_stenocephalus	AB526612	AB526626	AB526666	
Meristogenys_stigmachilus	AB526614	AB526629	AB526668	
Meristogenys_whiteheadi	AB526617		AB526671	EU076776
Nyctixalus_pictus	JN377342			
Odorrana_absita	EU861542		EF088245	EU076768
Odorrana_amamiensis	AB200947			
Odorrana_andersonii		KR264153	KR264386	DQ360049
Odorrana_anlungensis	KF185049			
Odorrana_aureola	DQ650567			
Odorrana_bacboensis	KT315385		EF088254	EU076777
Odorrana_banaorum	DQ650587			
Odorrana_chloronota	DQ650588		EF088256	EU076779
Odorrana_exiliversabilis	KF185056			
Odorrana_fengkaiensis	KT315380			
Odorrana_geminata	EU861548			
Odorrana_grahami	KF185051		EF088257	EU076780
Odorrana_graminea	KR338210		KP221672	KP191588
Odorrana_hainanensis	KF185032		KU840725	
Odorrana_hejiangensis	KR338103		KU840727	KU840788
Odorrana_hmongorum	EU861559		EF088258	EU076781
Odorrana_hosii_gombak	AB530591			
Odorrana_hosii_ranong	DQ650604			
Odorrana_hosii_swak	KR264087	KR264179	KR264432	KR264481
Odorrana_huanggangensis	KF185059			
Odorrana_ishikawae	AB576110			
Odorrana_jingdongensis	AF206483			
Odorrana_khalam	KU840606		EF088272	EU076783
Odorrana_kuangwuensis	KF185034			
Odorrana_leporipes	KF185036			

Odorrana_lipuensis	LC155911			
Odorrana_livida	EF453748		EF088260	EU076784
Odorrana_lungshengensis	KF185054			
Odorrana_margaretae	KF185035		EF088261	EU076785
Odorrana_morafkai	DQ650632		EF088263	EU076787
Odorrana_mutschmanni	KU356768			
Odorrana_narina	AB200948			
Odorrana_nasica			EF088264	EU076788
Odorrana_nasuta	KF185053			
Odorrana_schmackeri	KF185047		KP221673	KP191589
Odorrana_supranarina	AB200949			
Odorrana_swinhoana	KF185046			
Odorrana_tianmuii	KF185040			
Odorrana_tiannanensis	KF185044			
Odorrana_tormota	EF453754		EU076750	EU076766
Odorrana_utsunomiyaorum	AB200952			
Odorrana_versabilis	KF185055		KX269588	DQ360046
Odorrana_wuchuanensis	KF185043			
Odorrana_yizhangensis	KP710905		KU840730	
Papurana_arfaki	KR264048	KR264139		KR264497
Papurana_aurata	KR264101	KR264192	KR264419	KR264495
Papurana_daemeli	KR264056	KR264147	KR264382	KR264508
Papurana_garritor	KR264042	KR264133		
Papurana_jimiensis	KR264052	KR264143		
Papurana_kreffti	KR264050	KR264141		
Papurana_milneana	KR264044	KR264135		
Papurana_papua	KR264054	KR264145	KR264410	KR264485
Papurana_supragrisea	KR264055	KR264146	KR264383	
Papurana_volkerjane	KR264104	KR264195	KR264421	KR264498
Papurana_waliesa	KR264046	KR264137		
Pelophylax_bedriagae	AB640976	KU158382		
Pelophylax_bergeri	JN689222			
Pelophylax_caralitanus	AB640953	AB640981		
Pelophylax_cerigensis	DQ474196	DQ474144		
Pelophylax_chosenicus	EU386959	EU387059		
Pelophylax_cretensis	DQ474204	DQ474152		
Pelophylax_epeiroticus	AY147981	DQ474155		

Pelophylax_esculentus		AB980791		
Pelophylax_fukienensis		AB980780		GU978237
Pelophylax_grafi				KT879353
Pelophylax_hubeiensis	AF315137			
Pelophylax_kurtmuelleri	JF268493	DQ474176		
Pelophylax_lessonae	AY147982	EU047797		AY322347
Pelophylax_nigromaculatus	KF185062	AB980790	HQ902533	KU840778
Pelophylax_perezi	AY147985	DQ902145		KT879366
Pelophylax_plancyi	EU386960	AB980770		AB980678
Pelophylax_porosus_		AB980786		AB980674
Pelophylax_ridibundus	AB640933	AB980792		AB980690
Pelophylax_saharicus	KP177672	GU799108	KP177853	KT879338
Pulchrana_banjarana	KF739009			
Pulchrana_baramica	AB719231	KR264160	KR264398	
Pulchrana_centropeninsularis	EU604198			
Pulchrana_glandulosa	AB719223			
Pulchrana_grandocula	KF477677			
Pulchrana_laterimaculata	AB719228			
Pulchrana_mangyanum	KF477683			
Pulchrana_melanomenta	KF477692			
Pulchrana_moellendorffi	KF477694			
Pulchrana_picturata_penang	KF739007			
Pulchrana_picturata_sabah	KF477730	KR264159	KR264433	
Pulchrana_picturata_ZRC	KF477728			
Pulchrana_rawa	AB719222			
Pulchrana_siberu	KF477744			KF477593
Pulchrana_signata117	KR264086	KR264178	KR264406	KR264480
Pulchrana_signata_Bako	DQ835338			
Pulchrana_signata_matang	KF052074			
Pulchrana_signata_sabah	AB200963			
Pulchrana_signata_swak	AB719235			
Pulchrana_similis	KF477761	KR264168		
Rhacophorus_nigropalmatus	GQ204710	GQ204527	GQ204592	
Sanguirana_acai	KT881792			
Sanguirana_acai232	KT881795		KT881673	
Sanguirana_albotuberculata	KT881687			
Sanguirana_aurantipunctata728	KT881809		KT881676	KT881867
Sanguirana_aurantipunctata_ELR	KT881808		KT881677	KT881868
Sanguirana_everetti416	KT881689		KT881660	KT881841

Sanguirana_igorota	KT881827			
Sanguirana_luzonensis	KR264111	KR264202		KR264502
Sanguirana_sanguinea	KT881807			
Sanguirana_tipanan	KT881837			
Staurois_guttatus	KU840546		KU840739	KU840764
Staurois_latopalmatus	AB200966	AB259737	AB612049	EU076805
Staurois_natator	DQ347312	AB259734	DQ347250	DQ347155
Staurois_parvus	KU840548	AB259729	KU840741	EU076806
Staurois_tuberilinguis	KU840550	AB259731	KU840742	
Sylvirana_attigua	EU754872			
Sylvirana_cubitalis	KR264077		KR264396	KR264479
Sylvirana_faber767			EF088269	EU076796
Sylvirana_faber_camb	KR827801			
Sylvirana_faber_thai	KR827802			
Sylvirana_guentheri002	DQ360001			DQ360055
Sylvirana_guentheri109	KX269219	KX269363	KX269584	KX269810
Sylvirana_guentheri940	KR264039	KR264130		KR264440
Sylvirana_guentheri941	KR264040	KR264131	KR264365	
Sylvirana_guentheri_fujian	KF185060			
Sylvirana_guentheri_hanoi	AF206476			
Sylvirana_guentheri_viet	EU754860			
Sylvirana_latouchii	KF771284			
Sylvirana_latouchii_anhui		EU034927		
Sylvirana_latouchii_taipei	AB058880			
Sylvirana_maosonensis	KR264072	KR264162		
Sylvirana_menglaensis_chiangmai	KR827818			
Sylvirana_menglaensis_laos	KR827810			
Sylvirana_menglaensis_luangprabang	KR827820			
Sylvirana_menglaensis_thai	KR827816			
Sylvirana_mortenseni	KR264076	KR264166	KR264395	
Sylvirana_nigrovittata_camb	KR264079		KR264397	
Sylvirana_nigrovittata_msia	KF739002			
Sylvirana_nigrovittata_myan	KR264117	KR264205	KR264427	

Sylvirana_nigrovittata_thai	AB719238			
Sylvirana_nigrovittata_viet	KR264038			
Sylvirana_spinulosa	KR264093	KR264185	KR26441 1	KR264487
Theloderma_corticale	AF268256			DQ282904
Theloderma_gordoni	JN688167			
Theloderma_leprosum	AB847128			
Rana_amurensis	KX269203	KX269349	KX26956 8	KX269795
Rana_areolata	AY779229	KX269300	KX26951 4	KX269741
Rana_arvalis	KX269197	KX269344	KX26956 2	KX269789
Rana_asiatica	KX269200	KX269346	KX26956 5	KX269792
Rana_aurora	KX269212		KX26957 7	KX269803
Rana_berlandieri	AY779235	KX269301	KX26951 5	KX269742
Rana_blairi	AY779237			
Rana_boylii	KX269178	KX269299	KX26951 3	KX269740
Rana_bwana	AY779212			
Rana_capito	AY779231			
Rana_cascadae	KX269176	KX269302	KX26951 6	KX269743
Rana_catesbeiana	KX269208	KX269354	KX26957 3	DQ360044
Rana_chaochiaoensis	KX269192	KX269339	KX26955 7	KX269800
Rana_chensinensis	KX269186	KX269333	KX26955 1	KX269779
Rana_chiricahuensis	AY779225	KX269303	KX26951 7	KX269744
Rana_clamitans	AY779204	KX269304	KX26951 8	KX269745
Rana_coreana	KX269202	KX269348	KX26956 7	KX269794
Rana_culaiensis	KX269190	KX269337	KX26955 5	KX269783
Rana_dalmatina	KX269198		KX26956 3	KX269790
Rana_dunni	AY779222	KX269305	KX26951 9	KX269746
Rana_dybowskii	KX269188	KX269335	KX26955 3	KX269781

Rana_forreri	AY779233	GU184219	KX269520	KX269747
Rana_graeca	KX269199	KX269345	KX269564	KX269791
Rana_grylio	AY779201			
Rana_hanluica	KX269191	KX269338	KX269556	KX269784
Rana_heckscheri	AY779205			
Rana_huanrensis	KX269183	KX269330	KX269548	KX269776
Rana_iberica	KX269195	KX269342	KX269560	KX269787
Rana_japonica	KX269220	KX269364	KX269585	KX269811
Rana_jiemuxiensis	KX269221	KX269365	KX269586	KX269812
Rana_johnsi	KX269182	KX269328	KX269546	KX269774
Rana_juliani	AY779215		KX269521	KX269748
Rana_kobai	AB685778			
Rana_kukunoris	KX269185	KX269332	KX269550	KX269778
Rana_kunyuensis	KX269201	KX269347	KX269566	KX269793
Rana_latastei	AY147946	AY147967		
Rana_longicrus	KX269189	KX269336	KX269554	KX269782
Rana_luteiventris	KX269213	KX269358	KX269578	KX269804
Rana_macrocnemis	KX269194	KX269341	KX269559	KX269634
Rana_macroglossa	AY779243	KX269306	KX269522	KX269749
Rana_maculata	AY779207	KX269307	KX269523	KX269750
Rana_magnaocularis	AY779239	KX269308	KX269524	KX269751
Rana_montezumae	AY779223	KX269309	KX269525	KX269752
Rana_muscosa	AY779195			
Rana_neovolcanica	AY779236	KX269310	KX269526	KX269753
Rana_okaloosae	AY779203			
Rana_omeimontis	KX269193	KX269340	KX269558	KX269785

Rana_omiltemana	AY779238	KX269311	KX26952 7	KX269754
Rana_onca	AY779249			
Rana_ornativentris	KX269187	KX269334	KX26955 2	KX269780
Rana_palmipes	AY779211			
Rana_palustris	KX269207	KX269353	KX26957 2	KX269799
Rana_pipiens	AY779221			
Rana_pirica	KX269184	KX269331	KX26954 9	KX269777
Rana_psilonota	AY779217	KX269312	KX26952 8	KX269755
Rana_pustulosa	AY779220	KX269313		KX269756
Rana_pyrenaica	AY147950	AY147971		
Rana_sakuraii	KX269205	KX269351	KX26957 0	KX269797
Rana_sauteri	KX269204	KX269350	KX26956 9	KX269796
Rana_septentrionalis	KX269179	KX269314	KX26952 9	KX269757
Rana_sevosa	AY779230			
Rana_shuchinae	KX269210	KX269356	KX26957 5	DQ360057
Rana_sierrae	KX269211	KX269357	KX26957 6	KX269802
Rana_sierramadrensis	AY779216	KX269315	KX26953 0	KX269758
Rana_spectabilis	AY779227	KX269320	KX26953 7	KX269765
Rana_sphenocephala	AY779251	KX269321	KX26953 8	KX269766
Rana_sylvatica	KX269209	KX269355	KX26957 4	KX269801
Rana_tagoi	KX269214	KX269359	KX26957 9	KX269805
Rana_tarahumarae	AY779218	KX269322	KX26953 9	KX269767
Rana_temporaria	KX269196	KX269343	KX26956 1	KX269788
Rana_tlaloci	AY779234	KX269323	KX26954 0	KX269768
Rana_tsushimensis	KX269181	KX269329	KX26954 7	KX269775
Rana_uenoi	KX269177			

Rana_ulma	KX269360	KX269430	KX269655	KX269290
Rana_vaillanti	AY779214		KX269541	KX269769
Rana_vibicaria	KX269180	KX269324	KX269542	KX269770
Rana_warszewitschii	AY779209	KX269325	KX269543	KX269771
Rana_weiningensis	KX269217	KX269362	KX269582	KX269808
Rana_yavapaiensis	AY779240	KX269319	KX269535	KX269763
Rana_zhengi	KX269206	KX269352	KX269571	KX269798
Rana_zhenhaiensis	KX269218	JF939105	KX269583	KX269809
Rana_zweifeli	AY779219	KX269327	KX269545	KX269773
Rana_sp1	AY779213		KX269531	KX269759
Rana_sp2	AY779224	KX269316	KX269532	KX269760
Rana_sp3	AY779250	KX269317	KX269533	KX269761
Rana_sp4	AY779245			
Rana_sp5	AY779246	KX269318	KX269534	KX269762
Rana_sp6	AY779247			
Rana_sp7	AY779241	KX269326	KX269544	KX269772
Rana_sp8	AY779248	GU184220	KX269536	KX269764

Appendix Table 2. Raw morphological measurements and corresponding voucher specimens. See Materials and Methods for abbreviations.

Cat. #	Locality	sex	SVL	HL	HW	IND	SNL	FAL	FL	TBL	Fin3DW
ssm4	sedim	m	37.9	13.5	12.55	4.16	5.29	9.15	20.92	21.72	2.12
bhm5	sedim	m	35.12	12.6	12.3	3.52	5.64	8.88	18.9	19.95	2.45
bhm1	sedim	m	34.56	13.07	11.55	3.59	5.72	8	19.27	20.09	2.12
ssm5	sedim	m	34.72	12.63	11.36	3.84	5.26	8.54	19.29	20.5	2.18
bhm4	sedim	m	32.5	11.77	10.93	3.89	5.1	8.62	18.84	20.27	1.8
bhm8	sedim	m	34.8	12.58	11.85	3.58	5.45	7.89	18.83	20.08	2.14
ssm2	sedim	m	36.95	13.16	11.67	3.99	5.1	8.93	20.8	21.66	2
bhm2	sedim	m	35.52	13.78	12.32	3.76	5.77	9.26	19.95	21.19	2.46
bhm7	sedim	m	32.3	12.56	11.28	3.82	5.12	8.22	18.52	19.77	1.75
bhm6	sedim	m	35.81	13.2	12.77	3.82	5.68	8.35	20	20.81	2.4
bhm3	sedim	m	33	12.06	10.94	3.65	5.35	8.35	19.3	20.64	1.94
ssm3	sedim	m	35.65	13.31	11.35	3.77	5.26	8.38	20	21.31	2.1
ssm1	sedim	m	35	12.88	11.49	3.89	5.37	7.68	20.1	21.4	2.02
9639	sedim	m	33.65	12.81	11.66	3.54	5.2	7.85	19.5	21.34	1.8
ssf1	sedim	f	56	20	18.6	5.6	8.2	12	30	30.5	3.5
palong	palong	f	54.6	18.8	17.8	5.1	8	12.1	28	29.9	3.3
4893	lembing	f	52.7	19.8	17.6	6.1	8	11.9	28.8	31	3.6
4907	lembing	f	48.7	17.2	17	5.4	7.5	10.7	27.9	30.4	3.1
4899	lembing	f	54.8	19.3	18.4	6.1	7.9	13.1	31.1	33	3.6
4903	lembing	f	50.2	18.3	16.9	5.4	7.3	11.1	28.3	29.6	3.2
4905	lembing	f	56.7	20.3	18.3	5.8	8.5	11.5	29.7	32.7	3.5
4909	lembing	f	53.4	20.5	19.8	5.9	8.5	11.7	30.6	32.8	4.1
4932	lembing	f	53.9	19.6	17.9	5.4	8.5	12	28.7	32	3.4
4935	lembing	f	58	19.4	18.9	6	8.5	12.56	30.8	32.9	3.8
4913	lembing	f	52.3	17.7	17.2	5.4	7.9	12.5	28.2	30.9	3.5
4908	lembing	f	53.5	17.6	18.1	5.5	8.2	12.3	29.8	31.2	3.5
4898	lembing	f	53.1	20.1	18.5	5.8	8.2	12.6	29	31.2	3.4
4934	lembing	f	53.6	18.8	18.9	6	8.1	11.23	29.1	30.7	3.3
4894	lembing	f	50.7	18.9	18.2	5.7	7.8	12	29.3	31.4	3.4
4896	lembing	f	51.1	19	17.9	5.8	8	11.7	30.1	31.6	3.4
4895	lembing	f	56	19.8	18.6	5.78	8.5	12.1	30.8	33	3.54
4906	lembing	m	34	13.7	11.9	4	5.8	9.3	20.5	22.9	2.2
4904	lembing	m	35.9	14	12.5	4	5.3	9	21.4	23	2.4

4902	lembing	m	35.1	13.2	12.8	3.8	5.3	9.2	21	22	2.3
9926	fraser	f	60.5	22.8	21.5	6.4	9	14.7	32.8	35.4	3.6
9760	fraser	f	60	23.1	22	6.4	9	14	33.4	36.1	3.9
9759	fraser	f	64.6	24.6	23.3	6.6	9.9	15.7	37.7	38.6	4.2
21231	fraser	f	62.5	23.6	21.8	7	9.4	14.4	34.8	35.1	3.8
21229	fraser	f	59.7	22.4	20.9	5.9	8.7	14.3	33.5	34	3.5
21228	fraser	f	62.3	23.5	22.6	6.21	9.7	15.1	34.2	37.3	3.7
21230	fraser	f	63.8	25	24.4	7.1	10	15.3	37.1	38.1	4
21232	fraser	f	62.5	23.2	21.2	6	9.2	13.8	33.3	35.8	3.7
21238	fraser	m	38.5	14.8	13.7	4.5	6.3	7.8	22	23.3	2.4
21239	fraser	m	38.8	15.1	13	4	6.12	9.6	21	22.1	2.6
21236	fraser	m	39.8	14.9	13.7	4.5	5.8	10.9	23.2	23.4	2.3
21235	fraser	m	40.4	14.7	13	4	6	10.1	22	23.7	2.2
21233	fraser	m	38.2	14.7	13.7	4.5	6.3	10	22.4	23.4	2.6
21234	fraser	m	35.2	13.8	12	4	5.7	9.1	20.3	21.9	1.9
21252	genting	f	59	21.7	20.1	6.4	9	13.8	33.7	35.1	3.5
21253	genting	f	59.2	22.1	19.5	6.1	9	13.8	29.7	34.9	3.5
21251	genting	f	66	23.9	22.9	6.6	9.9	14.8	35.1	36	3.7
1.11127	genting	f	57.8	21	19.6	6	8.7	12.5	32.4	34.3	3.3
1.11126	genting	m	39.5	14.3	13.2	4.3	6	9.9	22.2	23.1	2.1
21247	genting	m	37.6	13.9	12.8	4.4	5.9	9.6	22.8	23.4	2.3
21246	genting	m	39.2	14.3	12.7	4.4	6.8	10.6	22.2	24.1	2.2
1.3528	sekayu	f	55.3	19.8	17.5	5.4	8.3	12.7	28.5	30.6	3.3
1.3527	sekayu	f	52.5	19.3	17.2	5.4	8	12.8	27.5	30.8	3.1
9958	chiling	f	48.6	17.8	16.2	5	7.6	10.8	26.3	26.5	2.6
9957	chiling	f	46.4	17.3	16.4	5.1	7.2	10.5	26.1	27.5	2.6
9954	chiling	m	35.9	13.1	12	4.1	5.5	8.6	19.2	19.5	2.1
9948	chiling	m	32.5	12.5	10.8	3.6	5.2	7.3	17.7	19.1	1.7
9951	chiling	m	35.3	13.4	12	4.4	5.4	8	19.7	20.2	2.1
9949	chiling	m	34.2	13.2	11.7	3.9	5.4	7.5	19.6	20.1	1.8
9952	chiling	m	33.7	12.8	11.8	3.7	5.4	9.2	18.9	19.4	2.1
9955	chiling	m	33.9	12.5	11	3.8	5.2	8.2	18.5	20	1.8
9950	chiling	m	35.1	12.6	11.2	4	5.5	8.6	18.5	19.8	1.9
9953	chiling	m	33.3	13.3	11	3.8	5.4	8.3	18.2	19.3	1.9
9947	chiling	m	31.5	12.3	10.7	3.3	5	8	18.3	19	1.9
21077	tebu	f	57.4	21.2	19.2	5.8	9	13	33	34.2	3.8
21088	tebu	f	56.4	21.4	20.3	6	8.7	12.9	30.6	31.6	3.7
21078	tebu	f	57.5	21.2	19.3	6.1	8.4	13	30.4	32	3.7
21079	tebu	f	58.2	21.4	20.2	6.2	9.1	13	31.3	33.6	3.8
21090	tebu	f	52.8	19.9	17.8	5.9	8.3	11.9	30.4	31.3	3.2
21081	tebu	f	55.7	20.5	18.9	6.4	8.4	11.2	31.3	34.4	3.5
21053	stong	f	54.8	19.7	18.2	5.6	8	12.6	30.9	33.2	3.2
21089	tebu	f	52.6	20.4	18.1	5.7	8.7	12.2	29.1	31.5	3.4

21080	tebu	m	37.2	13.8	12.7	3.9	6.1	8.6	22.3	22.5	2.5
21294	gombak	f	52.2	13.1	17.7	4.8	7.7	11.3	28.7	29.2	3.2
21293	gombak	f	55	20.5	18.7	5.5	8.2	12.2	29.3	30.1	3.1
21286	gombak	m	36.3	13.5	12.2	4	5.6	9.4	9.9	21.3	1.9
21289	gombak	m	36.8	13.1	12	4.1	5.5	8.9	20.7	21.6	2
21288	gombak	m	36.5	13.5	12.3	4.1	5.5	8.9	21.1	21.6	2
21287	gombak	m	36.2	13.6	12.3	4.3	5.5	9.2	20.9	21.5	2
21143	jeli	f	52	19.2	16.8	5.4	8.3	11.7	30.3	32.3	3.2
21142	jeli	f	57.1	20.1	19	5.8	9.1	11.7	30.7	33.1	3.5
21144	jeli	m	38.4	14.9	13.4	4	5.9	9.2	22.1	22.6	2.9
21146	jeli	m	36.4	14.6	13.1	3.9	5.9	8.2	21.4	21.8	2.1
21149	jeli	m	36.8	12.2	12.3	3.9	5.9	7.9	20.9	21.4	2.3
21145	jeli	m	36.9	14.7	13.1	4	5.9	8.5	21.3	21.9	2.9
21148	jeli	m	37.4	14.3	12.8	4.1	6	8.4	22	22.8	2.2
21147	jeli	m	34.8	13.5	12	3.8	5.8	8.3	20.4	21.2	2.1
7665	endau	f	46.3	16.5	15.8	4.8	7.4	9.8	25.6	26.7	3
7671	endau	f	47	17.9	16.2	5.2	7.4	9.3	24.7	25.8	3.3
7684	endau	f	45.8	17.1	15.5	4.9	7	8.6	22.9	25.1	3
7687	endau	f	46.8	17.5	15.7	4.9	7.6	10.3	24.7	25.4	2.7
8096	endau	f	46.6	16.7	14.3	4.9	7	9.2	23.1	25	2.6
7686	endau	m	31.5	12	10.6	3.5	5.3	6.6	17.4	18.2	1.7
7673	endau	m	32.1	12.1	11	3.5	5.3	6.7	17.2	17.4	1.8
8100	endau	m	28.7	11.5	10.4	3.5	4.8	7	15.4	17.4	1.9
7672	endau	m	32.7	12.5	11.3	3.3	5	7.1	17.7	18.6	2.1
1.10217	johor	f	47	17.2	15.4	5.3	7.8	10.1	25	27.2	2.8
1.10219	johor	m	32.1	12.6	11.1	3.8	5	7.1	19.5	19.5	1.8
1.10218	johor	m	29	12	10.1	3.6	5.1	6.5	17.2	17.6	1.7
1.10761	johor	m	32.3	12.6	11.2	3.6	5.5	7.1	18.5	18.9	1.8
1.10596	sendat	f	55.3	20.8	19.3	5.7	8.5	13	31.2	32.2	3.2
1.10597	sendat	f	58	21.9	20.1	5.8	9.1	13.3	32.5	33.4	3.5
9805	sendat	f	56.9	20.7	18.6	5.8	8.3	12.5	30	30.6	3.3
1.10059	uluyam	m	36.3	14	12.6	4.2	5.7	9	20.7	21.2	2.1
1.1006	uluyam	m	37.2	14.2	12.5	4.3	6.1	9.21	21.2	21.6	2.1
21264	iskandar	f	52.3	20.1	17.8	5.2	8.2	11.4	28.6	20.2	3.2
21266	iskandar	f	51.6	19.5	17.3	5.3	8	11.5	26.2	29.2	3.3
21268	iskandar	m	34.8	13.2	11.7	3.9	5.6	9.3	19.4	20.6	2.1
21267	iskandar	m	35.9	13.5	12.3	3.9	5.9	8.8	20	20.8	2.4
21262	robinson	m	39.4	14.9	12.9	3.7	5.8	10.4	22.1	23	2.4
21261	robinson	m	39.6	14.7	13.2	3.9	6	10.5	23.3	24	2.2
21263	robinson	m	36.9	14.1	12.4	4.1	5.8	10.3	21.9	23.1	2.2
5662	temengor	f	56.7	20.9	19.7	5.5	8.5	13.8	30	32.3	3.2
5663	temengor	m	32	12.4	11.6	3.9	5.7	9.2	19.4	20.8	2.1
21152	sglong	f	60.8	21.6	19.6	5.9	9.1	13.5	32.2	34.2	3.1

21153	sglong	f	62.7	22.7	20.8	6.4	9	13.3	34.2	35.2	3.7
21154	sglong	f	58.2	21.2	19.3	5.9	8.7	13.5	31	33.2	3.1
21136	belum	m	37.4	14.1	12.6	4	6.1	9.5	20.9	21.7	2.3
21137	belum	m	33.3	12.8	11.9	3.5	5.9	7.6	19.7	20.9	2.2
21141	belum	m	38.1	14.2	12.8	4	5.9	9.2	20.9	21.9	2.2
21138	belum	m	37.5	13.9	12.3	4.1	5.6	9.3	20.9	21.3	2.4
21135	belum	m	36.6	13.4	11.8	4.2	5.4	8.7	20.1	21.2	2.2
21139	belum	m	37	13.8	12.5	3.7	5.7	8	19.6	20.5	2.2
21140	belum	m	34.9	13.3	12.2	3.5	5.3	8.8	20.6	21.32	2.2
21012	stong	f	59.8	21.8	20	5.9	8.8	13.7	31.9	33.23	3.8
21010	stong	f	62.1	22.3	21	6.5	9.4	14.4	33.6	34.6	3.4
20988	stong	f	51.9	20.53	18.8	5.9	8.4	12.5	29.2	30.9	3.1
20989	stong	f	49.9	19.5	17.5	5.4	7.7	11.7	28.1	29.7	3.2
20990	stong	m	37.7	14.3	12.5	3.8	5.3	9.8	20.3	22.4	2.1
21011	stong	m	38.9	14.5	13	4.1	5.6	9.2	22	22.8	2.2
20978	stong	f	48.5	18.9	16.9	5.1	8	11.3	27.2	29.1	3.3
20981	stong	f	50.3	19.3	17.6	5.5	8	11.8	29.7	31.2	3.8
20979	stong	f	45	19.1	16.6	5.1	7.8	11	29.2	30.6	3.4
20980	stong	f	49.2	19	17	4.9	7.9	10.8	27.1	28.6	3.5
21009	stong	m	34.3	13.5	12.3	3.9	5.54	7.9	20.3	20.7	1.9
21008	stong	m	37	14.2	12.5	4.1	5.7	9.2	20.4	21.9	2.2
21013	stong	m	36.2	14	12.3	3.8	5.7	8.8	20.4	20.9	2.3

Appendix Table 3. List of morphological and genetic samples used in this study with corresponding GenBank accession numbers for mitochondrial samples. Genomic SNP data is available from the Dryad repository listed in the manuscript.

					GenBank #
Samples	**Locality**	**Source**	**Morpho**	**SNPs**	**16S**
Huia cavitympanum	Sabah, Borneo, Malaysia	GenBank			AB211489
Meristogenys jerboa	Jambi, Sumatra, Indonesia	GenBank			AB211493
Meristogenys orphnoncnemis	Sabah, Borneo, Malaysia	GenBank			AB211494
Amolops cremnobatus	Vietnam	GenBank			AB211483
Amolops cremnobatus	Vietnam	GenBank			FJ417143
Amolops cremnobatus	Vietnam	GenBank			AF206458
Amolops cremnobatus	Vietnam	GenBank			DQ204477
Amolops cremnobatus	Vietnam	GenBank			DQ204478
Amolops spinapectoralis	Vietnam	GenBank			DQ204488
Amolops torrentis	China	GenBank			DQ204489
Amolops hainanensis	China	GenBank			DQ204481
Amolops panhai	Thailand	GenBank			AB211487
Amolops panhai	Thailand	GenBank			AB211488
Amolops panhai	Thailand	GenBank			KR827705
Amolops panhai	Thailand	GenBank			KR827706
Amolops daiyunensis	China	GenBank			DQ204479
Amolops lifanensis	China	GenBank			DQ204482
Amolops viridimaculatus	China	GenBank			DQ204490
Amolops vitreus		GenBank			FJ417165
Amolops compotrix		GenBank			FJ417142
Amolops cucae		GenBank			FJ417146
Amolops bellulus	China	GenBank			DQ204473
Amolops archotaphus		GenBank			FJ17125
Amolops akhaorum		GenBank			FJ417158
Amolops ricketti		GenBank			AF458117
Amolops wuyiensis	China	GenBank			AB211476
Amolops daorum		GenBank			FJ417147
Amolops tuberodepressus		GenBank			FJ417162
Amolops iriodes		GenBank			FJ417154
Amolops wuyiensis	China	GenBank			KF771291

Amolops	mantzorum	China	GenBank			KF771289
Amolops	kangtingensis	China	GenBank			KF771287
Amolops	loloensis	China	GenBank			KF771288
Amolops	chunganensis	China	GenBank			KF771285
Amolops	mengyangensis	China	GenBank			KR827704
Amolops	indoburmanensis		GenBank			JF794458
Amolops	indoburmanensis		GenBank			JF794471
Amolops	indoburmanensis		GenBank			JF794429
Amolops	indoburmanensis		GenBank			JF794442
Amolops	indoburmanensis		GenBank			JF794438
Amolops	marmoratus	Thailand	GenBank			AB211486
Amolops	marmoratus		GenBank			JF794455
Amolops	marmoratus		GenBank			DQ204485
Amolops	marmoratus		GenBank			JF794432
Amolops	marmoratus		GenBank			EU861541
Amolops	larutensis	Kuala Lumpur, Malaysia	GenBank			AB511293
Amolops	larutensis	Gombak, Selangor, Malaysia	GenBank			AB530592
Amolops	larutensis	Malaysia	GenBank			AB211484
Amolops	larutensis	Thailand	GenBank			AB211485
21141		Belum, Perak	This study	x		MF061728
21140		Belum, Perak	This study	x		MF061727
21138		Belum, Perak	This study	x		MF061726
21137		Belum, Perak	This study	x		MF061725
21135		Belum, Perak	This study	x		MF061724
21136		Belum, Perak	This study	x		
21139		Belum, Perak	This study	x		
21142		Jeli, Kelantan	This study	x		
21143		Jeli, Kelantan	This study	x		
21144		Jeli, Kelantan	This study	x		
21145		Jeli, Kelantan	This study	x		
21146		Jeli, Kelantan	This study	x		
21147		Jeli, Kelantan	This study	x		
21148		Jeli, Kelantan	This study	x		
21149		Jeli, Kelantan	This study	x		
BHF2		Bukit Hijau, Kedah	This study			MF061748
BHM1		Bukit Hijau, Kedah	This study	x	x	
BHM2		Bukit Hijau, Kedah	This study	x	x	
BHM3		Bukit Hijau, Kedah	This study	x		
BHF1		Bukit Hijau, Kedah	This study			
BHM7		Bukit Hijau, Kedah	This study	x		
BHM5		Bukit Hijau, Kedah	This study	x		
BHF3		Bukit Hijau, Kedah	This study			

BHM4	Bukit Hijau, Kedah	This study	x				
BHM6	Bukit Hijau, Kedah	This study	x	x			
BHM8	Bukit Hijau, Kedah	This study	x				
BLF2	Bukit Larut, Perak	This study		x	MF061749		
BLF3	Bukit Larut, Perak	This study			MF061750		
BLF4	Bukit Larut, Perak	This study		x			
LSUHC610.larut	Bukit Larut, Perak	This study					
LSUHC611.larut	Bukit Larut, Perak	This study					
LSUHC613.larut	Bukit Larut, Perak	This study					
LSUHC604.larut	Bukit Larut, Perak	This study					
LSUHC605.larut	Bukit Larut, Perak	This study					
BLM3.larut	Bukit Larut, Perak	This study					
LSUHC606.larut	Bukit Larut, Perak	This study					
BLM1.larut	Bukit Larut, Perak	This study					
BLM4.larut	Bukit Larut, Perak	This study			x		
LSUHC601.larut	Bukit Larut, Perak	This study					
palongM1	Bukit Palong, Kedah	This study	x			MF061751	
10051	Cameron Highlands (Iskandar Falls)	This study					
21263	Cameron Highlands (Iskandar Falls)	This study	x			MF061735	
21264	Cameron Highlands (Iskandar Falls)	This study	x		x		
21265	Cameron Highlands (Iskandar Falls)	This study			x		
21266	Cameron Highlands (Iskandar Falls)	This study			x	MF061736	
21267	Cameron Highlands (Iskandar Falls)	This study	x		x		
21268	Cameron Highlands (Iskandar Falls)	This study	x		x		
21269	Cameron Highlands (Iskandar Falls)	This study			x		
21270	Cameron Highlands (Iskandar Falls)	This study			x		
622	Cameron Highlands (Parit Falls)	This study				MF061743	
21263	Cameron Highlands (Robinson Falls)	This study	x		x	MF061735	
21261	Cameron Highlands (Robinson Falls)	This study	x		x		
21262	Cameron Highlands (Robinson Falls)	This study	x		x		
LSUHC619.robinson	Cameron Highlands (Robinson Falls)	This study					
LSUHC616.robinson	Cameron Highlands (Robinson Falls)	This study					
LSUHC620.robinson	Cameron Highlands (Robinson Falls)	This study					
LSUHC621.robinson	Cameron Highlands (Robinson Falls)	This study					
7665	Endau-Rompin, Johor (Peta)	This study	x		x		
7684	Endau-Rompin, Johor (Peta)	This study	x				
7686	Endau-Rompin, Johor (Peta)	This study	x		x		
7687	Endau-Rompin, Johor (Peta)	This study	x				
7671	Endau-Rompin, Johor (Peta)	This study	x				

7672	Endau-Rompin, Johor (Peta)	This study	x		
7673	Endau-Rompin, Johor (Peta)	This study	x		MF061745
8095	Endau-Rompin, Johor (Selai)	This study		x	
8096	Endau-Rompin, Johor (Selai)	This study	x	x	
8100	Endau-Rompin, Johor (Selai)	This study	x	x	
8101	Endau-Rompin, Johor (Selai)	This study			MF061746
LSUHC7686.selai	Endau-Rompin, Johor (Selai)	This study			
LSUHC8098.selai	Endau-Rompin, Johor (Selai)	This study			
LSUHC8099.selai	Endau-Rompin, Johor (Selai)	This study			
LSUHC9926.fraser	Fraser's Hill, Pahang (back road)	This study	x		
LSUHC9925.fraser	Fraser's Hill, Pahang (back road)	This study			
21228	Fraser's Hill, Pahang (Jeriau Falls)	This study	x	x	
21229	Fraser's Hill, Pahang (Jeriau Falls)	This study	x	x	
21230	Fraser's Hill, Pahang (Jeriau Falls)	This study	x	x	
21231	Fraser's Hill, Pahang (Jeriau Falls)	This study	x	x	
21232	Fraser's Hill, Pahang (Jeriau Falls)	This study	x	x	
21233	Fraser's Hill, Pahang (Jeriau Falls)	This study	x	x	
21234	Fraser's Hill, Pahang (Jeriau Falls)	This study	x	x	
21235	Fraser's Hill, Pahang (Jeriau Falls)	This study	x	x	
21236	Fraser's Hill, Pahang (Jeriau Falls)	This study	x	x	
21237	Fraser's Hill, Pahang (Jeriau Falls)	This study		x	MF061733
21238	Fraser's Hill, Pahang (Jeriau Falls)	This study	x	x	
21239	Fraser's Hill, Pahang (Jeriau Falls)	This study	x	x	
10056	Fraser's Hill, Pahang (Jeriau Falls)	This study		x	
10058	Fraser's Hill, Pahang (Jeriau Falls)	This study			
10057	Fraser's Hill, Pahang (Jeriau Falls)	This study			
10060	Fraser's Hill, Pahang (Jeriau Falls)	This study			
8068	Fraser's Hill, Pahang (road up)	This study			
8069	Fraser's Hill, Pahang (road up)	This study		x	
8070	Fraser's Hill, Pahang (road up)	This study			
9958	Fraser's Hill, Pahang (Sungai Chiling)	This study	x		
9957	Fraser's Hill, Pahang (Sungai Chiling)	This study	x		
9947	Fraser's Hill, Pahang (Sungai Chiling)	This study	x		
9948	Fraser's Hill, Pahang (Sungai Chiling)	This study	x		
9949	Fraser's Hill, Pahang (Sungai Chiling)	This study	x		
9951	Fraser's Hill, Pahang (Sungai Chiling)	This study	x		
9954	Fraser's Hill, Pahang (Sungai Chiling)	This study	x		
9955	Fraser's Hill, Pahang (Sungai Chiling)	This study	x		
9953	Fraser's Hill, Pahang (Sungai Chiling)	This study	x		
9950	Fraser's Hill, Pahang (Sungai Chiling)	This study	x		
9952	Fraser's Hill, Pahang (Sungai Chiling)	This study	x		
6447	Fraser's Hill, Pahang (The Gap)	This study		x	MF061744
6445	Fraser's Hill, Pahang (The Gap)	This study			

6448	Fraser's Hill, Pahang (The Gap)	This study			
6450	Fraser's Hill, Pahang (The Gap)	This study			
6449	Fraser's Hill, Pahang (The Gap)	This study			
6444	Fraser's Hill, Pahang (The Gap)	This study			
6446	Fraser's Hill, Pahang (The Gap)	This study			
9759	Fraser's Hill, Pahang	This study	x		
9760	Fraser's Hill, Pahang	This study	x		
21246	Genting Highlands, Pahang	This study	x	x	
21247	Genting Highlands, Pahang	This study	x	x	
21248	Genting Highlands, Pahang	This study		x	
21249	Genting Highlands, Pahang	This study		x	
21250	Genting Highlands, Pahang	This study		x	MF061734
21251	Genting Highlands, Pahang	This study	x	x	
21252	Genting Highlands, Pahang	This study	x	x	
21253	Genting Highlands, Pahang	This study	x	x	
6594	Genting Highlands, Pahang	This study			
21286	Gombak, Selangor	This study	x	x	
21287	Gombak, Selangor	This study	x	x	
21288	Gombak, Selangor	This study	x	x	
21289	Gombak, Selangor	This study	x	x	
21290	Gombak, Selangor	This study		x	MF061737
21291	Gombak, Selangor	This study		x	
21292	Gombak, Selangor	This study		x	
21293	Gombak, Selangor	This study	x	x	
21294	Gombak, Selangor	This study	x	x	
LSUHC589.gombak	Gombak, Selangor	This study			
LSUHC592.gombak	Gombak, Selangor	This study			
LSUHC591.gombak	Gombak, Selangor	This study			
LSUHC587.gombak	Gombak, Selangor	This study			
LSUHC590.gombak	Gombak, Selangor	This study			
LSUHC588.gombak	Gombak, Selangor	This study			
574	Gunung Bubu, Perak	This study			MF061741
LSUHC575.bubu	Gunung Bubu, Perak	This study			
LSUHC582.bubu	Gunung Bubu, Perak	This study			
LSUHC580.bubu	Gunung Bubu, Perak	This study			
LSUHC577.bubu	Gunung Bubu, Perak	This study			
LSUHC573.bubu	Gunung Bubu, Perak	This study			
LSUHC578.bubu	Gunung Bubu, Perak	This study			
LSUHC579.bubu	Gunung Bubu, Perak	This study			
LSUHC581.bubu	Gunung Bubu, Perak	This study			
LSUHC576.bubu	Gunung Bubu, Perak	This study			
LSUHC585.bubu	Gunung Bubu, Perak	This study			
20989	Gunung Stong, Kelantan (Baha)	This study	x	x	MF061716

20978	Gunung Stong, Kelantan (Baha)	This study	x	x	
20979	Gunung Stong, Kelantan (Baha)	This study	x	x	MF061712
20980	Gunung Stong, Kelantan (Baha)	This study	x	x	
20981	Gunung Stong, Kelantan (Baha)	This study	x	x	MF061713
20982	Gunung Stong, Kelantan (Baha)	This study		x	
20983	Gunung Stong, Kelantan (Baha)	This study		x	
20984	Gunung Stong, Kelantan (Baha)	This study		x	
20985	Gunung Stong, Kelantan (Baha)	This study		x	
20986	Gunung Stong, Kelantan (Baha)	This study		x	
20987	Gunung Stong, Kelantan (Baha)	This study		x	MF061714
20988	Gunung Stong, Kelantan (Baha)	This study	x	x	
20990	Gunung Stong, Kelantan (Baha)	This study	x	x	
21053	Gunung Stong, Kelantan (Base)	This study	x	x	
21008	Gunung Stong, Kelantan (Rivendell)	This study	x	x	MF061717
21009	Gunung Stong, Kelantan (Rivendell)	This study	x	x	
21010	Gunung Stong, Kelantan (Rivendell)	This study	x	x	MF061718
21011	Gunung Stong, Kelantan (Rivendell)	This study	x	x	MF061719
21012	Gunung Stong, Kelantan (Rivendell)	This study	x	x	MF061720
21013	Gunung Stong, Kelantan (Rivendell)	This study	x	x	
21077	Gunung Tebu, Terengganu (Base camp)	This study	x		MF061721
21078	Gunung Tebu, Terengganu (Base camp)	This study	x	x	
21079	Gunung Tebu, Terengganu (Base camp)	This study	x	x	
21080	Gunung Tebu, Terengganu (Base camp)	This study	x	x	MF061722
21081	Gunung Tebu, Terengganu (Base camp)	This study	x	x	
21088	Gunung Tebu, Terengganu (Base camp)	This study	x	x	
21089	Gunung Tebu, Terengganu (Base camp)	This study	x	x	
21090	Gunung Tebu, Terengganu (Base camp)	This study	x	x	
9787	Lata Tembakah, Terengganu	This study		x	
9789	Lata Tembakah, Terengganu	This study		x	
9790	Lata Tembakah, Terengganu	This study			
9791	Lata Tembakah, Terengganu	This study			
9792	Lata Tembakah, Terengganu	This study			
594	Lentang, Pahang	This study			MF061741
595	Lentang, Pahang	This study			
4893	Sg. Lembing, Pahang	This study	x		
4894	Sg. Lembing, Pahang	This study	x		MF061738
9909	Sg. Lembing, Pahang	This study			
4896	Sg. Lembing, Pahang	This study	x		

4905	Sg. Lembing, Pahang	This study	x		
4906	Sg. Lembing, Pahang	This study	x		
4898	Sg. Lembing, Pahang	This study	x		
4899	Sg. Lembing, Pahang	This study	x	x	
4900	Sg. Lembing, Pahang	This study		x	
4903	Sg. Lembing, Pahang	This study	x		
4909	Sg. Lembing, Pahang	This study	x		
4901	Sg. Lembing, Pahang	This study		x	
4902	Sg. Lembing, Pahang	This study	x	x	
4904	Sg. Lembing, Pahang	This study	x	x	
4897	Sg. Lembing, Pahang	This study			
4895	Sg. Lembing, Pahang	This study	x		
4907	Sg. Lembing, Pahang	This study	x	x	
4908	Sg. Lembing, Pahang	This study	x	x	
4932	Sg. Lembing, Pahang	This study	x		
4934	Sg. Lembing, Pahang	This study	x		
4935	Sg. Lembing, Pahang	This study	x		
4913	Sg. Lembing, Pahang	This study	x		
21150	Sg. Long, Kelantan	This study			MF061729
21154	Sg. Long, Kelantan	This study	x		MF061732
21152	Sg. Long, Kelantan	This study	x		MF061730
21153	Sg. Long, Kelantan	This study	x		MF061731
9805	Sg. Sendat, Selangor	This study	x	x	MF061747
SSM1	Sungai Sedim, Kedah	This study	x		MF061752
LSUHC9639.sedim	Sungai Sedim, Kedah	This study	x		
LSUHC9636.sedim	Sungai Sedim, Kedah	This study			
LSUHC9637.sedim	Sungai Sedim, Kedah	This study			
SSM5.sedim	Sungai Sedim, Kedah	This study	x	x	
SSM4.sedim	Sungai Sedim, Kedah	This study	x	x	
SSM2.sedim	Sungai Sedim, Kedah	This study	x	x	
SSM3.sedim	Sungai Sedim, Kedah	This study	x	x	
SSF1	Sungai Sedim, Kedah	This study	x		
5662	Temenggor, Perak	This study	x		MF061739
5663	Temenggor, Perak	This study	x	x	MF061740
1.3528	sekayu	This study	x		
1.3527	sekayu	This study	x		
1.11127	genting	This study	x		
1.11126	genting	This study	x		
1.10761	johor	This study	x		
1.10597	sendat	This study	x		
1.10596	sendat	This study	x		
1.10219	johor	This study	x		
1.10218	johor	This study	x		

1.10217	johor	This study	x
1.1006	uluyam	This study	x
1.10059	uluyam	This study	x

Appendix Table 4. Results of the PCA analysis on adjusted morphological data.

Male	**PC1**	**PC2**	**PC3**	**PC4**	**PC5**	**PC6**	**PC7**	**PC8**	**PC9**
Standard deviation	1.868	1.254	1.062	0.86	0.788	0.735	0.681	0.509	0.432
Proportion of Variance	0.388	0.175	0.125	0.082	0.069	0.06	0.052	0.029	0.021
Cumulative Proportion	0.388	0.563	0.688	0.77	0.839	0.899	0.951	0.979	1
Eigenvalues	3.49	1.572	1.129	0.739	0.621	0.54	0.464	0.259	0.187
SVL	-0.271	0.294	-0.494	0.228	-0.222	0.67	-0.216	0.007	0.004
HL	-0.384	-0.171	0.13	0.489	0.166	0.117	0.647	-0.293	-0.148
HW	-0.331	-0.341	0.302	-0.05	-0.661	0.058	0.087	0.361	0.319
IND	-0.308	0.419	0.055	0.348	-0.299	-0.577	-0.293	-0.309	0.046
SNL	-0.216	-0.533	-0.227	0.403	0.358	-0.183	-0.437	0.319	0.02
FAL	-0.358	0.274	-0.397	-0.268	0.073	-0.328	0.367	0.519	-0.223
FL	-0.362	0.121	0.554	-0.169	0.145	0.213	-0.322	0.116	-0.579
TBL	-0.431	0.16	0.126	-0.304	0.463	0.085	-0.052	-0.132	0.662
Fin3DW	-0.288	-0.44	-0.337	-0.481	-0.167	-0.086	-0.091	-0.535	-0.224
Female									
Standard deviation	2.073	1.28	1.036	0.841	0.72	0.541	0.462	0.417	0.292
Proportion of Variance	0.477	0.182	0.119	0.079	0.058	0.033	0.024	0.019	0.009
Cumulative Proportion	0.477	0.659	0.779	0.857	0.915	0.948	0.971	0.991	1
Eigenvalues	4.296	1.639	1.074	0.707	0.519	0.293	0.213	0.174	0.085
SVL	-0.299	-0.015	0.284	-0.845	0.002	-0.05	0.332	-0.035	0.041
HL	0.285	0.31	0.409	0.047	0.741	-0.301	0.009	0.058	-0.1
HW	-0.263	0.471	0.436	0.027	-0.102	0.542	-0.444	0.065	0.115
IND	0.403	0.203	-0.245	-0.219	0.104	0.35	0.004	-0.744	-0.048
SNL	0.46	0.008	0.053	-0.031	-0.057	0.166	0.245	0.263	0.79
FAL	0.43	-0.06	-0.145	-0.294	0.001	0.376	-0.065	0.559	-0.496
FL	0.317	0.256	0.445	0.172	-0.594	-0.182	0.356	-0.104	-0.29
TBL	0.039	0.627	-0.435	-0.25	-0.203	-0.457	-0.264	0.158	0.1
Fin3DW	-0.317	0.424	-0.302	0.235	0.181	0.284	0.658	0.142	-0.088

Appendix Table 5. Results of the hybrid index analysis on the putative hybrid population (West 1). The populations East 1 and Larutensis were designated as reference and alternate populations respectively. H-values close to 1 indicate affinity with the reference population while values close to 0 indicate affinity to the alternate population.

Individual	H	ln(likelihood)	Low	Up
5663	0.003	-87.872	0.001	0.005
20988	0.002	-86.017	0.001	0.005
20989	0.002	-90.591	0.001	0.005
20990	0.006	-49.594	0.002	0.015
21010	0.002	-82.544	0.001	0.005
21012	0.004	-115.538	0.002	0.007
21011	0.549	-572.706	0.518	0.58

Appendix Table 6. Fastsimcoal2 results for the five models tested. Historical period (HP) 0: after L/W1 diverged (~3 mya to present); HP 1: after W3/W4 diverged, before L/W1 diverged (~6–3 mya ago); HP 2: after W2/W3/W4 diverged, before W3/W4 diverged (7–6 mya ago); HP 3: after L/W1 and W2/W3/W4 diverged, before W2/W3/W4 diverged (8–7 mya ago). Bootstrap support values are in parentheses. Asterisks(*) denotes scenarios with historical migration, where migration rates in these categories persist through time for populations that have not coalesced (e.g. W2 to W4 is fixed from historical period 0 through period 1, as neither of these populations had coalesced with other populations).

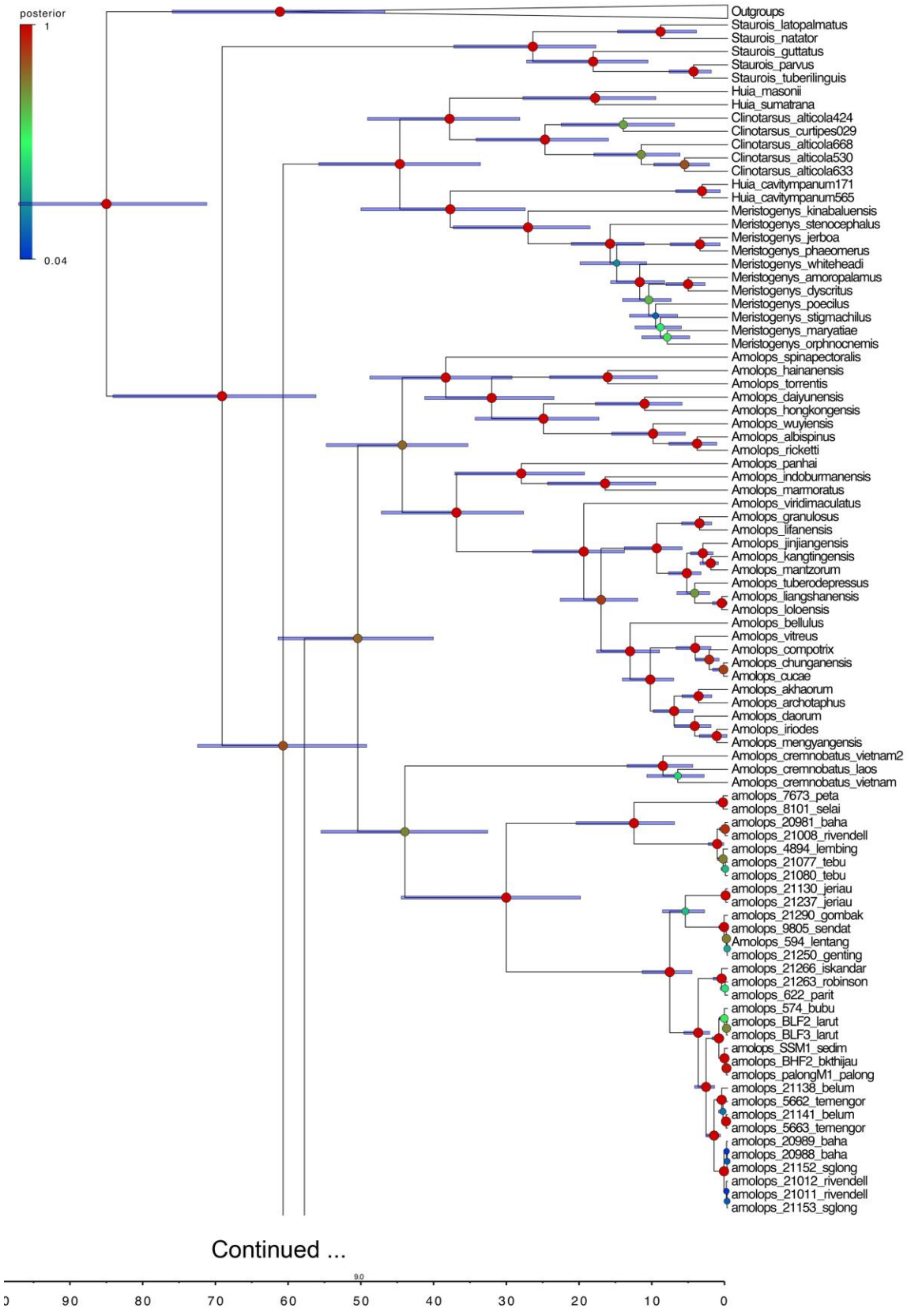
	Full migration (asymmetrical)	Contemporary migration only	Full migration (symmetrical)	Historical migration only	No migration
MaxEstLhood	-7,059.40	-7,763.50	-7,854.33	-8,659.63	-11,883.09
Total no. parameters	37	25	21	23	5
AIC	32583.73569	35802.23888	36212.51833	39925.0471	54733.6633
AIC Rank	1	2	3	4	5
∆i	0.00	3218.50	3628.78	7341.31	22149.93
HP 0 migration parameters (N_m)					
L to W1	0.5814 (0.3001--0.7947)	7.0222	0.1381	X	X
W1 to L	0.2377 (0.177--1.764)	0.1155	0.1639	X	X
L to W2	0.3331 (0.2272--0.4437)	0.9819	0.1094	X	X
W2 to L	0.3026 (0.212--0.8442)	0.1114	0.2602	X	X
W1 to W2	5.4939 (3.7697--7.6407)	3.0400	0.7145	X	X
W2 to W1	3.6221 (2.7419--15.8796)	1.6304	1.4314	X	X
W2 to W3*	1.0291 (0.6674--9.6109)	1.5182	2.0972	13.5186	X
W3 to W2*	0.6391 (0.3572--35.6598)	6.0901	0.6833	2.2373	X
W3 to W4*	5.7762 (0.4876--40.0024)	1.4479	1.2455	13.8007	X
W4 to W3*	1.2426 (0.7441--21.5485)	2.2224	2.3887	13.3971	X

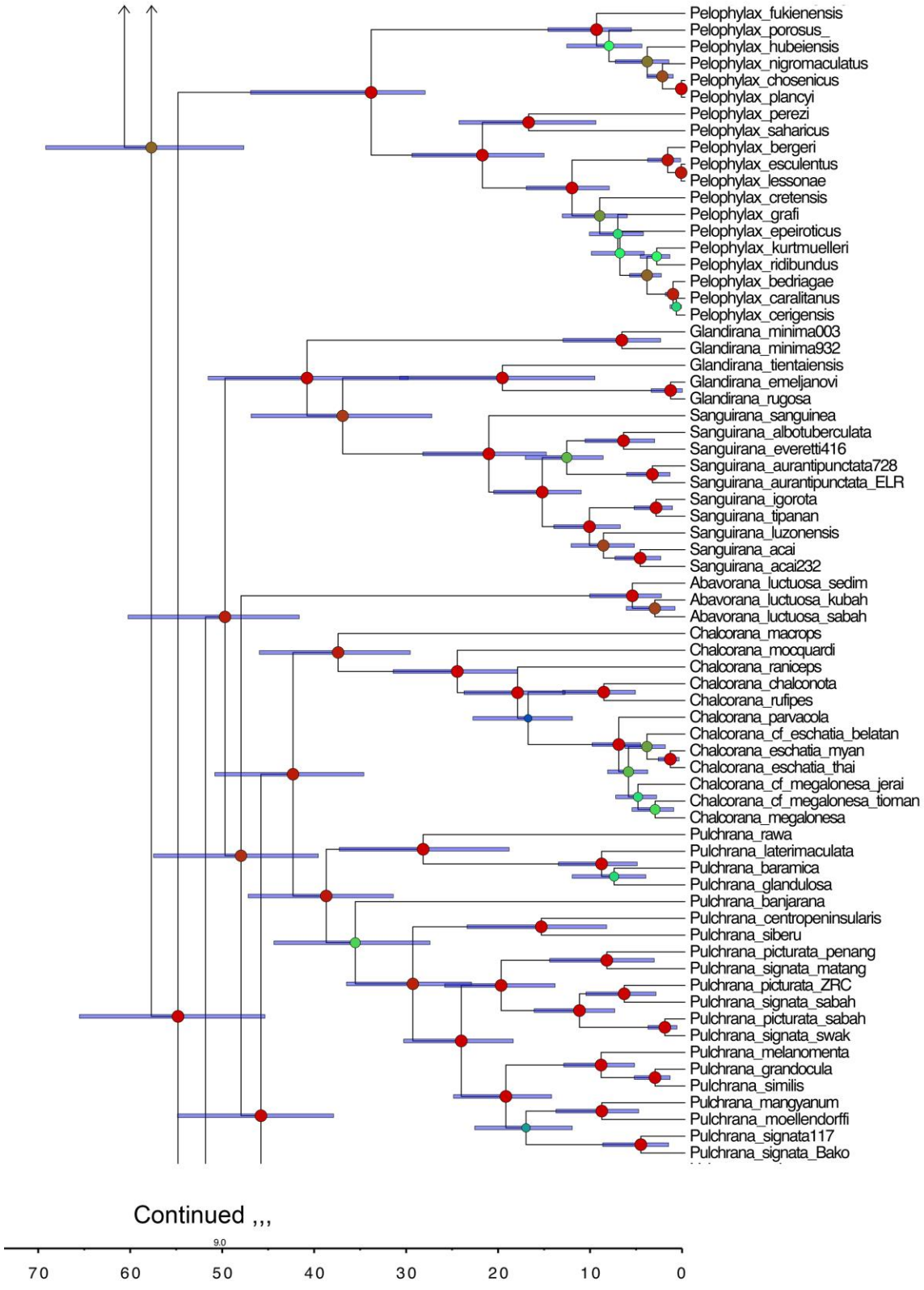
Continued ..

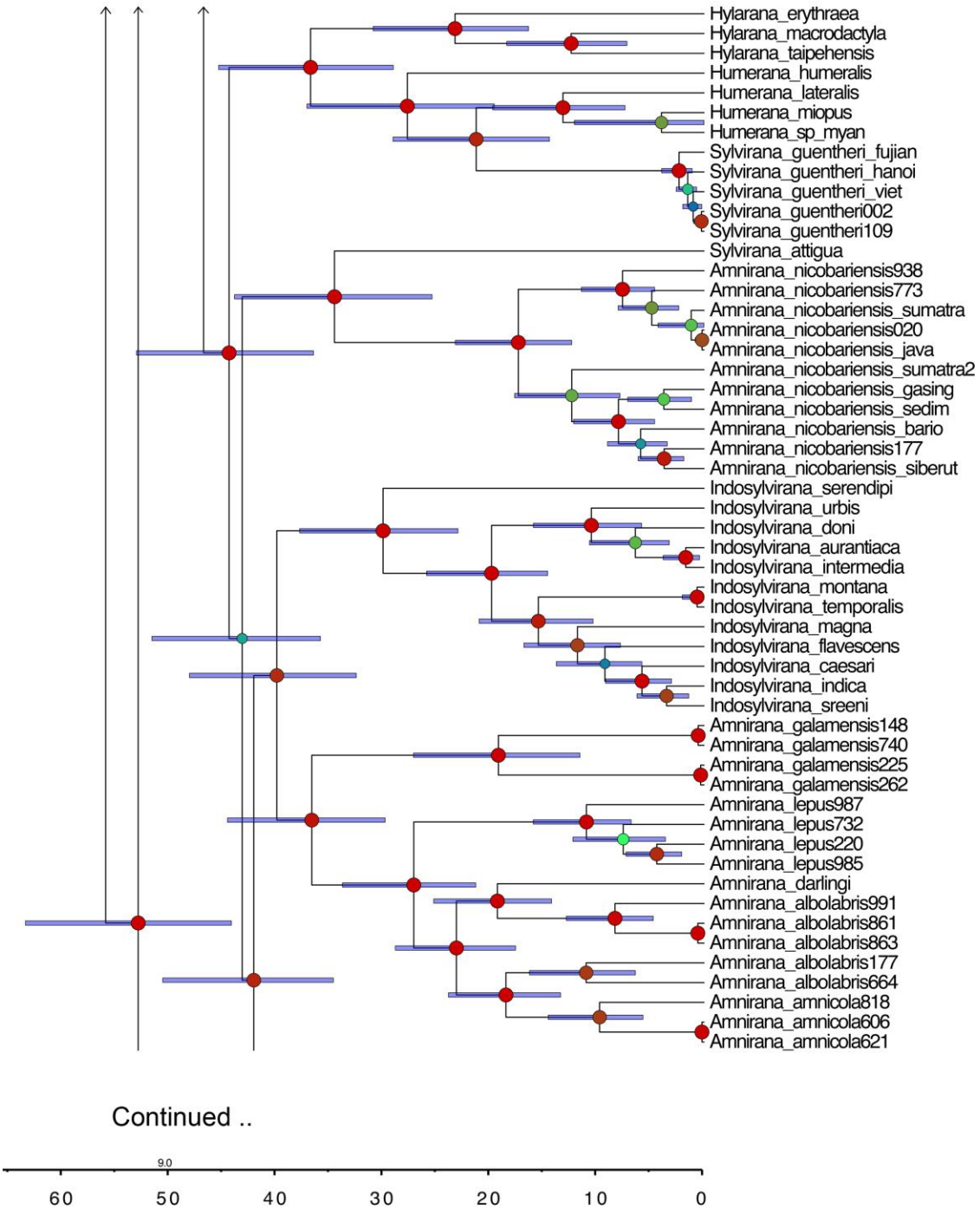
HP 1 migration parameters (N_m)					
Anc(L/W1) to W2*	0.6878 (0.1665--2.8119)	X	19.4837	1.6285	X
W2 to Anc(L1/W1)*	1591.2665 (132.2697--3021.2141)	X	85.3623	14.5032	X

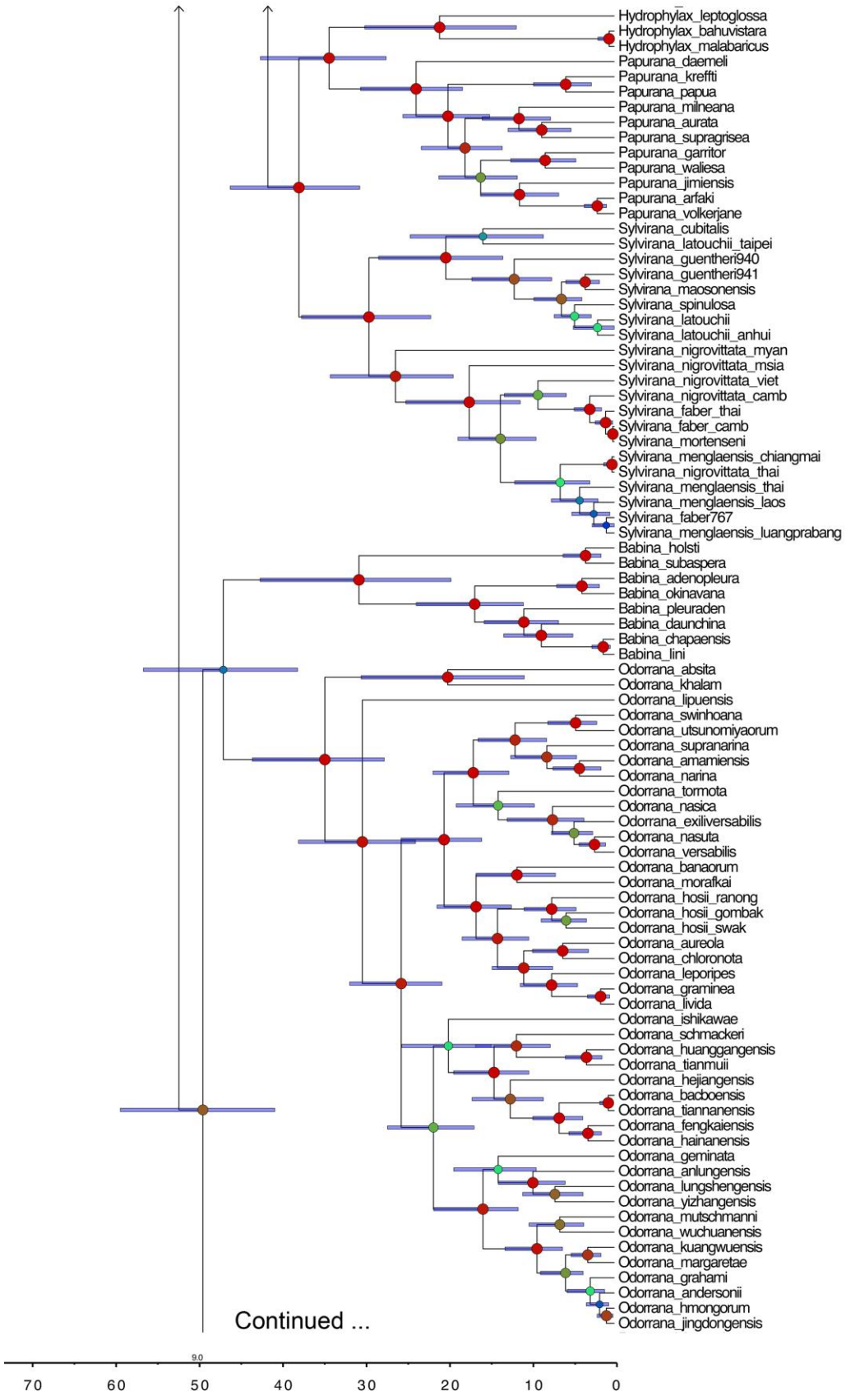
Anc(L/W1) to W3	240.3918 (173.1812--1070.6029)	X	13.1726	5.7871	X
W3 to Anc(L/W1)	13.7266 (6.0649--16.537)	X	18.8036	6.0155	X
Anc(L/W1) to W4	33.8203 (0.2995--119.4184)	X	33.4083	12.3160	X
W4 to Anc(L/W1)	141.6577 (52.3717--2727.1648)	X	91.4646	5.3057	X
HP 2 migration parameters (N_m)					
Anc(L/W1) to Anc(W3/W4)	14.9898 (8.7874--22.2458)	X	13.2968	4.9375	X
Anc(W3/W4) to Anc(L/W1)	11.9066 (7.18--17.9916)	X	12.4759	11.4628	X
W2 to Anc(W3/W4)	20.9471 (15.9191--58.2197)	X	11.2546	13.3294	X
Anc(W3/W4) to W2	0.6875 (0.5575--0.8011)	X	2.4103	12.7191	X
HP 3 migration parameters (N_m)					
Anc(L/W1) to Anc(W2/W3/W4)	11.3596 (8.1149--16.1302)	X	13.3400	7.9794	X
Anc(W2/W3/W4) to Anc(L/W1)	16.2122 (8.577--25.1072)	X	10.3088	3.3556	X

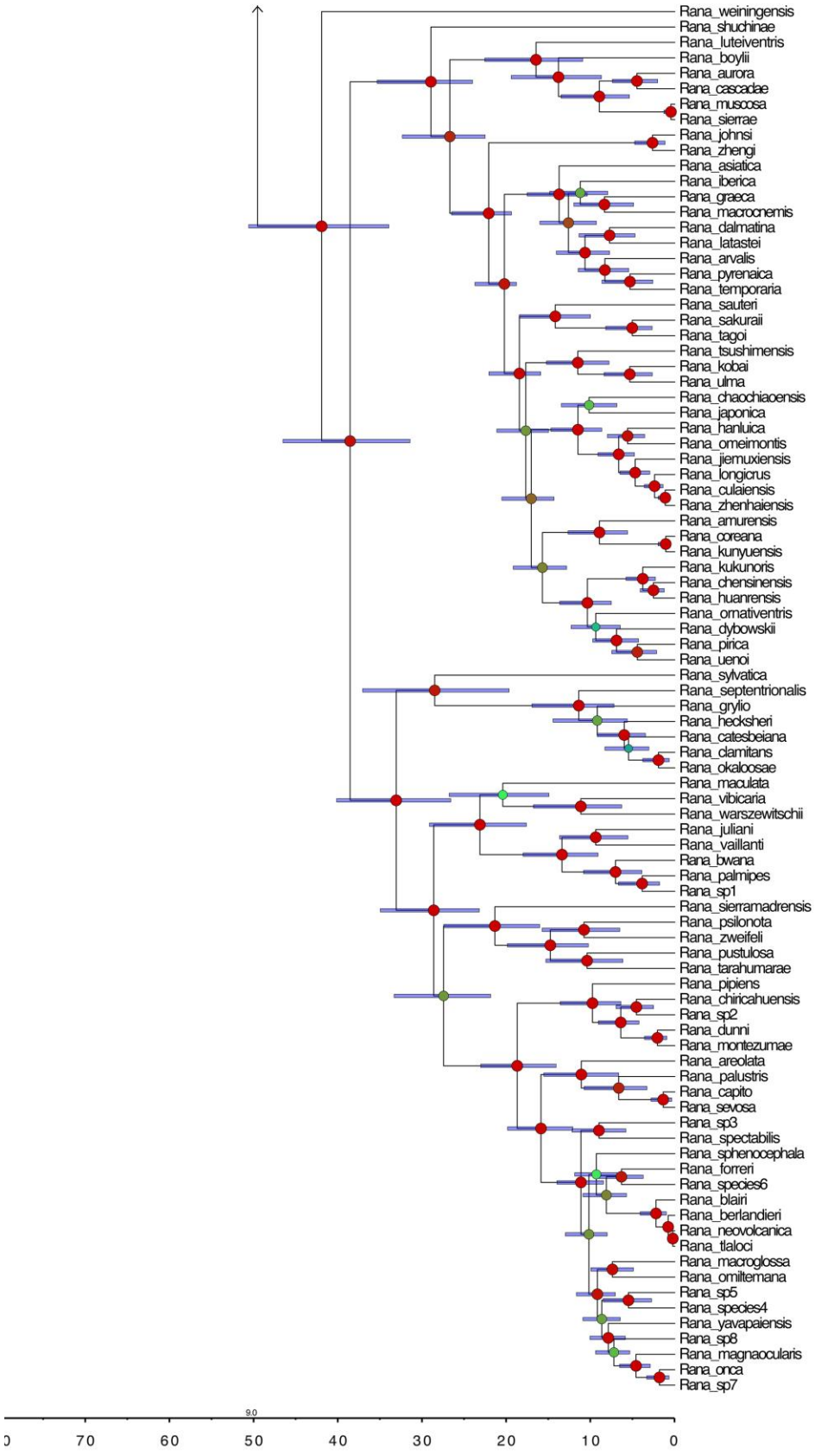
L = Larutensis; W1 = West1; W2 = West2; W3 = West3; W4 = West 4; Anc = Ancestor


posterior
1
0.04
Outgroups
Staurois_latopalmatus
Staurois_natator
Staurois_guttatus
Staurois_parvus
Staurois_tuberilinguis
Huia_masonii
Huia_sumatrana
Clinotarsus_alticola424
Clinotarsus_curtipes029
Clinotarsus_alticola668
Clinotarsus_alticola530
Clinotarsus_alticola633
Huia_cavitympanum171
Huia_cavitympanum565
Meristogenys_kinabaluensis
Meristogenys_stenocephalus
Meristogenys_jerboa
Meristogenys_phaeomerus
Meristogenys_whiteheadi
Meristogenys_amoropalamus
Meristogenys_dyscritus
Meristogenys_poecilus
Meristogenys_stigmachilus
Meristogenys_maryatiae
Meristogenys_orphnocnemis
Amolops_spinapectoralis
Amolops_hainanensis
Amolops_torrentis
Amolops_daiyunensis
Amolops_hongkongensis
Amolops_wuyiensis
Amolops_albispinus
Amolops_ricketti
Amolops_panhai
Amolops_indoburmanensis
Amolops_marmoratus
Amolops_viridimaculatus
Amolops_granulosus
Amolops_lifanensis
Amolops_jinjiangensis
Amolops_kangtingensis
Amolops_mantzorum
Amolops_tuberodepressus
Amolops_liangshanensis
Amolops_loloensis
Amolops_bellulus
Amolops_vitreus
Amolops_compotrix
Amolops_chunganensis
Amolops_cucae
Amolops_akhaorum
Amolops_archotaphus
Amolops_daorum
Amolops_iriodes
Amolops_mengyangensis
Amolops_cremnobatus_vietnam2
Amolops_cremnobatus_laos
Amolops_cremnobatus_vietnam
amolops_7673_peta
amolops_8101_selai
amolops_20981_baha
amolops_21008_rivendell
amolops_4894_lembing
amolops_21077_tebu
amolops_21080_tebu
amolops_21130_jeriau
amolops_21237_jeriau
amolops_21290_gombak
amolops_9805_sendat
Amolops_594_lentang
amolops_21250_genting
amolops_21266_iskandar
amolops_21263_robinson
amolops_622_parit
amolops_574_bubu
amolops_BLF2_larut
amolops_BLF3_larut
amolops_SSM1_sedim
amolops_BHF2_bkthijau
amolops_palongM1_palong
amolops_21138_belum
amolops_5662_temengor
amolops_21141_belum
amolops_5663_temengor
amolops_20989_baha
amolops_20988_baha
amolops_21152_sglong
amolops_21012_rivendell
amolops_21011_rivendell
amolops_21153_sglong
Continued ...
90
90
80
70
60
50
40
30
20
10
0

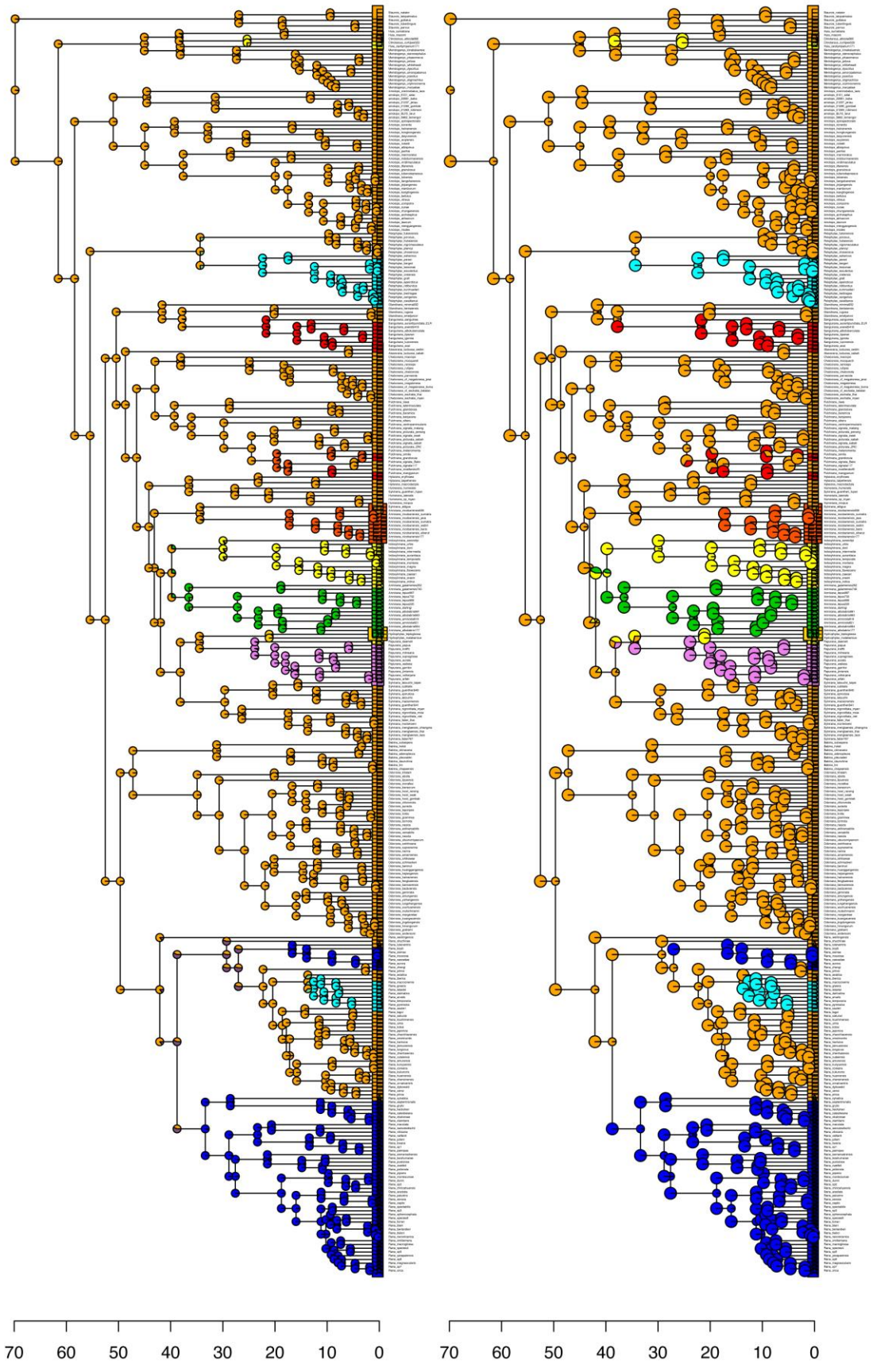

Pelophylax_fukienensis
Pelophylax_porosus_
Pelophylax_hubeiensis
Pelophylax_nigromaculatus
Pelophylax_chosenicus
Pelophylax_plancyi
Pelophylax_perezi
Pelophylax_saharicus
Pelophylax_bergeri
Pelophylax_esculentus
Pelophylax_lessonae
Pelophylax_cretensis
Pelophylax_grafi
Pelophylax_epeiroticus
Pelophylax_kurtmuelleri
Pelophylax_ridibundus
Pelophylax_bedriagae
Pelophylax_caralitanus
Pelophylax_cerigensis
Glandirana_minima003
Glandirana_minima932
Glandirana_tientaiensis
Glandirana_emeljanovi
Glandirana_rugosa
Sanguirana_sanguinea
Sanguirana_albotuberculata
Sanguirana_everetti416
Sanguirana_aurantipunctata728
Sanguirana_aurantipunctata_ELR
Sanguirana_igorota
Sanguirana_tipanan
Sanguirana_luzonensis
Sanguirana_acai
Sanguirana_acai232
Abavorana_luctuosa_sedim
Abavorana_luctuosa_kubah
Abavorana_luctuosa_sabah
Chalcorana_macrops
Chalcorana_mocquardi
Chalcorana_raniceps
Chalcorana_chalconota
Chalcorana_rufipes
Chalcorana_parvacola
Chalcorana_cf_eschatia_belatan
Chalcorana_eschatia_myan
Chalcorana_eschatia_thai
Chalcorana_cf_megalonesa_jerai
Chalcorana_cf_megalonesa_tioman
Chalcorana_megalonesa
Pulchrana_rawa
Pulchrana_laterimaculata
Pulchrana_baramica
Pulchrana_glandulosa
Pulchrana_banjarana
Pulchrana_centropeninsularis
Pulchrana_siberu
Pulchrana_picturata_penang
Pulchrana_signata_matang
Pulchrana_picturata_ZRC
Pulchrana_signata_sabah
Pulchrana_picturata_sabah
Pulchrana_signata_swak
Pulchrana_melanomenta
Pulchrana_grandocula
Pulchrana_similis
Pulchrana_mangyanum
Pulchrana_moellendorffi
Pulchrana_signata117
Pulchrana_signata_Bako
Continued ,,,
9.0
70
60
50
40
30
20
10
0


Hylarana_erythraea
Hylarana_macrodactyla
Hylarana_taipehensis
Humerana_humeralis
Humerana_lateralis
Humerana_miopus
Humerana_sp_myan
Sylvirana_guentheri_fujian
Sylvirana_guentheri_hanoi
Sylvirana_guentheri_viet
Sylvirana_guentheri002
Sylvirana_guentheri109
Sylvirana_attigua
Amnirana_nicobariensis938
Amnirana_nicobariensis773
Amnirana_nicobariensis_sumatra
Amnirana_nicobariensis020
Amnirana_nicobariensis_java
Amnirana_nicobariensis_sumatra2
Amnirana_nicobariensis_gasing
Amnirana_nicobariensis_sedim
Amnirana_nicobariensis_bario
Amnirana_nicobariensis177
Amnirana_nicobariensis_siberut
Indosylvirana_serendipi
Indosylvirana_urbis
Indosylvirana_doni
Indosylvirana_aurantiaca
Indosylvirana_intermedia
Indosylvirana_montana
Indosylvirana_temporalis
Indosylvirana_magna
Indosylvirana_flavescens
Indosylvirana_caesari
Indosylvirana_indica
Indosylvirana_sreeni
Amnirana_galamensis148
Amnirana_galamensis740
Amnirana_galamensis225
Amnirana_galamensis262
Amnirana_lepus987
Amnirana_lepus732
Amnirana_lepus220
Amnirana_lepus985
Amnirana_darlingi
Amnirana_albolabris991
Amnirana_albolabris861
Amnirana_albolabris863
Amnirana_albolabris177
Amnirana_albolabris664
Amnirana_amnicola818
Amnirana_amnicola606
Amnirana_amnicola621
Continued ..
9.0
60
50
40
30
20
10
0

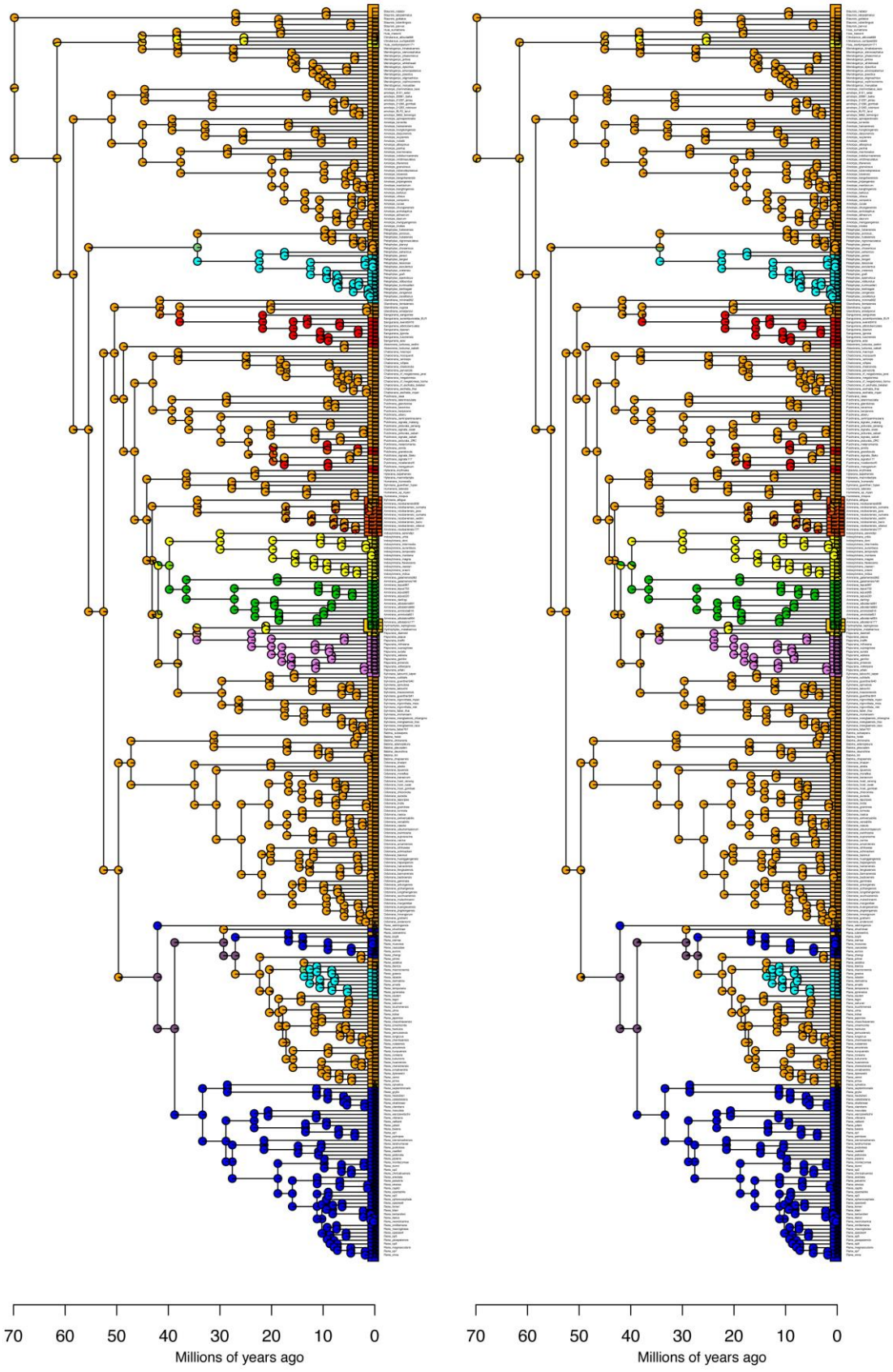

Hydrophylax_leptoglossa
Hydrophylax_bahuvistara
Hydrophylax_malabaricus
Papurana_daemeli
Papurana_kreffti
Papurana_papua
Papurana_milneana
Papurana_aurata
Papurana_supragrisea
Papurana_garritor
Papurana_waliesa
Papurana_jimiensis
Papurana_arfaki
Papurana_volkerjane
Sylvirana_cubitalis
Sylvirana_latouchii_taipei
Sylvirana_guentheri940
Sylvirana_guentheri941
Sylvirana_maosonensis
Sylvirana_spinulosa
Sylvirana_latouchii
Sylvirana_latouchii_anhui
Sylvirana_nigrovittata_myan
Sylvirana_nigrovittata_msia
Sylvirana_nigrovittata_viet
Sylvirana_nigrovittata_camb
Sylvirana_faber_thai
Sylvirana_faber_camb
Sylvirana_mortenseni
Sylvirana_menglaensis_chiangmai
Sylvirana_nigrovittata_thai
Sylvirana_menglaensis_thai
Sylvirana_menglaensis_laos
Sylvirana_faber767
Sylvirana_menglaensis_luangprabang
Babina_holsti
Babina_subaspera
Babina_adenopleura
Babina_okinavana
Babina_pleuraden
Babina_daunchina
Babina_chapaensis
Babina_lini
Odorrana_absita
Odorrana_khalam
Odorrana_lipuensis
Odorrana_swinhoana
Odorrana_utsunomiyaorum
Odorrana_supranarina
Odorrana_amamiensis
Odorrana_narina
Odorrana_tormota
Odorrana_nasica
Odorrana_exiliversabilis
Odorrana_nasuta
Odorrana_versabilis
Odorrana_banaorum
Odorrana_morafkai
Odorrana_hosii_ranong
Odorrana_hosii_gombak
Odorrana_hosii_swak
Odorrana_aureola
Odorrana_chloronota
Odorrana_leporipes
Odorrana_graminea
Odorrana_livida
Odorrana_ishikawae
Odorrana_schmackeri
Odorrana_huanggangensis
Odorrana_tianmuii
Odorrana_hejiangensis
Odorrana_bacboensis
Odorrana_tiannanensis
Odorrana_fengkaiensis
Odorrana_hainanensis
Odorrana_geminata
Odorrana_anlungensis
Odorrana_lungshengensis
Odorrana_yizhangensis
Odorrana_mutschmanni
Odorrana_wuchuanensis
Odorrana_kuangwuensis
Odorrana_margaretae
Odorrana_grahami
Odorrana_andersonii
Odorrana_hmongorum
Odorrana_jingdongensis
Continued ...
9.0
70
60
50
40
30
20
10
0


Rana_weiningensis
Rana_shuchinae
Rana_luteiventris
Rana_boylii
Rana_aurora
Rana_cascadae
Rana_muscosa
Rana_sierrae
Rana_johnsi
Rana_zhengi
Rana_asiatica
Rana_iberica
Rana_graeca
Rana_macrocnemis
Rana_dalmatina
Rana_latastei
Rana_arvalis
Rana_pyrenaica
Rana_temporaria
Rana_sauteri
Rana_sakuraii
Rana_tagoi
Rana_tsushimensis
Rana_kobai
Rana_ulma
Rana_chaochiaoensis
Rana_japonica
Rana_hanluica
Rana_omeimontis
Rana_jiemuxiensis
Rana_longicrus
Rana_culaiensis
Rana_zhenhaiensis
Rana_amurensis
Rana_coreana
Rana_kunyuensis
Rana_kukunoris
Rana_chensinensis
Rana_huanrensis
Rana_ornativentris
Rana_dybowskii
Rana_pirica
Rana_uenoi
Rana_sylvatica
Rana_septentrionalis
Rana_grylio
Rana_hecksheri
Rana_catesbeiana
Rana_clamitans
Rana_okaloosae
Rana_maculata
Rana_vibicaria
Rana_warszewitschii
Rana_juliani
Rana_vaillanti
Rana_bwana
Rana_palmipes
Rana_sp1
Rana_sierramadrensis
Rana_psilonota
Rana_zweifeli
Rana_pustulosa
Rana_tarahumarae
Rana_pipiens
Rana_chiricahuensis
Rana_sp2
Rana_dunni
Rana_montezumae
Rana_areolata
Rana_palustris
Rana_capito
Rana_sevosa
Rana_sp3
Rana_spectabilis
Rana_sphenocephala
Rana_forreri
Rana_species6
Rana_blairi
Rana_berlandieri
Rana_neovolcanica
Rana_tlaloci
Rana_macroglossa
Rana_omiltemana
Rana_sp5
Rana_species4
Rana_yavapaiensis
Rana_sp8
Rana_magnaocularis
Rana_onca
Rana_sp7
9.0
70
60
50
40
30
20
10
0

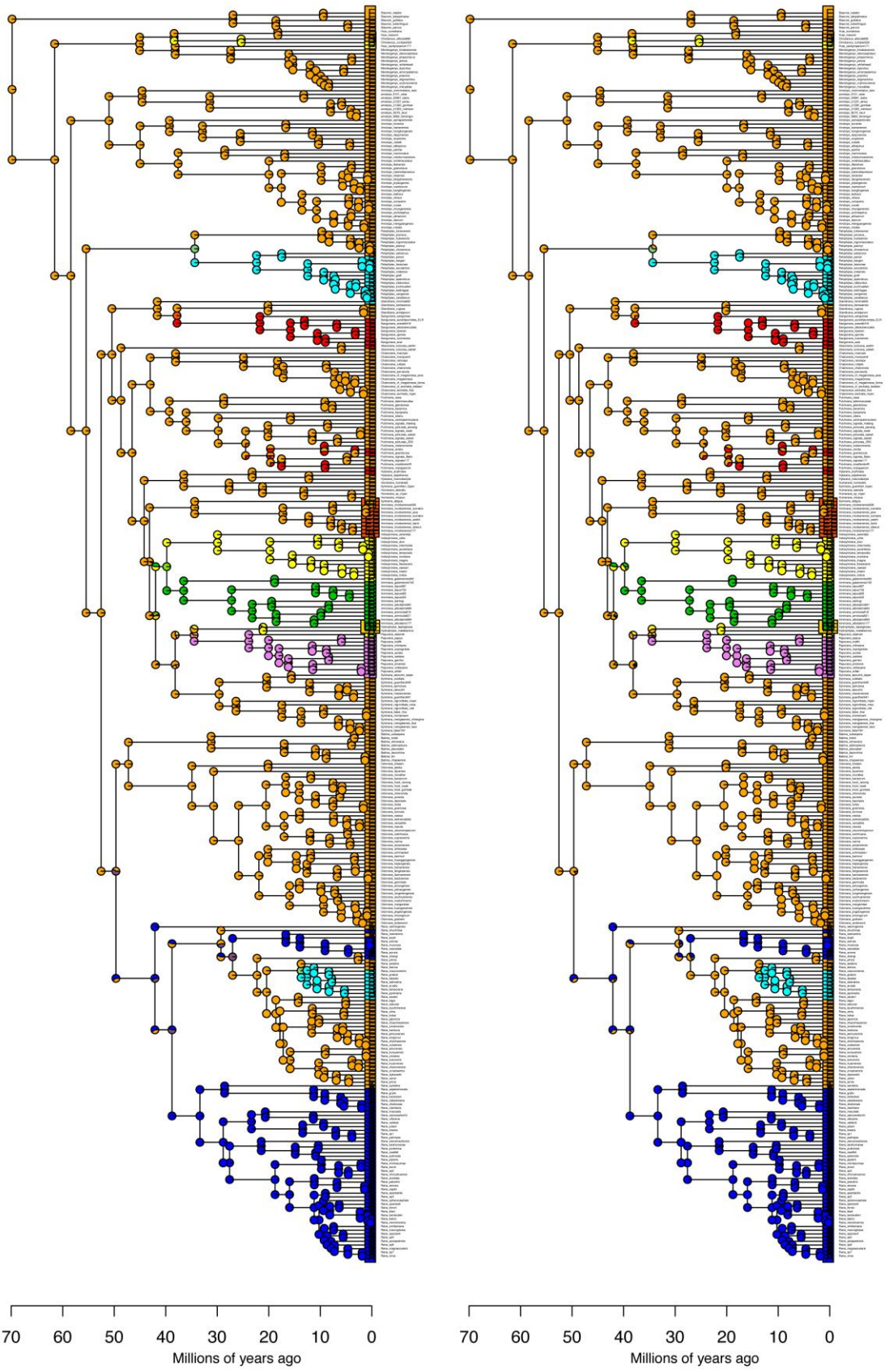
Appendix Fig. 1. Time-calibrated maximum clade credibility Bayesian phylogeny of Ranidae with corresponding posterior probabilities and 95% HPD age intervals. Posterior probabilities are scaled to size and color.


70
60
50
40
30
20
10
0
70
60
50
40
30
20
10
0

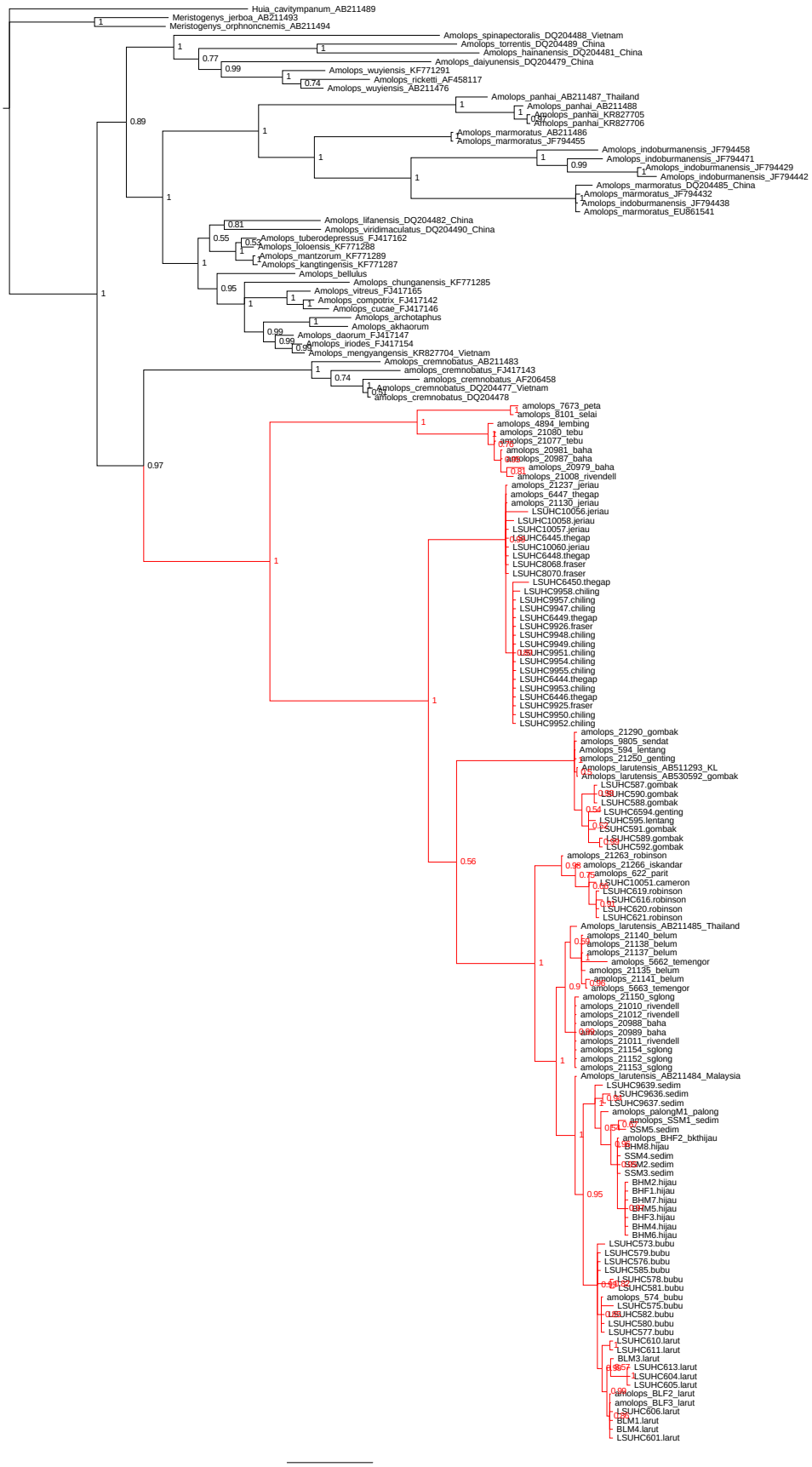
Appendix Fig. 2. Unconstrained BioGeoBEARS analysis using the BAYAREALIKE model. Left: without the *j* parameter (LnL = -155.72); right: with the *j* parameter (LnL = -99.48).


70
60
50
40
30
20
10
0
Millions of years ago
70
60
50
40
30
20
10
0
Millions of years ago

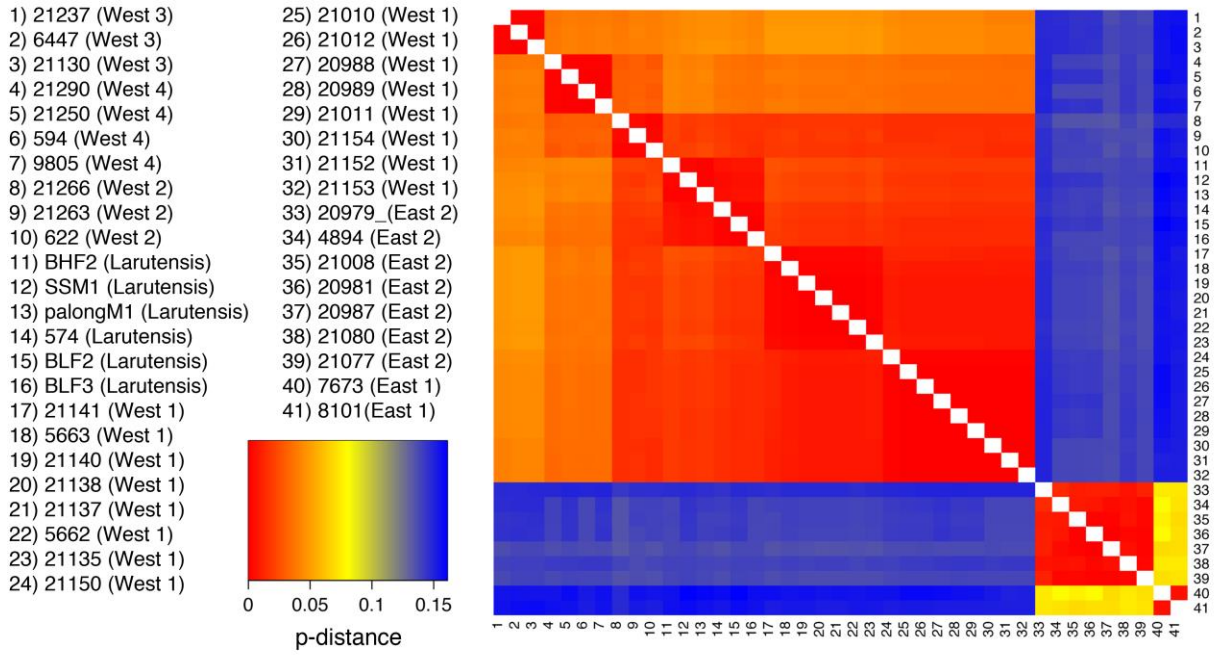
Appendix Fig. 3. Unconstrained BioGeoBEARS analysis using the DEC model. Left: without the *j* parameter (LnL = -147.6); right: with the *j* parameter (LnL = -141.89).


70
60
50
40
30
20
10
0
Millions of years ago
70
60
50
40
30
20
10
0
Millions of years ago

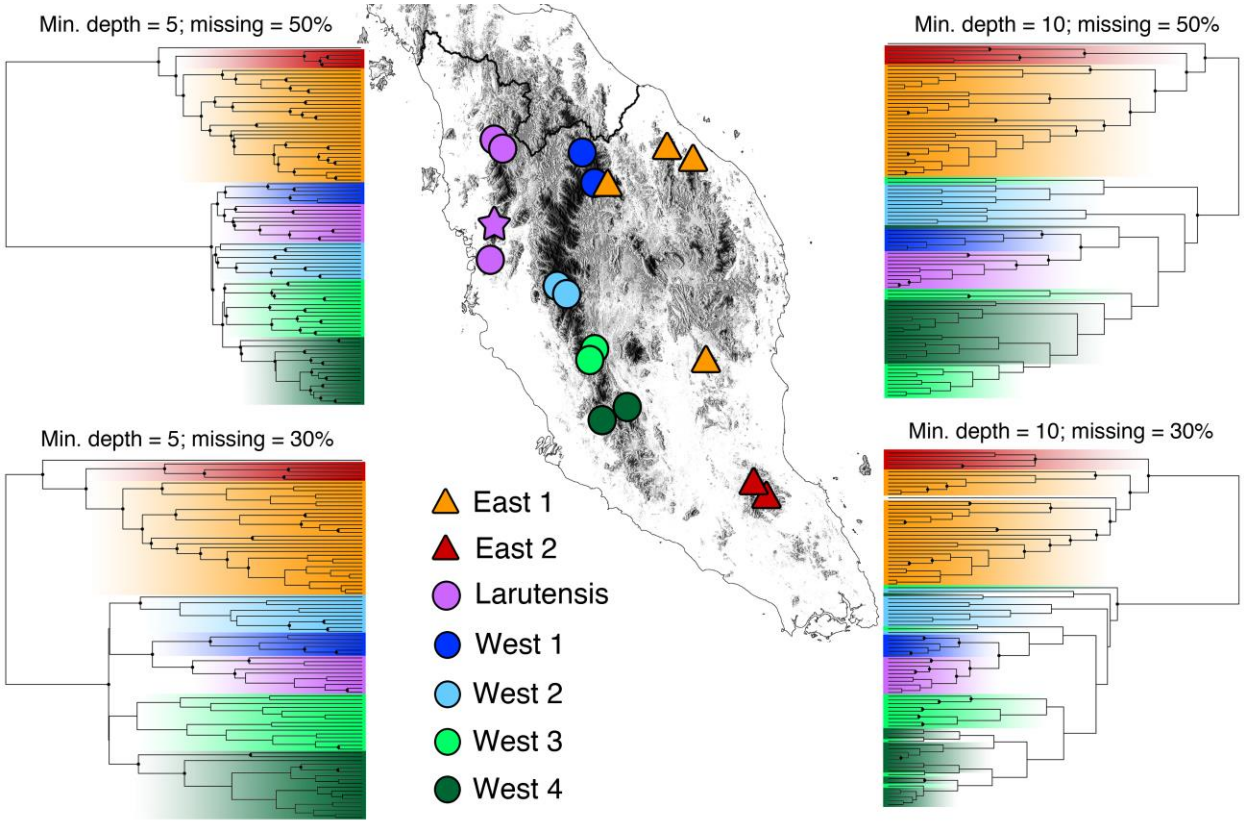
Appendix Fig. 4. Unconstrained BioGeoBEARS analysis using the DIVALIKE model. Left: without the *j* parameter (LnL = -157.87); right: with the *j* parameter (LnL = -152.31).


Huia_cavitympanum_AB211489
Meristogenys_jerboa_AB211493
Meristogenys_orphnocnemis_AB211494
Amolops_spinapectoralis_DQ204488_Vietnam
Amolops_torrentis_DQ204489_China
Amolops_hainanensis_DQ204481_China
Amolops_daiyunensis_DQ204479_China
Amolops_wuyiensis_KF771291
Amolops_ricketti_AF458117
Amolops_wuyiensis_AB211476
Amolops_panhai_AB211487_Thailand
Amolops_panhai_AB211488
Amolops_panhai_KR827705
Amolops_panhai_KR827706
Amolops_marmoratus_AB211486
Amolops_marmoratus_JF794455
Amolops_indoburmanensis_JF794458
Amolops_indoburmanensis_JF794471
Amolops_indoburmanensis_JF794429
Amolops_indoburmanensis_JF794442
Amolops_marmoratus_DQ204485_China
Amolops_marmoratus_JF794432
Amolops_indoburmanensis_JF794438
Amolops_marmoratus_EU861541
Amolops_lifanensis_DQ204482_China
Amolops_viridimaculatus_DQ204490_China
Amolops_tuberodepressus_FJ417162
Amolops_loloensis_KF771288
Amolops_mantzorum_KF771289
Amolops_kangtingensis_KF771287
Amolops_bellulus
Amolops_chunganensis_KF771285
Amolops_vitreus_FJ417165
Amolops_compotrix_FJ417142
Amolops_cucae_FJ417146
Amolops_archotaphus
Amolops_akhaorum
Amolops_daorum_FJ417147
Amolops_iriodes_FJ417154
Amolops_mengyangensis_KR827704_Vietnam
Amolops_cremnobatus_AB211483
amolops_cremnobatus_FJ417143
amolops_cremnobatus_AF206458
Amolops_cremnobatus_DQ204477_Vietnam
amolops_cremnobatus_DQ204478
amolops_7673_peta
amolops_8101_selai
amolops_4894_lembing
amolops_21080_tebu
amolops_21077_tebu
amolops_20981_baha
amolops_20987_baha
amolops_20979_baha
amolops_21008_rivendell
amolops_21237_jeriau
amolops_6447_thegap
amolops_21130_jeriau
LSUHC10056.jeriau
LSUHC10058.jeriau
LSUHC10057.jeriau
LSUHC6445.thegap
LSUHC10060.jeriau
LSUHC6448.thegap
LSUHC8068.fraser
LSUHC8070.fraser
LSUHC6450.thegap
LSUHC9958.chiling
LSUHC9957.chiling
LSUHC9947.chiling
LSUHC6449.thegap
LSUHC9926.fraser
LSUHC9948.chiling
LSUHC9949.chiling
LSUHC9951.chiling
LSUHC9954.chiling
LSUHC9955.chiling
LSUHC6444.thegap
LSUHC9953.chiling
LSUHC5446.thegap
LSUHC9925.fraser
LSUHC9950.chiling
LSUHC9952.chiling
amolops_21290_gombak
amolops_9805_sendat
Amolops_594_lentang
amolops_21250_genting
Amolops_larutensis_AB511293_KL
Amolops_larutensis_AB530592_gombak
LSUHC587.gombak
LSUHC590.gombak
LSUHC588.gombak
LSUHC6594.genting
LSUHC595.lentang
LSUHC591.gombak
LSUHC589.gombak
LSUHC592.gombak
amolops_21263_robinson
amolops_21266_iskandar
amolops_622_parit
LSUHC10051.cameron
LSUHC619.robinson
LSUHC616.robinson
LSUHC620.robinson
LSUHC621.robinson
Amolops_larutensis_AB211485_Thailand
amolops_21140_belum
amolops_21138_belum
amolops_21137_belum
amolops_5662_temengor
amolops_21135_belum
amolops_21141_belum
amolops_5663_temengor
amolops_21150_sglong
amolops_21010_rivendell
amolops_21012_rivendell
amolops_20988_baha
amolops_20989_baha
amolops_21011_rivendell
amolops_21154_sglong
amolops_21152_sglong
amolops_21153_sglong
Amolops_larutensis_AB211484_Malaysia
LSUHC9639.sedim
LSUHC9636.sedim
LSUHC9637.sedim
amolops_palongM1_palong
amolops_SSM1_sedim
SSM5.sedim
amolops_BHF2_bkthijau
BHM8.hijau
SSM4.sedim
SSM2.sedim
SSM3.sedim
BHM2.hijau
BHF1.hijau
BHM7.hijau
BHM5.hijau
BHF3.hijau
BHM4.hijau
BHM6.hijau
LSUHC573.bubu
LSUHC579.bubu
LSUHC576.bubu
LSUHC585.bubu
LSUHC578.bubu
LSUHC581.bubu
amolops_574_bubu
LSUHC575.bubu
LSUHC582.bubu
LSUHC580.bubu
LSUHC577.bubu
LSUHC610.larut
LSUHC611.larut
BLM3.larut
LSUHC613.larut
LSUHC604.larut
LSUHC605.larut
amolops_BLF2_larut
amolops_BLF3_larut
LSUHC606.larut
BLM1.larut
BLM4.larut
LSUHC601.larut
0.03

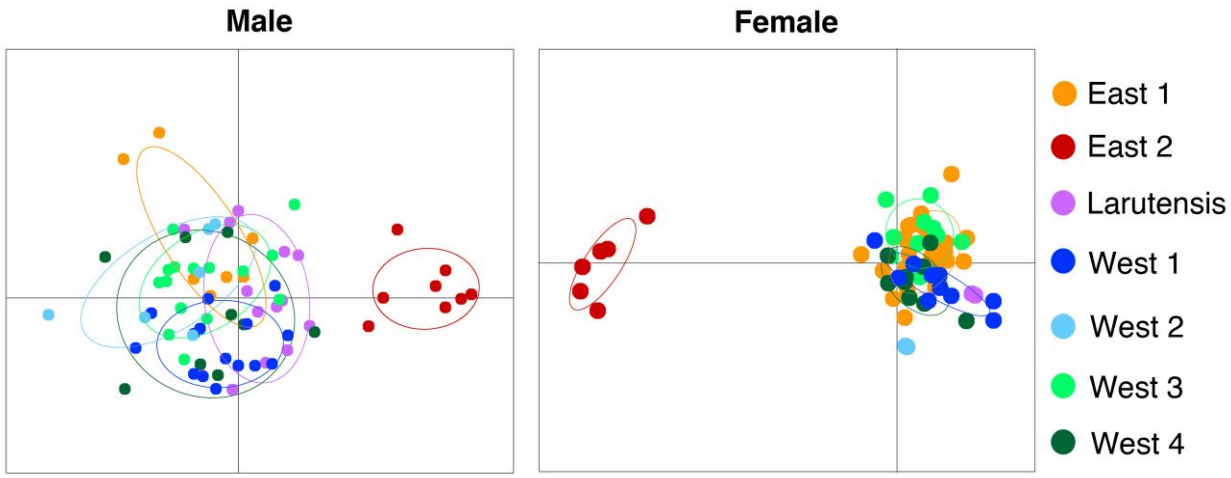
Appendix Fig. 5. Bayesian phylogeny of *Amolops* inferred from 1,466 base pairs of the 16S rRNA mitochondrial gene. Node support represent posterior probabilities and the Peninsular Malaysian clade is highlighted in red.



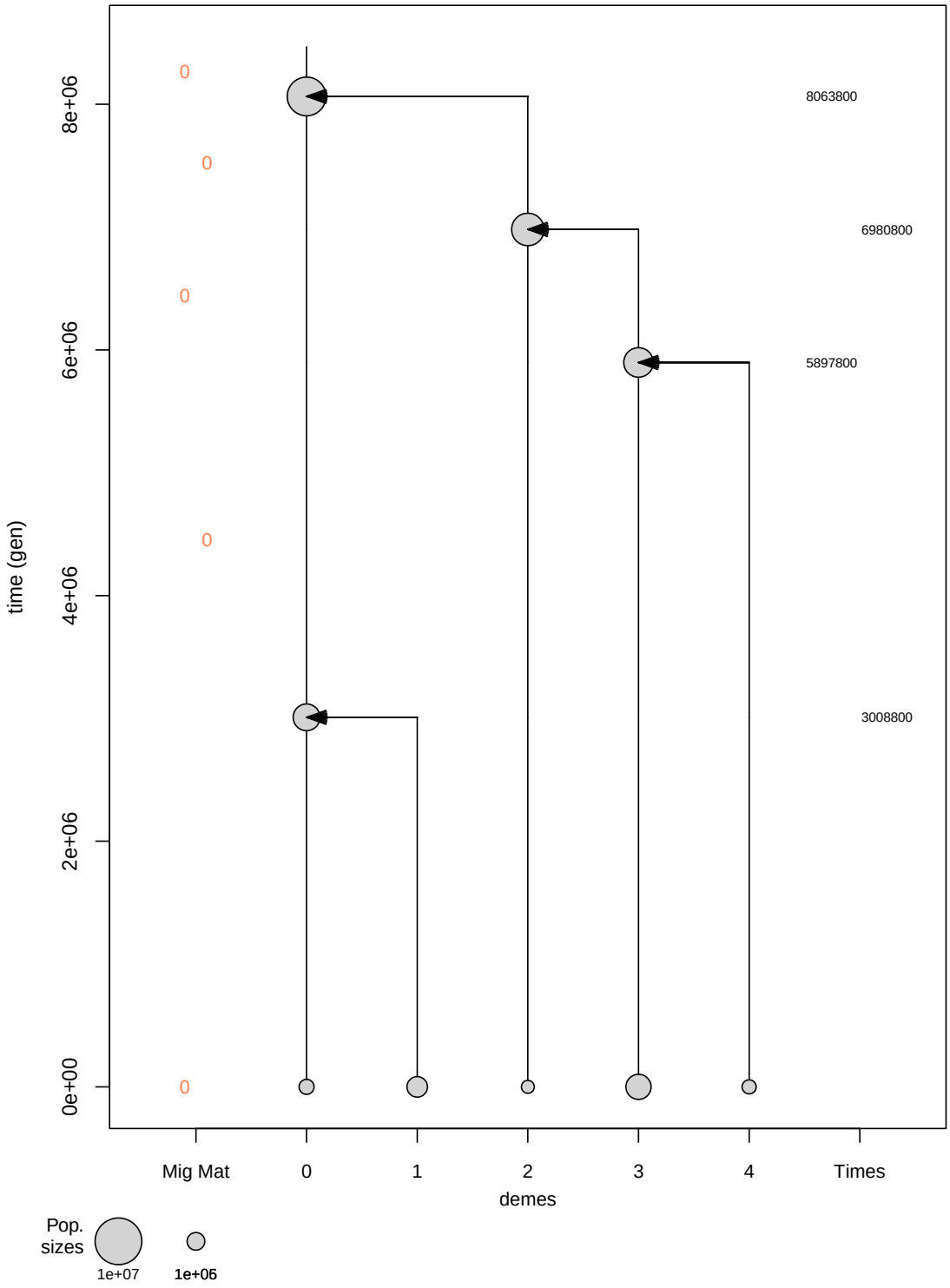
Appendix Fig. 6. Heatmap of pairwise uncorrected mitochondrial p-distances calculated from the 16S gene.



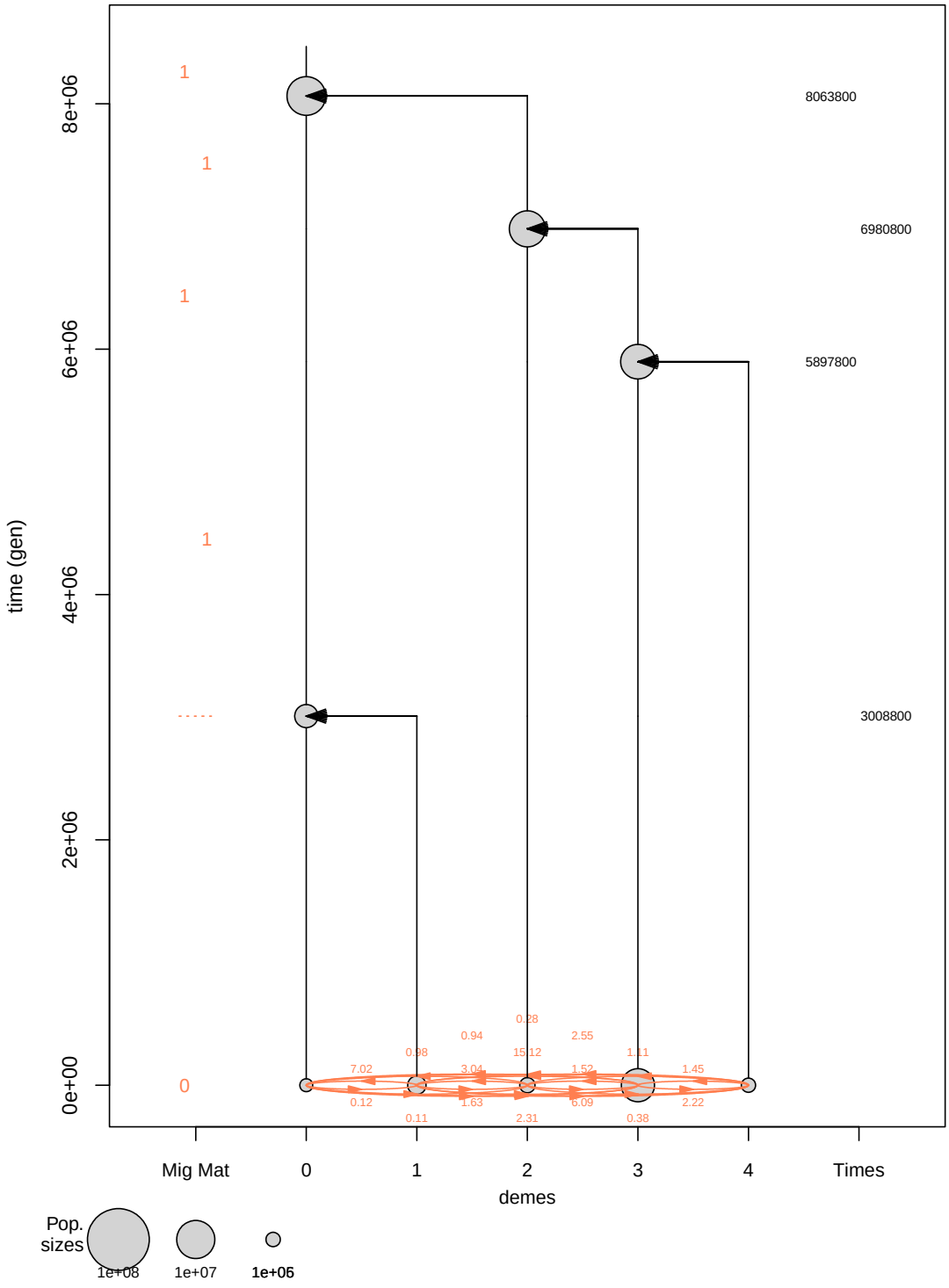
Appendix Fig. 7. Maximum likelihood phylogenies inferred from genome-wide SNPs filtered at different minimum depths and missing data. Black dots represent nodes with >90% bootstrap values.



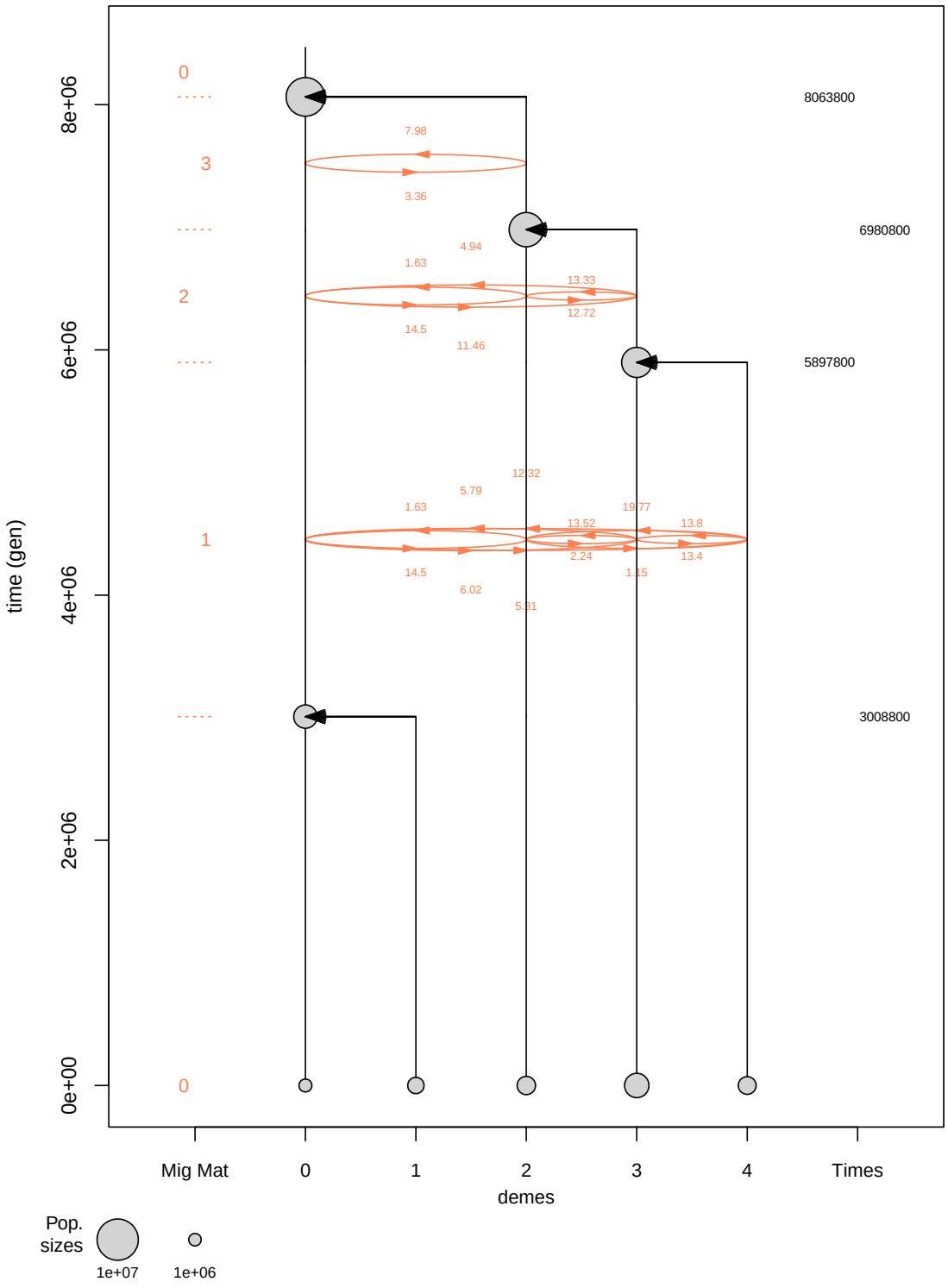
Appendix Fig. 8. Results of the DAPC analysis on adjusted morphological data.



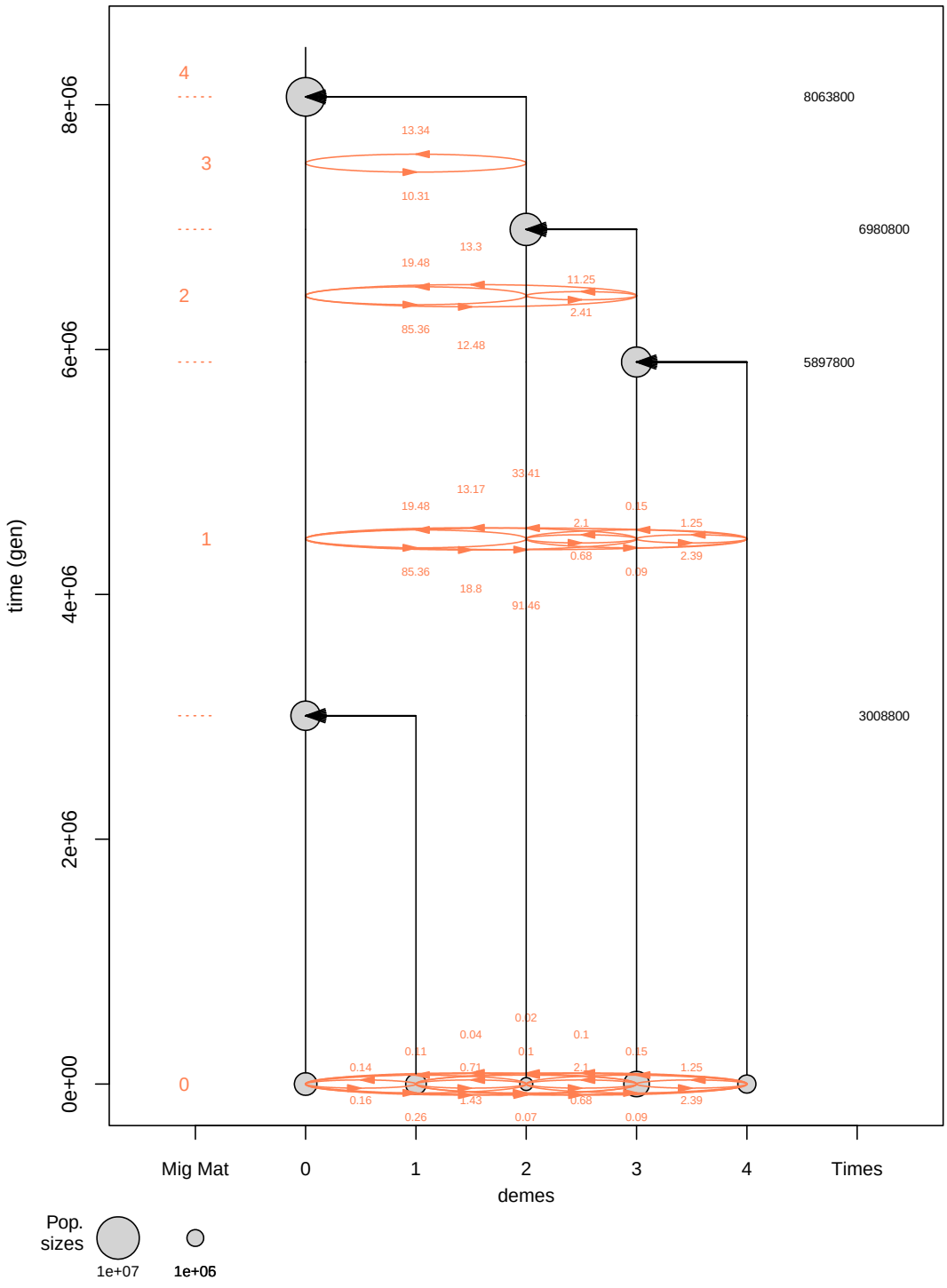
Appendix Fig. 9. Fastsimcoal2 model with no migration



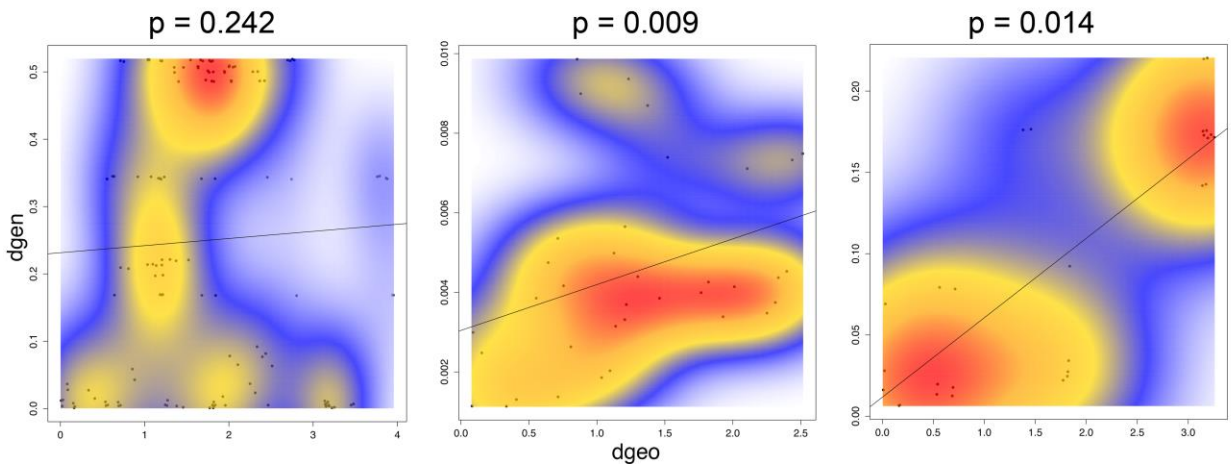
Appendix Fig. 10. Fastsimcoal2 model with contemporary migration defined as the time period from the present up to the first diversification even at ~3 mya.



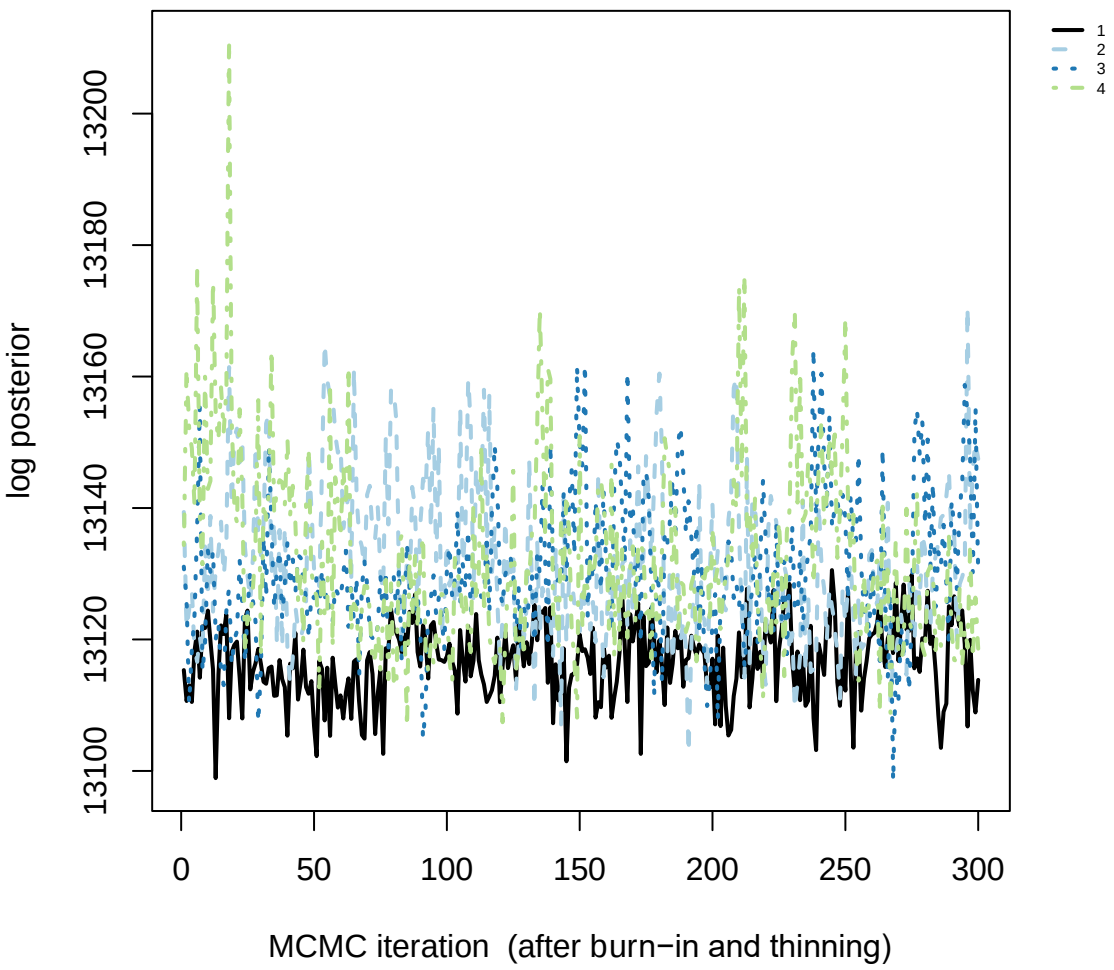
Appendix Fig. 11. Fastsimcoal2 model with historical migration only including three different historical time periods (see materials and methods for more information).



Appendix Fig. 12. Fastsimcoal2 model that accommodated both contemporary and historical migrations. This full model was tested twice, once where migrations were constrained to be symmetrical and the other with asymmetrical migrations.



Appendix Fig. 13. Kernel density estimation plots between genetic and geographic distances with corresponding P-values from the Mantel test. Left: entire dataset; middle: western clade only; right: eastern clade only.



Appendix Fig. 14. Traceplots of the four independent MCMC chains from the EEMS analysis at 300 demes.